Systems-level characterization and reconstitution of genetic regulatory networks in Escherichia coli for rational strain optimization : towards improved bioprocesses

Kang, A Ram

2012

Kang, A. R. (2012). Systems-level characterization and reconstitution of genetic regulatory networks in Escherichia coli for rational strain optimization : towards improved bioprocesses. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50704>

<https://doi.org/10.32657/10356/50704>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**SYSTEMS-LEVEL CHARACTERIZATION AND
RECONSTITUTION OF GENETIC REGULATORY NETWORKS
IN *ESCHERICHIA COLI* FOR RATIONAL STRAIN
OPTIMIZATION: TOWARDS IMPROVED BIOPROCESSES**

A RAM KANG

SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING

2012

**SYSTEMS-LEVEL CHARACTERIZATION AND
RECONSTITUTION OF GENETIC REGULATORY NETWORKS
IN *ESCHERICHIA COLI* FOR RATIONAL STRAIN
OPTIMIZATION: TOWARDS IMPROVED BIOPROCESSES**

KANG A.

A RAM KANG

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to Lord, who sincerely leads me all the way now and forever. I believe that this thesis would not have been possible without the grace of God and the strength He has put in me. Also, I would like to give my utmost gratitude to my parents for their love and endless prayers. My parents, especially my dad, supported my decision to pursue a Ph. D degree in Singapore when he was diagnosed with lung cancer and just started receiving chemotherapy for cancer. My mother and younger brother went through hard time with him fighting against his illness in Korea while I was away in Singapore.

It is impossible for me to overemphasize my gratitude to my supervisor, Prof. Matthew Wook Chang. He is a great researcher and a mentor whom I would like to model myself after. I sincerely appreciate all inspiring instructions, kind encouragement and untiring supports that he has provided to me. I was lucky to join his lab, and I believe that I had many opportunities to learn and experience various researches under his supervision. Also, I thank my lab colleagues. Dr. Ling Hua helped me to learn basic lab skills and gave many practical advices on troubleshooting of experiments. We equally contributed to the works presented in Chapter 4. I would like to give my thanks to Mui Hua, Binbin, Sam and Simon. They not only assisted me in doing experiments, but also helped me enjoy my Ph.D. in Singapore. We had a lot of fun moments together, and their great sense of humor and fine care helped me recover my strength to continue to do research when I was exhausted. I thank to Inyoung for her help to read this thesis and give thoughtful comments and suggestions to me. Her help cleared up ambiguity and polished sentences of this thesis. I also

appreciate all technical and administrative staffs in CBE, especially Dr. Ong and Woan Ting for their sincere supports.

Lastly, all studies done in this thesis could not be done without financial supports from Nanyang Technological University and the following funding agencies: Chapter 4 by the Ministry of Education of Singapore (AcRF RG31/07) and the National Medical Research Council of Singapore (NIG07NOV035), and Chapters 5 and 6 by the Ministry of Education of Singapore (AcRF RG31/07) and the National Research Foundation of Singapore (NRF-CRP5-2009-03).

LIST OF PUBLICATIONS

Kang A, Tan MH, Ling H and Chang MW. "Systems-level oxidative stress response of *Escherichia coli* under anaerobic condition and essential transcription factors." under review

Kang A and Chang MW (2012) "Identification and reconstitution of genetic regulatory networks for improved microbial tolerance to isooctane." Molecular BioSystems **8**(4): 1350-1358.

Ling H, **Kang A** (co-first author), Tan MH, Qiao B and Chang MW (2010) "The absence of the *luxS* gene increases swimming motility and flagella synthesis in *Escherichia coli* K12." Biochemical Biophysical Research Communications **401**(4): 521-526.

Choi YH, Park CH, Kim W, Ling H, **Kang A**, Chang MW, Im SK, Jeong HW, Kong YY, Kim KT (2010) "Vaccinia-related kinase 1 is required for the maintenance of undifferentiated spermatogonia in mouse male germ cells." PloS One **5**(12): e15254.

Qiao Y, Li CM, Lu Z, Ling H, **Kang A** and Chang MW (2009) "A time-course transcriptome analysis of *Escherichia coli* with direct electrochemistry behavior in microbial fuel cells." Chemical Communications (41): 6183-6185.

TABLE OF CONTENTS

ACKNOWLEDGEMENT.....	1
LIST OF PUBLICATIONS.....	3
TABLE OF CONTENTS.....	4
LIST OF FIGURES.....	8
LIST OF TABLES	11
SUMMARY	13
CHAPTER 1 INTRODUCTION	14
1.1 BACKGROUNDS	14
1.2 OBJECTIVES AND SCOPE.....	16
1.3 SIGNIFICANCES	19
1.4 THESIS OUTLINE	20
CHAPTER 2 LITERATURE REVIEW.....	22
2.1 BIOTECHNOLOGY AND BIOPROCESS.....	22
2.2 <i>ESCHERICHIA COLI (E. COLI)</i> AS A HOST IN BIOPROCESSES	22
2.2.1 <i>E. coli</i> physiology.....	22
2.2.2 <i>E. coli</i> genomics and genetic manipulations.....	23
2.2.3 <i>Limitations and improvement of E. coli strains for bioprocesses</i>	24
2.2.3.1 Improper protein folding.....	25
2.2.3.2 Biased codon usage	26
2.2.3.3 Optimization of <i>E. coli</i> strains to improve bioprocess: group behavior, stress responses and tolerance	26
2.3 QUORUM SENSING AND BIOPROCESS	28
2.3.1 <i>Quorum sensing and autoinducers</i>	28
2.3.2 <i>AI-2 Synthesis in E. coli</i>	31
2.3.3 <i>Quorum sensing in E. coli</i>	33

2.3.4	<i>Quorum sensing in various bacterial phenotypes.....</i>	35
2.4	OXIDATIVE STRESS RESPONSES.....	41
2.4.1	<i>Oxidative stress and reactive oxygen species (ROS)</i>	41
2.4.2	<i>Oxidative stress and bioprocess.....</i>	44
2.4.2.1	Endogenous production of ROS and respiration	44
2.4.2.2	Oxidative stress and bacterial cell death.....	45
2.4.2.3	Oxidative stress and organic solvents in bioprocess	47
2.4.2.4	Anaerobic environments and oxidative stress	48
2.4.3	<i>Cellular responses to oxidative stress in E. coli.....</i>	50
2.5	HYDROCARBON STRESS RESPONSES.....	54
2.5.1	<i>Biofuel production.....</i>	54
2.5.2	<i>Isooctane (2,2,4-Trimethylpentane, C₈H₁₈)</i>	55
2.5.3	<i>Toxicity of organic solvents.....</i>	55
2.5.4	<i>Defense mechanisms against organic solvents.....</i>	58
CHAPTER 3	METHODOLOGY AND APPROACHES	60
3.1	SYSTEMS BIOLOGY	60
3.2	MICROARRAY, FOLD CHANGE ANALYSIS AND GO ANALYSIS.....	63
3.2.1	<i>Microarray experimental methods</i>	63
3.2.2	<i>Fold Change analysis.....</i>	64
3.2.3	<i>Gene Ontology (GO) Term analysis</i>	65
3.3	TRANSCRIPTIONAL REGULATIONS AND ITS ANALYSIS.....	66
3.3.1	<i>Promoter and RNA polymerase</i>	66
3.3.2	<i>Role and action of transcription factors in prokaryotes</i>	68
3.3.3	<i>Hierarchy of transcriptional regulation by TF</i>	69
3.3.4	<i>Network component analysis</i>	71
3.4	PHYSIOLOGICAL TRAITS ANALYSIS	72

CHAPTER 4 THE ABSENCE OF THE *LUXS* GENE INCREASES SWIMMING MOTILITY AND FLAGELLA SYNTHESIS IN *ESCHERICHIA COLI* K12.....74

4.1	ABSTRACT	74
4.2	INTRODUCTION.....	74
4.3	MATERIALS AND METHODS.....	75
4.3.1	<i>Bacterial strains, media, and growth conditions.....</i>	75
4.3.2	<i>Motility assay.....</i>	76
4.3.3	<i>Microarray and real-time PCR assays</i>	76
4.3.4	<i>Dot blotting of flagellin and flagella visualization</i>	77
4.4	RESULTS AND DISCUSSION.....	78
4.5	CONCLUSIONS	95

CHAPTER 5 SYSTEMS-LEVEL OXIDATIVE STRESS RESPONSE OF *ESCHERICHIA COLI* UNDER ANAEROBIC CONDITION: IDENTIFICATION OF ESSENTIAL REGULATORY FACTORS 96

5.1	ABSTRACT	96
5.2	INTRODUCTION.....	97
5.3	MATERIALS AND METHODS.....	98
5.3.1	<i>Strains, growth condition and H₂O₂ treatment</i>	98
5.3.2	<i>RNA extraction and microarray experiments</i>	99
5.3.3	<i>Data analysis.....</i>	99
5.3.4	<i>NADH and NAD⁺ quantification.....</i>	100
5.4	RESULTS AND DISCUSSION.....	101
5.4.1	<i>Systems-level cellular response to oxidative stress under anaerobic condition</i>	101
5.4.1.1	Class of iron ion transport.....	109
5.4.1.2	Class of nitrogen metabolism and pyrimidine metabolism	111
5.4.1.3	Class of TCA cycle and cellular respiration	116

5.4.2	<i>Transcription factors the activity of which was perturbed by oxidative stress under anaerobic condition.....</i>	124
5.4.3	<i>NADH: NAD⁺ ratio and oxidative stress.....</i>	128
5.5	CONCLUSIONS.....	131
CHAPTER 6 IDENTIFICATION AND RECONSTITUTION OF GENETIC REGULATORY NETWORKS FOR IMPROVED MICROBIAL TOLERANCE TO ISOCTANE.....		132
6.1	ABSTRACT.....	132
6.2	INTRODUCTION	133
6.3	MATERIALS AND METHODS.....	135
6.3.1	<i>Strains, growth conditions and isooctane treatment.....</i>	135
6.3.2	<i>Knock-out mutant construction, gene overexpression, and viability testing</i>	138p
6.3.3	<i>Total RNA extraction and microarray analysis.....</i>	139
6.3.4	<i>Transcriptome data processing and analysis.....</i>	139
6.3.5	<i>Transmission electron microscopy (TEM).....</i>	140
6.4	RESULTS AND DISCUSSION	140
6.4.1	<i>Essential genetic determinants underlying cellular response to isooctane....</i>	140
6.4.2	<i>Functional transcriptome analysis and corresponding physiological and metabolic responses</i>	148
6.4.3	<i>Reconstitution of genetic regulatory networks of isooctane response by knocking out and overexpressing related transcription regulators.....</i>	157
6.5	CONCLUSIONS.....	164
CHAPTER 7 CONCLUSIONS AND FUTURE CONSIDERATIONS		166
BIBLIOGRAPHY		173

LIST OF FIGURES

Figure 1-1 A schematic diagram of this thesis.....	18
Figure 2-1 Chemical structures of diverse autoinducers.....	30
Figure 2-2 Activated methyl cycle (AMC) and LuxS	32
Figure 2-3 Schematic diagram of AI-2 quorum sensing in <i>E. coli</i>	34
Figure 2-4 Extracellular and intracellular routes of ROS production in microorganisms	43
Figure 3-1 Schematic diagram of the promoter region.....	67
Figure 3-2 Graphical presentations of transcriptional regulatory networks (Balaji, Babu <i>et al.</i> 2007).....	70
Figure 4-1 Swimming and swarming motility of WT and <i>luxS</i> mutants	80
Figure 4-2 Functional classification of differentially expressed genes in <i>luxS</i> mutants and a schematic representation of gene regulation and flagella structure	86
Figure 4-3 Flagella visualized by Field Emission Scanning Electron Microscopy (FESEM).....	89
Figure 4-4 Conceptual cellular response networks related to flagella synthesis in <i>luxS</i> mutants.....	94
Figure 5-1 Heat maps of expression levels of genes involved in defense mechanisms against oxidative stress	106

Figure 5-2 Heat maps (A) of relative expression levels and metabolic pathways (B) of the genes in the GO term “alanine, aspartate and glutamate metabolism (ecj000250)”	112
Figure 5-3 Heat maps of relative expression levels and metabolic pathways of genes involved in the TCA cycle	117
Figure 5-4 Scatter plots and heat maps of genes involved in cellular respiration	123
Figure 5-5 Relative viability of the wild type (W3110, WT) and <i>arcA</i> -, <i>arcB</i> -, <i>fnr</i> -deleted isogenic mutants upon H ₂ O ₂ treatment under anaerobic (A) or aerobic condition (B)	126
Figure 5-6 NADH: NAD ⁺ ratio without (empty) or with H ₂ O ₂ (filled) under aerobic and anaerobic condition (A). Effect of FDH1 expression on relative viability without (empty) or with H ₂ O ₂ (filled) under aerobic (B) or anaerobic condition (C)	129
Figure 6-1 Temporal transcriptome responses of <i>E. coli</i> to isooctane.....	145
Figure 6-2 Essentiality of various stress mechanisms in isooctane-tolerance of <i>E. coli</i>	147
Figure 6-3 Clustering and functional annotation analysis of significantly induced or repressed genes	149
Figure 6-4 Transmission electron microscopy images of <i>E. coli</i> incubated without (a and b) or with (c and d) isooctane.....	156
Figure 6-5 Relative viability of single transcription factor-deleted mutants in the presence of isooctane.	159

Figure 6-6 Effects of overexpression of the transcription factors whose activity was perturbed on <i>E. coli</i> viability in the presence of isooctane	160
Figure 6-7 Relative viability of <i>E. coli</i> lacking two transcription factor genes in the presence of isooctane.	163
Figure 6-8 Conceptual genetic regulatory networks in isooctane stress responses	165

LIST OF TABLES

Table 2-1 Effects of <i>luxS</i> -deletion on bacterial phenotypic traits.....	37
Table 2-2 Toxicity levels for next-generation biofuels (Dunlop 2011).....	57
Table 4-1 List of genes and specific primers used for RT-PCR.....	77
Table 4-2 Significantly up- or down-regulated genes in <i>luxS</i> mutants.....	82
Table 4-3 Relative transcription factor (TF) activities of <i>luxS</i> mutants (L) and <i>E. coli</i> K12 wild type (W)	91
Table 5-1 List of genes regulated by OxyR, SoxRS or genes involved in SOS response, of which transcription was significantly changed by H ₂ O ₂ under anaerobic condition ($p < 0.05$, Fold change > 2).....	102
Table 5-2 GO terms and KEGG pathways analyzed by GO analysis from GeneSpring GX and DAVID functional annotation analysis.	108
Table 5-3 List of genes in “iron ion transport (GO:0006826 and GO:0015681)” with fold changes (FC) by H ₂ O ₂ under anaerobic condition.	110
Table 5-4 List of genes in “purine metabolism (ecj00230)” and “pyrimidine metabolism (ecj00240)” with fold changes (FC) by H ₂ O ₂ under aerobic or anaerobic condition	114
Table 5-5 List of genes in “Cellular respiration (GO: 0045333)” with fold changes (FC) by H ₂ O ₂ under aerobic or anaerobic condition.....	118
Table 6-1 A list of bacterial strains, plasmids and primers used in Chapter 6.	136

Table 6-2 Top 10 genes out of 225 significantly up-regulated ($p < 0.05$, fold change > 2) genes at 10min after isooctane treatment	142
Table 6-3 Top 10 genes out of 696 significantly up-regulated ($p < 0.05$, fold change > 2) genes at 60 min after isooctane treatment	143
Table 6-4 Annotation clusters and associated terms of genes of Clusters 0 and 1.....	151
Table 6-5 Significantly up- or down-regulated genes related to membrane and surface structure and properties.....	152

SUMMARY

Biological processes have demonstrated the potential to offer cost-effective and environment-friendly strategies for a wide range of applications such as food and chemical production, material synthesis, wastewater treatment, and environmental remediation. Despite this potential, biological processes frequently suffer from low productivity and stability stemming from the inadequate phenotypic characteristics of the cells involved. This limitation in biological processes has so far been tackled through various strain optimization strategies, which mostly require time-consuming, labor-intensive selection and screening methods. In this light, with recent advances in systems biology, which offers a holistic approach to understanding cells, the work presented in this thesis aimed to develop systems-level, rational strain engineering framework that could complement current optimization techniques. Towards this aim, cellular responses were elucidated at the systems level, and essential cellular functions and regulatory factors were characterized during the response of *Escherichia coli* to three important conditions in bioprocesses: (i) genetic disruption in quorum sensing, (ii) oxidative stress under anaerobic condition, and (iii) hydrocarbon solvent stress. Genetic regulatory networks underlying the respective responses were identified based on bioinformatics analysis of gene-transcription factor relationships, and these genetic networks were reconstituted to obtain desirable phenotypic characteristics of *E. coli*. In particular, it was shown that this strain-engineering framework, driven by re-wiring of regulatory components, significantly improved *E. coli* tolerance against oxidative stress and hydrocarbon stress. Taken together, this thesis demonstrates the potential of the systems-level cellular engineering framework developed in this work to design rational strain optimization strategies for improved bioprocesses, and to construct synthetic microbes that exhibit other desired phenotypic characteristics.

CHAPTER 1 INTRODUCTION

1.1 Backgrounds

Microbial organisms have been a subject of numerous studies to develop various bioprocesses such as food production and valuable chemical synthesis. Initially, microbiologists or bioprocess engineers were able to identify only a small number of genes related to microbial physiology. However, genome-sequencing projects provided sequences of genes and intergenic regions, which drastically extended our knowledge on bacterial cell morphology, metabolisms and genetic regulations. Based on the available DNA sequences, *in vitro* gene cloning and protein expression have enabled scientists to discover useful enzymatic proteins that perform desired chemical reactions, and to characterize their enzyme kinetics activities. Furthermore, based on bioinformatics approaches, scientists and engineers can now obtain information on biological functions of genes and proteins by studying structural homology of proteins across different species and organisms.

While information on single genes or proteins is more available, molecular mechanisms that enable multiple genes and proteins to work in concert with each other to exert biological functions have recently become of great interest. With progress of high-throughput biotechnology, it has been possible for researchers to investigate a number of genes, proteins and metabolites simultaneously in different experimental settings. Among these high-throughput technologies, DNA microarrays present important technological advances because they can readily be used to quantify transcriptomic level, and to study how living organisms develop dynamic profiles of transcripts in response to various environmental conditions. For this reason, transcriptomic profiling has been exploited to find multiple genes and enzymes

responsible for important cellular physiologies and useful metabolic pathways, of which characterization has been restricted due to technical difficulties.

The rationale behind this approach is that genes of which expression level increases or decreases are essential or beneficial in response to a certain stimulus, and therefore those genes are potential candidates for in-depth investigation. This notion further implies that living organisms have mechanisms that sense various stimuli in the intracellular and extracellular environments, and that organisms also develop genetic regulatory networks that determine gene expression levels under specific conditions. Due to this nature of the regulatory mechanisms, transcriptomic profiling also facilitates genetic circuit designing. Identification of genes of which expression level changes in response to a specific stimulus leads to elucidation of condition-specific regulatory factors and sequences of intergenic region. For example, a genetic circuit can be regulated by a certain stimulus by introducing an upstream regulatory region for a gene. Furthermore, bioinformatics algorithms enable statistical and efficient analysis of microarray data and enrich the interpretation. These bioinformatics tools reduce microarray data sizes and help to identify differentially expressed genes, transcriptomic regulations and cellular functions related to a specific condition. Therefore, microarray-driven transcriptomic profiling has become a promising tool in the field of biological sciences and biochemical engineering.

On the other hand, it has been of great interest for biochemical engineers to develop innovative strategies to systematically improve bioprocesses that utilize whole-cell microbial systems for production of valuable chemicals. Particularly, accumulation of genetic information and development of high-throughput analytical technologies with bioinformatics algorithms has extended the dimension of engineering approaches.

Traditionally, a bioprocesses development relied on evolutionary engineering to obtain a host strain with better characteristics. Evolutionary strain engineering approaches employ a selective pressure to screen strains with desirable physiological traits, and these approaches are based on the emergence of several random and unknown mutations in bacterial genome. On the other hand, rational approach aims to identify a certain target, modification of which increases possibility of obtaining desired outcome by designing artificial mutations to be introduced. Unlike evolutionary approaches, rational approach determines engineering target by examining several candidates by compiling available information and prediction made by computational algorithms.

Among many factors that can be optimized to improve a bioprocess, intrinsic robustness of a microbial system has increased in importance. This is mainly attributed to the rationale that innate capacity of a microbial host to perform designed biochemical reactions is fundamentally restricted by growth ability and resistance to detrimental factors generated during the reactions. This rationale has been supported by studies showing that expression of heterologous proteins inhibit growth of host cells, decreasing the yield of the protein production. Therefore, robustness and stability of a microbial system often become a bottleneck in development of efficient bioprocesses, which can be improved by increased robustness of a microbial host.

1.2 Objectives and Scope

This thesis aims to develop a systems-level cellular engineering framework for rational strain optimization that leads to improved bioprocesses, with a particular emphasis on characterization and reconstruction of genetic regulatory networks underlying cellular responses to specific bioprocess-relevant conditions. Towards this

aim, systems-level cellular responses have been elucidated, and essential cellular functions, and regulatory factors have been characterized. Then, underlying genetic regulatory networks have been identified based on bioinformatics analysis of gene-transcription factor relationships. These identified genetic networks have been reconstructed to obtain desirable phenotypes of *E. coli*. Finally, this novel systems cellular engineering framework has been harnessed to address strain optimization strategies for three specific, bioprocess-relevant conditions: genetic disruption in quorum sensing (Chapter 4), oxidative stress under anaerobic condition (Chapter 5), and hydrocarbon solvent stress (Chapter 6). Figure 1-1 shows a schematic diagram of this thesis.

With a particular emphasis on three aforementioned conditions, Chapter 4 aims to characterize transcriptomic responses in the absence of *luxS*, a gene responsible for production of autoinducer-2 (a quorum sensing molecule for *E. coli*) and to identify regulatory factors that link *luxS* and swimming motility. Chapter 5 aims to comparably investigate cellular responses, particularly anaerobic metabolisms and oxidative stress responses in the presence of hydrogen peroxide (H_2O_2) under aerobic and anaerobic condition. Furthermore, Chapter 5 aims to reveal essential regulatory factors and improve *E. coli* tolerance against H_2O_2 . Lastly, Chapter 6 aims to describe systems-level cellular responses to isooctane and to improve robustness of *E. coli* against isooctane by reconstituting underlying genetic regulatory networks.

Figure 1-1 A schematic diagram of this thesis

1.3 Significances

The work presented in this thesis provides a cellular engineering framework that can be exploited for developing microbes with desired phenotypes, contributing to identification of novel molecular targets for improved bioprocesses. For example, this thesis work will help scientists and researchers to engineer microbial hosts to be more tolerant to biofuel molecules. With the soaring need for alternative energy sources, microorganisms have been genetically and metabolically engineered to produce alcohols such as ethanol and butanol. However, these organic solvents have bactericidal effects, and as a result, production of biofuel molecules has been limited. Therefore, it is imperative to improve microbial resistance against toxicity of metabolites and organic solvents. So far, there have been a few examples to engineer yeast for higher tolerance to organic solvents: Iinoya *et al* (2009) (Iinoya, Kotani *et al.* 2009) engineered *Saccharomyces cerevisiae* (*S. cerevisiae*) Sigma1278b to improve oxidative stress tolerance because stresses imposed by ethanol are related to oxidative stress responses. While Iinoya and colleagues improved yeast viability by engineering its own enzyme involved in oxidative stress, Ukibe and colleagues (Ukibe, Hashida *et al.* 2009) constructed more tolerant *S. cerevisiae* against oxidative stress by introducing bacterial genes and over-expressing antioxidants.

In addition to bioprocess, the cellular engineering framework presented in this thesis, can be exploited for development of drugs to treat bacterial infections. For example, elucidation of stress responses of microbial cells may facilitate discovery of new antimicrobial agents that can specifically target defense mechanisms in the microbial cells. This higher specificity may help us to avoid drawbacks of current antibiotics treatment. Antibiotics have been extensively used to treat bacterial infections. Most

antibiotics target bacterial cell walls, cytoplasmic membranes or essential enzymes, of which disruption abolish normal growth of microbial cells. However, administration of antibiotics has increased the possibility for an advent of superbugs that show higher tolerance and resistance to antibiotics used for human. In addition, it has been reported that under stress conditions such as treatment of antibiotics or immune responses, bacterial cells form subpopulations of non-dividing dormant cells that more likely elude antibiotics or immune responses. Interestingly, it has been found that formation of dormant cells correlated with persistent and recursive infections even after long period of antibiotics treatment (Lewis 2007). These reports suggested that more susceptible and specific targets could be identified by investigating responses of pathogen under sub-lethal stress towards development of better antimicrobial reagents. Therefore, this thesis presents a framework that can be applied to obtain a desired microbial phenotype towards improved bioprocesses and to identify novel targets for development of new antimicrobial strategies.

1.4 Thesis Outline

This thesis consists of seven chapters. In Chapter 1, the specific objectives and significance, and motivation of the works presented in the thesis are described. In Chapter 2, the advantages of using *E. coli* as a microbial host for various biological processes and relevant main technological difficulties are discussed. Further, backgrounds and relevance of the following specific conditions are reviewed: *luxS*-mediated quorum sensing, oxidative stress responses and hydrocarbon stress responses. In Chapter 3, the methodology and approaches of the systems engineering framework developed in this thesis are discussed. In Chapter 4, regulatory connectivity between quorum sensing and swimming motility are discussed. In

Chapters 5 and 6, systems-level cellular responses and engineering of genetic regulatory networks for improved stress tolerance are presented, with a particular emphasis on oxidative stress under anaerobic condition, and hydrocarbon isooctane stress, respectively. Finally, in Chapter 7, the conclusions and significance of this thesis, and considerations for future research are discussed.

CHAPTER 2 LITERATURE REVIEW

2.1 Biotechnology and bioprocess

Bioprocess can be defined as any process that utilizes a whole biological system or biochemistry reactions in living cells to obtain desired product. For example, fermentation is one of the traditional bioprocess where defined or undefined bacterial cultures are used to produce food such as butter, cheese, and soybean and chili pastes. Development of biotechnological tools has extended areas of bioprocesses by allowing bioprocess engineers to introduce a gene or implement a pathway in a heterologous host. Heterologous protein expression has met the need in production of human enzymes, antibodies, vaccines and other therapeutic proteins. In addition, enzymes and a whole organism have been engineered to catalyze difficult reactions in organic chemical and pharmaceutical industries providing highly value added chemical precursors and medicinal products. Therefore, introduction of biotechnology has facilitated and advanced bioprocesses to provide more solutions in human society.

2.2 *Escherichia coli* (*E. coli*) as a host in bioprocesses

E. coli has been a primary host in the above-mentioned bioprocesses. In this Section, general physiologies including morphology and growth of *E. coli* will be briefly introduced (Section 2.2.1) followed by current progress in *E. coli* genomic studies and available biotechnological tools for *E. coli* (Section 2.2.2). In addition, current limitations faced in the field regarding the use of *E. coli* will be addressed.

2.2.1 *E. coli* physiology

E. coli is a Gram negative bacterial species that belongs to the *Enterobacteriaceae* family. *E. coli* is rod-shaped, about 2-3 μm -long and 0.7-0.8 μm in diameter, but the

dimension of *E. coli* varies depending on culture medium composition and growth states (Pierucci 1978). *E. coli* is a facultative anaerobe, which is able to grow under both aerobic and anaerobic conditions utilizing various substrates and final electron acceptors such as oxygen, nitrate, nitrite, and dimethylsulfoxide (DMSO) in respiration (Ingledew and Poole 1984). *E. coli* can grow in the minimal medium containing a carbon source such as glucose and glycerol and in general cell culture is prepared in rich medium such as Luria Broth (consisting of 10 g tryptophan, 5 g yeast extract and 10 g NaCl in 1 L distilled water). *E. coli* undergoes typical bacterial growth phases: lag phase, exponential phase, stationary phase and death phase. After adapting to a new condition during a lag phase, *E. coli* exponentially doubles its number about every 30 min at its optimal growth temperature, 37 °C, although the growth rate varies depending on strains and medium composition.

2.2.2 *E. coli* genomics and genetic manipulations

Strain is defined as a group of organisms of the same species, having distinctive characteristics but not usually considered a separate breed or variety. *E. coli* species is divided to various strains according to serotypes determined by antigens on the surface of *E. coli* (O-, H- and K-antigen) or phylogeny characterized by evolutionary relations of genomes. Most of *E. coli* strains, such as K-12 and B strains, are non-pathogenic and non-toxic to human. However, *E. coli* enterohemorrhagic strains (EHEC), such as *E. coli* O157:H7 and *E. coli* O104:H4, are pathogenic.

E. coli is one of the most extensively studied prokaryotic microorganisms. In addition to the simplicity of the culture conditions, it is easy to manipulate *E. coli* genomic DNA and plasmid DNA by means of homologous recombination, transduction, conjugation and chemical or electrical DNA transformation. Therefore, a plethora of

genetic engineering tools for *E. coli* have been developed and widely established. Since the complete genome sequence of *E. coli* K12 was published (Blattner, Plunkett *et al.* 1997), 4503 genes, 4463 transcription units and 1991 biological reactions have been annotated in the public database EcoCyc (Keseler, Collado-Vides *et al.* 2011). Functional characterization of enzymes, transporters and transcription factors encoded in *E. coli* genome has facilitated identification and prediction of protein functions, metabolic pathways and genetic regulatory mechanisms in other microorganisms based on sequence similarity and structural homology of peptide domains. Therefore, *E. coli* has various advantages as an engineering host for bioprocesses and it has been extensively used to produce various proteins and small molecules. A few examples are production of recombinant human interleukin-2 (IL-2) (Roifman, Mills *et al.* 1985), erythropoietin (Johnson, Middleton *et al.* 1996), poly-hydroxybutyrate for a biodegradable plastic (Steinbüchel 2005), 1,3-propanediol as a carpet fiber monomer (Nakamura and Whited 2003), and amino acids for animal feed additives (Park and Lee 2008).

2.2.3 Limitations and improvement of *E. coli* strains for bioprocesses

Major limitations of *E. coli* in bioprocess have been directly related to expression of heterologous proteins originating from eukaryotes due to inherent differences between eukaryotes and prokaryotes. Given the long history of this issue, various approaches have been suggested to fill the gap between eukaryotes and prokaryotes, and some approaches have improved *E. coli* strains for heterologous protein expression. On the other hand, recently, there has been more effort to improve bioprocess by characterizing and optimizing microbial hosts at systems level. In this section, problems and improvements of heterologous protein expression in *E. coli* will be discussed in relation to improper protein folding (Section 2.2.3.1) and codon usage

bias (Section 2.2.3.2) followed by the systems-level optimization of microbial host in Section 2.2.3.3.

2.2.3.1 Improper protein folding

When heterologous proteins, especially eukaryote-originated proteins, are expressed in *E. coli*, the expressed proteins often lose or change activity of the original proteins. This inactivity is originated from a fact that the translated peptides do not form correct tertiary or quaternary protein structures. Proper protein folding often requires post-translational modification and proper environmental conditions. However, *E. coli* strains lack post-translational modification mechanisms that are essential for proper folding and activity of proteins in eukaryotes, by changing conformation of proteins. This may be mainly attributed to the absence of endoplasmic reticulum, which is full of various enzymes, chaperones and sensors that help proteins to fold and assemble in eukaryotes (Gething and Sambrook 1992). To solve this problem, there has been much effort to assist protein-folding ability of *E. coli*. For example, glycosylation function of other bacterial has been introduced in *E. coli*. The *pgl* gene cluster of *Campylobacter jejuni* encodes proteins that have activity of N-linked glycosylation function. When the *pgl* cluster was co-expressed, *C. jejuni*-derived PEB3 and ArcA glycoproteins were functionally produced in *E. coli* (Wacker, Linton *et al.* 2002). Disulfide bond formation is another important post-translational modification that determines structural configuration and stability of expressed proteins. In *E. coli*, disulfide bond formation could be enhanced by co-expression of cytoplasmic chaperones and by targeting and secreting expressed proteins into periplasmic space. Periplasmic space provides more oxidizing conditions, which makes disulfide bond formation more favorable compared to cytoplasm (de Marco 2009).

2.2.3.2 Biased codon usage

Codon usage bias refers to the fact that different organisms have different preferences for genetic codons and concentration of cognate tRNAs (Gustafsson, Govindarajan *et al.* 2004). Insolubility and low expression level of heterologous proteins is well correlated with codon usage bias between *E. coli* and an original host from which the expressed proteins are derived. In addition, computational analyses of *E. coli* genomes suggested that low-frequency-usage codons in coding region of nucleotides determine protein translation rates and specific structural domains of proteins (Angov 2011). One approach to solve this problem was to co-express genes coding rare tRNAs in *E. coli*, and this approach has been shown to improve expression of several heterologous proteins (Kane 1995). While cost of *de novo* gene synthesis decreases, it is now possible to synthesize a gene sequence of which codons are optimized by various codon optimization algorithms. For instance, two reports have shown that *de novo* synthesis of codon-optimized sequences can increase expression levels of human phosphatidylcholine transfer protein and mouse c-Fos proteins in *E. coli* by 10-20% (Gustafsson, Govindarajan *et al.* 2004). In summary, co-expression of assisting proteins and optimization of protein sequences have solved the problems pertaining to improper protein folding and codon usage bias in *E. coli* for improved expression of heterologous proteins.

2.2.3.3 Optimization of *E. coli* strains to improve bioprocess: group behavior, stress responses and tolerance

While the above-mentioned problems of *E. coli* have been much resolved by approaches to overcome innate differences between prokaryotes and eukaryotes, there have recently been an increasing number of attempts to understand and engineer a whole microorganism at systems level for improved bioprocess. These attempts have

aimed to minimize unnecessary responses of microbial cells, to remove harmful effects of engineering, to optimize and balance metabolic fluxes, and to fine-tune the engineered microbial organism.

These goals have been pursued by elucidating molecular mechanism of microbial interactions and responses to detrimental environments and to other neighboring microbial cells. Stress responses of microbial cells during bioprocess due to toxicity of intermediates or final products of engineered pathways have been of particular interest. While native host of the introduced genes or pathways have innate mechanisms to deal with the metabolites, heterologous host *E. coli* is more vulnerable to the unnatural toxic metabolites. For example, production and accumulation of toxic chemicals such as alcohols and aromatic compounds inhibit growth and metabolisms of *E. coli*. Retarded growth and increased death rate of *E. coli* cells hinder production of these chemicals due to the reduced number of cells available to produce the chemicals. In addition, biosynthesis of foreign molecules may disturb the existing cellular metabolisms by altering availability and redirecting flow of substrates and intermediates required in the essential metabolic pathways.

To optimize microbial hosts for a desired bioprocess, it is therefore essential to elucidate mechanisms of interactions with other bacterial cells and cellular responses to toxic chemicals in *E. coli*. Towards this goal, profiling transcriptomic, proteomic, metabolic and physiological changes was suggested to characterize cellular responses and microbial behaviors at systems level. In addition, system-level data in combination with bioinformatics analysis would narrow down high-throughput changes to the manageable number of targets.

In the rest of this chapter, quorum sensing, oxidative stress responses and isooctane stress responses of microbial cells, particularly *E. coli*, will be reviewed focusing on the cellular stress responses and bacterial group behaviors in microbial bioprocesses. In Section 2.3, the term ‘quorum sensing’, will be introduced to approach bacterial group behavior phenomenon. Also, mechanisms of quorum sensing will be discussed in different strains followed by the role of quorum sensing in various aspects. In Section 2.4, generation of oxidative stresses and defense mechanisms in microbial cells will be summarized. Oxidative stress phenomenon will be further reviewed in relation to four different aspects: aerobic respiration, bacterial infection, bacterial cell death and solvent stresses. Furthermore, oxidative stress will be discussed in terms of anaerobic conditions. Lastly, toxicity of organic solvents will be discussed in relation to microbial biofuel production and defense mechanisms in Section 2.5.

2.3 Quorum sensing and bioprocess

2.3.1 Quorum sensing and autoinducers

Quorum sensing refers to a communication-like phenomenon among neighboring bacterial cells. Quorum sensing is mediated by “autoinducers (AI)”, which are small, diffusible and signaling molecules (Vendeville, Winzer *et al.* 2005). These molecules are synthesized, secreted and up-taken by neighboring cells and the producer cells. With the increased concentration of these signal molecules in proportion to the bacterial cell number, bacterial cells in the vicinity recognize these signal molecules and subsequently, the signal molecules synchronistically change gene expressions of individual microbial cell. Therefore, ‘Quorum sensing’ is a term coined for this phenomenon whereby the accumulation of autoinducers (i.e. ‘small, diffusible and signaling molecules’) enables bacteria to monitor population density and adjust their gene regulation (Vendeville, Winzer *et al.* 2005; Hardie and Heurlier 2008). Quorum

sensing among bacterial cells results in coordinated phenomena such as bioluminescence in *Vibrio harveyi*, biofilm formation and secretion of virulence factors (Bassler and Losick 2006).

Autoinducers from different species have diverse chemical structures with different cognate receptors and mechanisms of signal transduction. Figure 2-1 shows some examples of autoinducers from different microbial species. N-Acyl Homoserine Lactones (AHLs) (Figure 2-1a) and oligopeptides (Figure 2-1b) are the most well characterized autoinducers in Gram-negative bacteria and Gram-positive bacteria, respectively. Different bacteria produce AHLs with different R-groups and autoinducer peptides (AIP) with different amino acids sequences. Autoinducer-2 (AI-2; Figure 2-1c) is converted from 4,5-dihydroxy-2,3-pentanedione (DPD), the precursor of AI-2, and exists in two forms: S-THMF-borate in the presence of boron or R-THMF in the absence of boron. Autoinducers of *Vibrio cholerae* (CAI-1 and amino-CAI-1, Figure 2-1d) and *Pseudomonas aeruginosa* (PQS, Figure 2-1e) have relatively long carbon chains. Many quorum sensing molecules are species-specific as Figure 2-1 shows that quorum sensing molecules are diverse in their chemical structures (Choudhary and Schmidt-Dannert 2010). For example, AHLs are strain-specific and only sensed by the host strain that produces a specific AHL. On the contrary, AI-2 can be broadly recognized among Gram-negative and Gram-positive strains, which can result in interspecies quorum sensing behaviors (Ng and Bassler 2009).

Figure 2-1 Chemical structures of diverse autoinducers¹.

(a) N-acyl homoserine lactone autoinducers with different R-groups, which are specifically synthesized and recognized by different Gram-negative bacteria. (b) Linear and circular oligopeptides autoinducers utilized among Gram-positive bacteria. Four different circular autoinducer peptides (AIPs) are produced by *Staphylococcus aureus*. (c) DPD, the precursor of autoinducer-2 (AI-2). AI-2 exists as S-THMF-borate or R-THMF depending on the presence of borates. (d) Structure of *Vibrio cholerae* autoinducers, CAI-1 and Amino-CAI-1. (e) PQS autoinducer of *Pseudomonas aeruginosa*.

¹ Strain names and corresponding quorum sensing molecules were referred to Ng and Bassler 2009 and chemical structures were generated by Symyx draw v3.3

2.3.2 AI-2 Synthesis in *E. coli*

In *E. coli*, quorum sensing is mediated by AI-2 (Surette and Bassler 1998), a cyclized form of DPD. AI-2 is derived from S-ribosyl homoserine (SRH), one of metabolites involved in activated methyl cycle (AMC) (Hardie and Heurlier 2008). Figure 2-2 shows AMC pathway and AI-2 synthesis pathway from DPD. S-ribosyl homocysteine lyase coded in *luxS* is responsible for conversion of SRH to homocysteine and production of DPD. It is interesting that the homologous enzymes of S-ribosylhomocysteine lyase have been found in 55 different bacterial species (Xavier and Bassler 2003). This finding suggests that the enzymatic reaction mediated by the S-ribosylhomocysteine lyase (EC 4.4.1.21) may be essential in bacterial metabolism. In addition, this global existence of S-ribosylhomocysteine lyase that produces AI-2 might explain unspecific quorum sensing mediated by AI-2 among different species.

2.3.3 Quorum sensing in *E. coli*

Roles of quorum sensing to regulate gene expressions in *E. coli* K-12 strains are still obscure, but regulation of *lsr* promoter by AI-2 and LsrR/LsrK system has been well characterized. Figure 2-3 shows a schematic diagram of regulation of *lsr* regulon by a quorum-sensing molecule, AI-2, in *E. coli*. Once AI-2 is synthesized and secreted, extracellular AI-2 molecules are transported into the neighboring *E. coli* cells or into the producer itself by Lsr transporters coded in *lsr* operon (*lsrA*, *lsrC*, *lsrD*, *lsrB*, *lsrF* and *lsrG*). LsrK phosphorylates imported AI-2 molecules, and phosphorylated AI-2 molecules induce expression of *lsr* operon by inhibiting a repressor, LsrR. Increased expression of Lsr transporters facilitate *E. coli* cells to uptake more AI-2, which increases expression level of Lsr transporters resulting in a positive feedback loop. This mechanism shows that secreted extracellular AI-2 increases its own uptake from outside of cells by inducing gene expressions of *lsr* operon. In addition, it suggests that expression of genes regulated by quorum sensing molecule, AI-2, can be initiated and amplified depending on the size of bacterial population, which increases extracellular AI-2 concentration in *E. coli*.

Figure 2-3 Schematic diagram of AI-2 quorum sensing in *E. coli*

Black dots symbolize AI-2 molecules. Four empty squares represent Lsr transporter complex coded in *lsrACDBFG*. Three empty ovals represent proteins (LsrK, LsrR and LuxS). LsrK phosphorylates transported AI-2 molecules, and phosphorylated AI-2 (a black dot with PO₄) inhibits LsrR, a repressor of *lsrACDBFG* operon allowing transcriptional induction of *lsr* operon promoter.

2.3.4 Quorum sensing in various bacterial phenotypes

Initially, the roles of quorum sensing have been extensively studied in relation to bacterial infection and disease because many bacterial phenotypes reportedly regulated by quorum sensing are well correlated with emergence of resistance and persistence in bacterial infection treatment. For example, quorum sensing has been highlighted as a decisive process for biofilm formation, virulence factor secretion, antibiotics production, sporulation and competence for DNA uptake (Ng and Bassler 2009), which help bacterial cells to persist and survive under bactericidal conditions. Furthermore, participation of quorum sensing in bacterial infection has been studied more extensively because the above-mentioned bacterial behaviors are reinforced when a large number of bacterial cells synchronize their genetic regulations and exert collective and accumulative effects. Therefore, quorum sensing has been investigated to understand mechanisms of bacterial tolerance and persistence in bacterial infection. Also, disruption of quorum sensing has been suggested as a novel strategy to eradicate bacterial infection.

In this context, many studies have studied physiological changes of various bacterial species in the absence of AI-2 production by knocking out *luxS* gene. For example, one study showed that deletion of *luxS* gene altered expression level of more than 400 genes, which are involved in DNA processing, biofilm formation, carbohydrate metabolism, surface- and outer membrane-associated functions and virulence in *Streptococcus mutans* (Wen and Burne 2004). Table 2-1 summarizes the effects of *luxS* deletion on cellular processes in various bacterial strains highlighting the regulatory role of *luxS* genes in various phenotypes. The investigated cellular processes are biofilm formation, flagella biogenesis, motility and virulence as well as individual phenotypes such as growth rate and stress resistance. These studies imply

that LuxS, S-ribosyl homocysteine lyase, or AI-2 may play a role in global transcriptional regulation in various species.

Table 2-1 Effects of *luxS*-deletion on bacterial phenotypic traits.

Class	Effect ^a	Bacteria species	Reference
Growth rate	-	<i>Streptococcus pyogenes</i> <i>Campylobacter jejuni</i> strain 81-176 (37 °C)	(Lyon, Madden <i>et al.</i> 2001; He, Frye <i>et al.</i> 2008)
	+	<i>Escherichia coli</i> O157:H7	(Sperandio, Torres <i>et al.</i> 2001)
	N/A	<i>Campylobacter jejuni</i> NCTC 11168, <i>E. coli</i> K-12 <i>Campylobacter jejuni</i> strain 81-176 (42 °C)	(Elvers and Park 2002; Wang, Li <i>et al.</i> 2005; He, Frye <i>et al.</i> 2008)
Quorum sensing	-	<i>Porphyromonas gingivalis</i> <i>Serratia spp</i> <i>E. coli</i> O157:H7 <i>Campylobacter jejuni</i> strain 81-176	(Burgess, Kirke <i>et al.</i> 2002; Coulthurst, Kurz <i>et al.</i> 2004; Soni, Jesudhasan <i>et al.</i> 2007; He, Frye <i>et al.</i> 2008)
	N/A	<i>E. coli</i> K-12 <i>Staphylococcus aureus</i> <i>Erwinia amylovora</i> <i>Streptococcus mutans</i> UA159	(Wang, Li <i>et al.</i> 2005; Doherty, Holden <i>et al.</i> 2006; Rezzonico and Duffy 2007; Sztajer, Lemme <i>et al.</i> 2008)
Biofilm formation	-	<i>Streptococcus mutans</i> <i>Klebsiella pneumoniae</i> <i>Photobacterium luminescens</i> subsp. <i>laumondii</i> TT01	(Wen and Burne 2004; Balestrino, Haagenensen <i>et al.</i> 2005; Krin, Chakraborty <i>et al.</i> 2006)
	+	<i>Aeromonas hydrophila</i>	(Kozlova, Popov <i>et al.</i> 2008)
	N/A	<i>E. coli</i> K-12, <i>Staphylococcus aureus</i>	(Wang, Li <i>et al.</i> 2005; Doherty, Holden <i>et al.</i> 2006)
Resistance	-	<i>Streptococcus anginosus</i> (erythromycin and ampicillin) <i>Campylobacter jejuni</i> strain 81-176 (H ₂ O ₂) <i>Erwinia amylovora</i> strain	(Ahmed, Petersen <i>et al.</i> 2007; He, Frye <i>et al.</i> 2008; Gao, Song <i>et al.</i> 2009)

		NCPB1665 (H ₂ O ₂)	
	+	<i>Streptococcus mutans</i> (H ₂ O ₂) <i>Porphyromonas gingivalis</i> LY2001 (50 °C, H ₂ O ₂ , pH 9)	(Wen and Burne 2004; Yuan, Hillman <i>et al.</i> 2005)
	N/A	<i>Campylobacter jejuni</i> NCTC 11168 (paraquat, H ₂ O ₂)	(Elvers and Park 2002)
Flagella biogenesis	-	<i>E. coli</i> O157:H7 <i>Vibrio alginolyticus</i>	(Sperandio, Torres <i>et al.</i> 2001; Tian, Wang <i>et al.</i> 2008)
	N/A	<i>Campylobacter jejuni</i> strain 81116 <i>Helicobacter pylori</i> G27	(Jeon, Itoh <i>et al.</i> 2003; Rader, Campagna <i>et al.</i> 2007)
Motility	-	<i>E. coli</i> O157:H7 <i>Campylobacter jejuni</i> NCTC 11168 <i>Campylobacter jejuni</i> strain 81116 <i>Helicobacter pylori</i> HPKY08 <i>Helicobacter pylori</i> G27 <i>Aeromonas hydrophila</i> <i>Campylobacter jejuni</i> strain 81- 176 (swarming, 37 °C) <i>Vibrio alginolyticus</i> <i>Erwinia amylovora</i> strain NCPB1665	(Sperandio, Torres <i>et al.</i> 2001; Elvers and Park 2002; Jeon, Itoh <i>et al.</i> 2003; Osaki, Hanawa <i>et al.</i> 2006; Rader, Campagna <i>et al.</i> 2007; Soni, Jesudhasan <i>et al.</i> 2007; He, Frye <i>et al.</i> 2008; Kozlova, Popov <i>et al.</i> 2008; Tian, Wang <i>et al.</i> 2008; Gao, Song <i>et al.</i> 2009)
	+	<i>Photobacterium luminescens</i> subsp. <i>laumondii</i> TT01 (twitching)	(Krin, Chakroun <i>et al.</i> 2006)
	N/A	<i>E. coli</i> K-12 <i>Campylobacter jejuni</i> strain 81- 176 (swarming, 42 °C)	(Wang, Li <i>et al.</i> 2005; He, Frye <i>et al.</i> 2008)
Virulence	-	<i>Streptococcus pyogenes</i> <i>Serratia spp.</i>	(Lyon, Madden <i>et al.</i> 2001; Coulthurst, Kurz <i>et al.</i> 2004;

	<i>Vibrio harveyi</i>	Defoirdt, Bossier <i>et al.</i> 2005;
	<i>E. coli</i> O157:H7	Jordan, Sperandio <i>et al.</i>
	<i>Photorhabdus luminescens</i>	2005; Krin, Chakroun <i>et al.</i>
	subsp. laumondii TT01	2006; Osaki, Hanawa <i>et al.</i>
	<i>Helicobacter pylori</i> HPKY08	2006; Walters, Sircili <i>et al.</i>
	<i>Vibrio alginolyticus</i>	2006; Tian, Wang <i>et al.</i>
	<i>Erwinia amylovora</i> strain NCPPB1665	2008; Gao, Song <i>et al.</i> 2009)
+	<i>Erwinia carotovora ssp.</i>	(Laasik, Andresen <i>et al.</i>
	<i>Carotovora</i>	2006; Kozlova, Popov <i>et al.</i>
	<i>Aeromonas hydrophila</i>	2008)
N/A	<i>Porphyromonas gingivalis</i>	(Burgess, Kirke <i>et al.</i> 2002;
	<i>Campylobacter jejuni</i> NCTC 11168	Elvers and Park 2002;
	<i>Klebsiella pneumonia</i>	Balestrino, Haagensen <i>et al.</i>
	<i>E. coli</i> O157:H7	2005; Jeon and Itoh 2007)

^a -, decreased; +, increased; N/A, no significant effects

Even though many studies have shown that LuxS/AI-2 quorum sensing might be involved in various bacterial physiologies, a molecular mechanism of LuxS/AI-2-regulation is still unknown. Neither a transcriptional regulatory factor of *luxS* gene nor a protein interaction of LuxS has been discovered. Furthermore, roles of LuxS/AI-2 are still controversial depending on strains and cellular processes. For example, roles of quorum sensing in bacterial responses to oxidative stress have been studied by knocking out *luxS* gene in six different strains, but the relationships between the absence of *luxS* gene and bacterial resistance to oxidative stress were not conclusive (Table 2-1). When quorum sensing was inhibited by deletion of *luxS* gene, *Streptococcus mutans* (Wen and Burne 2004) and *Porphyromonas gingivalis* LY2001 (Yuan, Hillman *et al.* 2005) became more resistant against hydrogen peroxide (H₂O₂) treatment while *Campylobacter jejuni* strain 81-176 (He, Frye *et al.* 2008) and *Erwinia amylovora* strain NCPPB 1665 (Gao, Song *et al.* 2009) were more vulnerable to H₂O₂. However, *Campylobacter jejuni* NCTC 11168 did not show significant difference in cell viability against H₂O₂ and paraquat (Elvers and Park 2002).

Another example of the obscure correlations can be found between LuxS/AI-2 quorum sensing and bacterial motility, both swimming and swarming motility. Swimming motility refers to a type of bacterial motility in fluids such as a broth medium. Swimming motility is often described as “running” (for a linear movement) and “tumbling” (for a change of directions) motions depending on the directional rotations of flagella. While swimming motility is related to movement of an individual bacterial cell, swarming motility is related to movement of a group of bacteria on a solid surface. The swarming motility also requires flagella and this surface translocation has been known to be important for bacterial colonization (Harshey 2003).

In this context, previous studies showed that LuxS-based metabolism and/or quorum sensing might be importantly connected to motility (Table 2-1). However, the roles of *luxS* gene in motility and flagella biogenesis were not conclusive. While deletion of *luxS* gene decreased swimming motility of *E. coli* O157:H7 (Sperandio, Torres *et al.* 2001), *luxS*-deleted mutants of *E. coli* K12 did not show difference in their swimming motility (Wang, Li *et al.* 2005). In addition, at 37 °C, swarming motility of *Campylobacter jejuni* strain 81-176 was decreased by *luxS* gene deletion, but the swarming motility did not show a significant difference at 42 °C compared to its normal counterpart (He, Frye *et al.* 2008). This variance suggests that oxidative stress responses and swimming motility of *luxS*-deleted mutants and wild type are complicatedly regulated depending on species, strain genotypes, modes of motility and experimental conditions. Therefore, it is essential to provide a more complete linkage between the *luxS* gene and essential cellular processes at the molecular mechanism level.

2.4 Oxidative stress responses

2.4.1 Oxidative stress and reactive oxygen species (ROS)

Microbial cells are maintained by sophisticated control mechanisms and regulated by biomolecules such as proteins, nucleic acids and lipids. The function of these macromolecules can be modulated by numerous physico-chemical reactions such as oxidation, reduction and phosphorylation reactions. Some by-products of these chemical reactions are often deleterious to cells, but their production is inevitable because they result from essential cellular pathways. Reactive oxygen species (ROS) are one example of the harmful by-products, generated from aerobic respiration pathways. In the cell, produced ROS oxidize essential macromolecules due to intrinsic chemical instability of ROS. In other words, ROS extract electrons from proteins and

DNAs in the vicinity. Oxidized macromolecules change their biological activity and as a result, disrupt cellular processes where these oxidized molecules are involved. Therefore, oxidative stress refers to a biological condition where cells are exposed to detrimental effects of strong oxidizing agents, particularly ROS.

There are three major forms of ROS in the cells: hydrogen peroxide (H_2O_2), superoxide (O_2^-) and hydroxyl radicals ($-\text{OH}^\cdot$). Aerobic respiration is a major source of ROS generation in the cells, and Figure 2-4 shows key routes of ROS generation in microbial cells. Superoxide (O_2^-) can be produced by NADPH oxidase (Figure 2-4-1), redox cycling drugs (Figure 2-4-2) and autoxidation of flavoproteins (Figure 2-4-3). Extracellular hydrogen peroxide (H_2O_2) can be diffused through the membrane and converted to hydroxyl radicals ($-\text{OH}^\cdot$) by Fenton reaction mediated by ferrous (Fe^{2+} ; Figure 2-4-4) inside the cell. The generated hydroxyl radicals react with other molecules in the cells damaging and changing activity of normal cellular processes.

Figure 2-4 Extracellular and intracellular routes of ROS production in microorganisms

2.4.2 Oxidative stress and bioprocess

As shown in Figure 2-4, the sources of ROS are ubiquitous and therefore, oxidative stress has been highlighted in various aspects. In this Section, oxidative stress will be discussed in relation to aerobic respiration, bacterial cell death and alcohols production.

2.4.2.1 Endogenous production of ROS and respiration

Oxidative stress is an inevitable problem in aerobic organisms that reside in environmental conditions where substantial amount of oxygen is present. To sustain life, all living organisms need energy in the forms of NADH, FADH₂ and ATP. In the presence of oxygen molecules, these energy molecules are mostly produced by aerobic respiration. During aerobic respiration, cellular redox enzymes extract and transfer electrons from nutrients to final electron acceptors via electron transport chains in the membrane. However, ROS is also endogenously formed in the process of aerobic respiration. For example, superoxide and hydrogen peroxide are formed when oxygen molecules react with reduced flavoenzymes such as NADH dehydrogenase II, succinate dehydrogenase, sulfite reductase, and fumarate reductase (Imlay 1995; Messner and Imlay 1999; Messner and Imlay 2002). In addition, it was shown that a large amount of hydrogen peroxide is produced by a flavin-dependent desaturating dehydrogenase (NadB), where reduced flavin is oxidized by molecular oxygen while producing hydrogen peroxide as a product (Korshunov and Imlay 2010). Also, superoxide is reportedly produced by autooxidation of dihydromenaquinone in the cytoplasmic membrane and released to periplasmic space of *E. coli* (Korshunov and Imlay 2006).

Given that ROS production is closely related to aerobic respiration, it is important for aerobes to have cellular mechanisms to protect themselves from damages by ubiquitous ROS while preferably utilizing oxygen molecules as final electron acceptors. Higher tolerance to ROS-driven oxidative stress allows microbial cells to vigorously grow and actively perform metabolic reactions engineered for desired bioprocesses.

2.4.2.2 Oxidative stress and bacterial cell death

Oxidative stress is one of the stresses that induce, so-called “bacterial programmed cell death” (Engelberg-Kulka, Amitai *et al.* 2006). Bacterial cell death is accompanied by cell lysis, which is represented by lost of cell membrane integrity. There are several genes and operons known to be involved in bacterial cell lysis. The *cidABC* and *lrgAB* operons are one of two-component regulatory systems, which control cell lysis in *Staphylococcus aureus* by holin (*cidA*) and antiholin (*lrgA*) interactions (Rice and Bayles 2008). Transcription of these two operons is induced by excessive glucose in the culture medium, and *cidC* encode a pyruvate oxidase. Interestingly, mutation of *cidC* was shown to increase cell viability in the stationary phase in *S. aureus* (Patton, Rice *et al.* 2005). Moreover, mutation of *spxB*, a homologue of *cidC* in *Streptococcus pneumonia*, increased viability in the stationary phase by decreasing production of ROS, particularly hydrogen peroxide generated as a by-product of pyruvate oxidase coded in *spxB* (Regev-Yochay, Trzcinski *et al.* 2007). It is noteworthy that the characteristics of the process of death induced by hydrogen peroxide in *S. pneumonia* were similar to those of apoptosis in eukaryotic cells. This result suggests that oxidative stress might play a role in bacterial cell death and regulation of cell lysis.

Furthermore, ROS generation has been reported to be a common mechanism of death caused by bactericidal antibiotics (Kohanski, Dwyer *et al.* 2007). Kohanski and colleagues investigated antibiotics of which major targets were cellular components involved in DNA replication, repair of damaged DNA, protein synthesis, and cell wall turnover in bacterial cells. Even though initial drug-target interactions vary among antibacterial drugs, it was shown that all three bactericidal antibiotics resulted in production of hydroxyl radicals, which induce bacterial cell death in both Gram negative and Gram-positive bacteria (Kohanski, Dwyer *et al.* 2007). It was shown that oxidative stress was generated via disrupted tricarboxylic acid cycle, temporary shortage of NADH, deactivation of iron-sulfur clusters, and Fenton reaction. Interestingly, bacteriostatic antibiotics, which do not kill bacterial cells but inhibit growth of bacterial cells, did not produce ROS unlike bactericidal antibiotics.

Another interesting fact is that ROS is a common defense mechanism in bacterial cells, mammalian macrophages and plant cells against other bacterial competitors or bacterial infections (Imlay 2008). For example, lactic acid bacteria produce a large amount of hydrogen peroxide to inhibit growth of competitors. Also, NADPH oxidase is known to generate hydrogen peroxide molecules in physically damaged spots of plants. The use of NADPH oxidase in mammalian cells is more complicated. Upon infecting mammalian cells, bacterial cells produce virulence factors, which are toxic to host cells often causing growth arrest and disruption of normal cellular functions. Host cells show inflammatory responses of activated immune system against these virulence factors produced by infectious bacteria. The immune system consists of innate and adaptive immune systems, which recruit humoral and cell-mediated immunity. Both immune systems involve antigen-presenting cells (APC) such as macrophages, neutrophils and dendritic cells. These cells play a major role in

internalizing antigens of invading pathogens by phagocytosis and presenting the antigens on the surface. To kill the internalized bacteria, phagocytes, macrophages and other lymphocytes employ ROS such as superoxide (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radicals ($OH\cdot$) (Chang, Small *et al.* 2006).

Although the molecular mechanism between regulation of bacterial cell death and ROS is yet conclusive, these results imply that ROS play an important role in bacterial cell death. Given that the growth and death of bacterial cells are among the essential factors that determine productivity of bioprocesses, understanding of bacterial responses to oxidative stresses will provide strategies to engineer microbial cells for an improved bioprocess.

2.4.2.3 Oxidative stress and organic solvents in bioprocess

As a routine laboratory practice, 70% ethanol has been generally used to sterilize working bench, glassware, pipettes and gloves. Although alcohols have been known for bactericidal effects, the killing mechanisms of alcohols have not been fully understood. Stress responses to organic solvents, which will be discussed in detail later in this chapter (Section 2.5), have been also highlighted in the field of bioprocesses such as whole cell catalysis reactions in organic solvents, biodegradation of organic pollutants and microbial production of biofuel molecules. Therefore, there have been studies to identify genes, proteins and pathways, responsible for cellular responses to alcohols.

In *E. coli*, it was initially shown that alcohols interact with lipid bilayers of bacterial membranes, increase fluidity of membrane and often result in malfunction of membrane components and disintegration of bacterial cell membranes (Borden and Papoutsakis 2007). However, recent studies have suggested that cellular responses to

alcohols accompany oxidative stress responses in various strains. For example, a transcriptomic study showed that ethanol treatment induced expression of oxidative stress related genes in *S. aureus* (Korem, Gov *et al.* 2009). Similarly, Rutherford and colleagues (Rutherford, Dahl *et al.* 2010) reported that n-butanol stress increased expression level of genes related to oxidative stress in *E. coli*. In addition, Brynildsen and colleagues (Brynildsen and Liao 2009) showed that transcription factors regulating oxidative stress responses are involved in a transcriptomic regulatory network of cellular responses to isobutanol in *E. coli*. These results implied that oxidative stress responses are part of stress responses to alcohols. Therefore, understanding oxidative stress responses provides helpful insights into increasing microbial tolerance and improving bioprocesses such as biological production of fuel molecules.

2.4.2.4 Anaerobic environments and oxidative stress

Microbial studies are conducted in artificial laboratory conditions. It is interesting that some strains cannot be cultured in rich media such as Luria broth (LB) and tryptic soy broth (TSB). More surprisingly, microbiologists speculate that only 5-10% microorganisms on earth have been successfully cultivated in the laboratory condition. These two points suggest that gene expression data obtained from artificial environment may not reflect a true existing phenomenon because bacteria have different gene expression profiles that are dependent on culture environments. Therefore, it is critical to provide appropriate conditions *in vitro* when it comes to microbiological studies, especially for gene expression studies under oxidative stresses.

It is noteworthy that biological niches of pathogens (*Escherichia coli*, *Salmonella*, *Pseudomonas*, *Staphylococcus*, *Streptococcus* and *Vibrio cholera*) and bioreactors are often micro-aerobic or anaerobic conditions. A gastrointestinal tract is an example of micro-aerobic niches, where many *E. coli* strains reside along with pathogenic strains and species. Furthermore, oxygen gradient is developed in biofilm due to limited oxygen penetration through the thick biofilm layers, and different oxygen consumption rates depending on depth of the biofilm. Facultative aerobes in the upper layer of biofilm are likely to consume oxygen more vigorously, which limits oxygen supply to the cells in the lower layer of biofilms. As result, most cells in the biofilms are often under hypoxic or anaerobic conditions rather than aerobic conditions.

Under microaerobic or anaerobic conditions, bacterial metabolisms and transcriptomic profiles are different from those under aerobic conditions. For example, unlike aerobic conditions where oxygen is a preferred electron acceptor with high electro-potential in energy production for aerobes, alternative molecules such as nitrate (NO_3^-), nitrite (NO_2^-), fumarate and dimethylsulfoxide (DMSO) serve as electron acceptors in respiration process under anaerobic condition. If all these inorganic electron acceptors are not available, bacterial cells utilize organic substance such as acetyl-CoA, fumarate, formate and pyruvate as final electron acceptors, producing mixed acids, ethanol, CO_2 and H_2 as fermentative products. Utilization of different metabolic pathways changes transcriptomic and proteomic profiles, and bacterial responses to oxidative stress would be different from the responses obtained under aerobic conditions. Therefore, it would be interesting to investigate bacterial responses against oxidative stress under anaerobic condition, which is more relevant to true environment of bacterial infection and anaerobic or fermentative bioprocesses.

There are only few reports available that mention indirect relationships between anaerobic metabolism and oxidative stress resistance. For example, in *Haemophilus influenza*, deletion of aerobic respiration-controlling protein increased susceptibility to H₂O₂ and gene expression levels of anaerobic respiratory chains (Wong, Alugupalli *et al.* 2007). In another study, Chang and colleagues (Chang, Small *et al.* 2006) reported that H₂O₂ treatment up-regulated genes involved in anaerobic metabolism, and especially, genes essential for oxygen-limiting conditions in *Staphylococcus aureus*. Subsequently, these genes have been reported to protect *Staphylococcus aureus* from the damage by oxidants. On the contrary, Johnston and colleagues (Johnston, Lin *et al.* 2009) showed that improved resistance was accompanied by reduced anaerobic metabolism in the sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. Although transcriptomic profiles were investigated in *E. coli* during transitions between anaerobic and aerobic culture conditions (Partridge, Scott *et al.* 2006; Partridge, Sanguinetti *et al.* 2007), no study has so far shown cellular responses of *E. coli* to oxidants such as H₂O₂ under anaerobic conditions. Even though it is known that oxidative stress response play an important role in adaptation to toxicity of oxygen molecules under aerobic condition, oxidative stress response would be regulated by different mechanisms under anaerobic condition. In these regards, anaerobic condition needs to be taken into account to study the oxidative stress response of *E. coli*.

2.4.3 Cellular responses to oxidative stress in *E. coli*

Since oxidation of biologically important molecules can result in malfunction of various cellular processes, bacterial cells have multiple ways to protect their cellular components, DNA and proteins, from oxidative stress. In *E. coli*, the functions of genes and proteins involved in defense mechanism against oxidative stress have been identified by many studies (Imlay 2003; Imlay 2008). First of all, SoxR/SoxS and

OxyR were identified as master transcriptional factors regulating responses to oxidative stress. SoxR/SoxS and OxyR are activated by superoxide (O_2^-) and H_2O_2 , respectively, and these transcription factors induce expression of target genes, which are regulon members.

Peroxidases (*ahpC* and *ahpF*) and catalases (*katG*) are H_2O_2 scavengers regulated by OxyR, and these two enzymes remove H_2O_2 in *E. coli* by two different molecular mechanisms. At millimolar concentrations of H_2O_2 , catalases are more highly induced; the catalases remove H_2O_2 while producing oxygen and two water molecules ($H_2O_2 + H_2O_2 \rightarrow O_2 + 2 H_2O$). On the other hand, H_2O_2 oxidizes cysteine 46 of peroxidase AhpC, and the resulting disulfide bond can be reduced by NADH-reducible flavoprotein AhpF ($RH_2 + H_2O_2 \rightarrow R + 2H_2O$) (Imlay 2008). Therefore, catalases and peroxidases protect cells from H_2O_2 toxicity by directly reacting with and removing H_2O_2 molecules.

When excessive ROS are present inside the cells, ROS damages various macromolecules such as DNA and proteins. For example, H_2O_2 causes mutagenesis of DNA by facilitating Fenton reaction (Henle, Han *et al.* 1999). The Fenton reaction is described by the following two consecutive reactions:

In the Fenton reaction, free intracellular ferrous irons (Fe^{2+}) are oxidized to ferric irons (Fe^{3+}) by H_2O_2 while generating hydroxyl radicals (OH^*). When this reaction occurs in the vicinity of DNA, the produced hydroxyl radicals (OH^*) oxidize nucleotides, which

result in distortion in DNA duplex and mismatches during the process of replication and transcription.

To repair mutated DNA, which contains fundamental genetic information of life, *E. coli* utilizes various DNA repair systems to prevent mutations from transcending to daughter cells. For example, DNA glycosylase, which belongs to a base excision repair (BER) system, detects distortions in DNA duplex and excises the oxidized base before replication (Tchou, Kasai *et al.* 1991). In addition to DNA glycosylase, *E. coli* uses endonucleases to excise mutated DNA sequences. While endonuclease III and endonuclease VIII are constitutively expressed in *E. coli*, endonuclease VI (coded in *nfo*) is specifically induced by SoxRS system, which is activated by oxidative stress (Cunningham, Saporito *et al.* 1986; Li and Demple 1994). On the other hand, UvrABC nucleotide excision repair complex is activated as a back-up system via SOS response when BER system fails to remove all damaged DNA lesions (Imlay and Linn 1987).

While genetic information of DNA can be protected and conserved by the repair systems, oxidation of protein amino acids by ROS is irreversible. Some sulfur-containing amino acids such as cysteine and methionine can be reduced back to normal structures by methionine sulfoxide reductase and disulfide reducing systems (Imlay 2008). Therefore, it seems that oxidized peptides are more likely to be subjected to protein degradation by proteasomal system although the correlation of oxidized proteins and protein degradation in prokaryotes has not been clearly reported. Since oxidation of amino acids cannot be recovered in most of the cases, the reported repairing mechanisms of damaged proteins are based on re-synthesis and recovery of catalytic cofactors of enzymes under stress responses. In particular, iron-sulfur cluster is a ubiquitous co-factor of metalloproteins and one of the most vulnerable targets of

oxidative stress. For example, stability of 4 iron-4 sulfur clusters ([4Fe-4S]) is impaired in reaction to superoxide (O_2^-) and as a result, [4Fe-4S] clusters lose a ferrous ion (Fe^{2+}) as shown in the following equations:

Upon the loss of a ferrous ion from [4Fe-4S] clusters, enzymes cannot function normally, resulting in a failure of certain cellular pathways. Since these iron-sulfur clusters are catalytically important, the cells possess mechanisms to keep the iron (Fe)-sulfur (S) clusters functional. Some enzymes exist as isoforms, which have the same function but are specifically induced under oxidative stress. For instance, fumarase C (an isozyme of fumarase A) and aconitase A (an isozyme of aconitase B) are more resistant to oxidative stress than their other isoforms (Liochev and Fridovich 1992; Cunningham, Gruer *et al.* 1997). Furthermore, *E. coli* uses the specialized system Suf machinery to synthesize iron-sulfur clusters under oxidative stress conditions. Under normal conditions, iron-sulfur clusters are synthesized by the Isc system. However, when iron supply is limited or H_2O_2 is present in the culture, the Suf system is activated for Fe-S cluster synthesis (Imlay 2008). On the other hand, oxidized iron-sulfur clusters can be restored without *de novo* biosynthesis, which requires the Isc or Suf systems. For example, dehydratase was inactivated when the [4Fe-4S] clusters were oxidized by O_2^- or H_2O_2 to $[3Fe-4S]^+$. However, dehydratase was still active under O_2^- or H_2O_2 stress even though neogenesis of this protein was inhibited (Barras, Loiseau *et al.* 2005). The result suggested the existence of repair mechanisms for oxidized iron-sulfur clusters, although the mechanism has not been clear yet (Imlay

2008). Therefore, functional availability of iron sulfur clusters is maintained by the repair mechanisms, the isozymes and the backup systems induced under oxidative stress conditions.

In summary, bacterial cells are equipped with multi-layered protection mechanisms to deal with oxidative stresses. There are scavenging enzymes, systems to preserve iron-sulfur clusters, DNA and protein damage repair systems and even mechanisms regulating intracellular ferrous ion concentration and its storage and uptake from extracellular nutrients.

2.5 Hydrocarbon stress responses

2.5.1 Biofuel production

Alternatives to fossil fuels have been extensively searched for and developed to supply petroleum-based energy and chemical substrates. While biological production of alcohols has been well established, microbial production of alkanes has recently drawn much attention in need of fuels for automobiles. In nature, alkanes are found in microorganisms, insects and plants (Ladygina, Dedyukhina *et al.* 2006) and can be directly converted from fatty acid metabolites; for example, by decarbonylation of fatty aldehydes (Cheesbrough and Kolattukudy 1984). Therefore, much effort has been made to identify and characterize metabolic pathways that can be introduced into microorganisms for alkane synthesis. As a consequence, alkane biosynthesis pathways from cyanobacteria were recently re-constructed in *E. coli* (Schirmer, Rude *et al.* 2010). This study demonstrated production of the mixture of alkanes and alkenes in the range of C13 to C17.

2.5.2 Isooctane (2,2,4-Trimethylpentane, C₈H₁₈)

Isooctane (2,2,4-Trimethylpentane, C₈H₁₈), a medium chain alkane, has been widely utilized in petroleum refinery process because it is an important component of gasoline that prevents automobile engine-knocking. Furthermore, isooctane is an organic solvent frequently used in biochemical processes. In general, introduction of organic solvents in biochemical reactions is advantageous because organic solvents can dissolve water-insoluble substrates or products, which enhances reaction efficiency and product yield of the last purification step (Kanda, Miyata *et al.* 1998). Also, organic solvents make enzymatic reactions more specific by increasing substrate selectivity, stereo-selectivity, region-selectivity and chemo-selectivity (Klibanov 2001). For example, isooctane enhanced stereo-specificity of epoxide hydrolase, stability and solubility of glycidyl phenyl ether (GPE) in two-phase system that requires an organic solvent and an aqueous buffer for whole-cell catalyst reactions mediated by *Bacillus megaterium* ECU1001 (Gong and Xu 2005). Moreover, isooctane was used to provide stereo-selectivity in reduction of ethyl 3-oxobutanoate to ethyl (S)-3-hydroxybutanoate performed by *Saccharomyces cerevisiae* mutant strain, KK-211 (Kanda, Miyata *et al.* 1998). On the other hand, many microorganisms are able to perform a series of complicated chemical reactions with multiple enzymes coded in their genome. Therefore, the combinational usage of living microorganisms and organic solvents enhances bioconversion and biosynthesis of highly valuable chemicals (Kanda, Miyata *et al.* 1998).

2.5.3 Toxicity of organic solvents

It is noteworthy that both microbial production of hydrocarbons and whole catalyst reaction in hydrophobic organic solvents can be limited by cytotoxicity of hydrocarbons. Organic solvent toxicity is not a very new issue; its toxicity to

microorganisms has been highlighted in biochemical processes, such as whole-cell catalyzed biotransformation in organic solvents and bioremediation of organic pollutants (Heipieper, Weber *et al.* 1994). Initially, toxicity studies were focused on aromatic alcohols (e.g. phenol), aromatic solvents (e.g. benzene and toluene) and hexane, of which partition coefficients are in the range from 1 to 5. Studies on the tolerance of *E. coli* and *Pseudomonas putida* strains have proposed mechanisms of the toxicity caused by organic solvents such as hexane and toluene. One of the studies suggested that these organic solvents destroy the integrity of cell membranes by accumulating in the lipid bilayer of plasma membranes (Miura, Zou *et al.* 2000). In addition, other mechanisms have been suggested. Firstly, hydrophobic organic solvents disrupted microbial membranes by direct insertion, and the cytotoxicity of these chemicals showed a strong inverse correlation with the logarithm of their partition coefficients ($\log P_{ow}$) (Aono and Kobayashi 1997; Rutherford, Dahl *et al.* 2010). Secondly, organic solvent tolerance might be related to lipopolysaccharides modification (Pinkart, Wolfram *et al.* 1996) and induction of efflux systems (Aono, Kobayashi *et al.* 1995; Nakajima, Kobayashi *et al.* 1995; Isken and de Bont 1996) in *E. coli*, *P. putida* and *P. aeruginosa*. Lastly, organic solvents affected the chain length and the composition of fatty acids (Ingram 1977; Heipieper, Weber *et al.* 1994; Sikkema, de Bont *et al.* 1995), outer membrane compositions (Ingram 1977; Weber, Isken *et al.* 1994; Pinkart, Wolfram *et al.* 1996), and cell surface hydrophobicity (Asako, Nakajima *et al.* 1997). Due to the strong need of biofuel molecules, the number of studies on their toxicity and the corresponding defense mechanisms in various microorganisms has exponentially increased. These studies are summarized in Table 2-1.

Table 2-2 Toxicity levels for next-generation biofuels (Dunlop 2011)²

	Tolerance (% v/v)	Microorganisms
Biogasolines		
Butanol	1.5%	<i>Escherichia coli</i>
	1.5%	<i>Zymomonas mobilis</i>
	1.5%	<i>Pseudomonas putida</i>
	1.6%	<i>Clostridium acetobutylicum</i>
	2.0%	<i>Saccharomyces cerevisiae</i>
2-Methyl-1-butanol (2MB)	0.1%	<i>Escherichia coli</i>
3-Methyl-1-butanol (3MB)	0.1%	<i>Escherichia coli</i>
Geraniol	0.1%	<i>Escherichia coli</i>
1-Propanol	5.0%	<i>Escherichia coli</i>
Biodiesels		
Farnesyl hexanoate	1.0% ^a	<i>Escherichia coli</i>
Geranyl acetate	0.5% ^a	<i>Escherichia coli</i>
Bioaviation fuels		
Pinene	0.5% ^a	<i>Escherichia coli</i>
Limonene	0.0%	<i>Escherichia coli</i>

^aBiofuel is inhibitory, but not lethal; percentages listed are where growth drops to 75% of the normal, unstressed levels.

² © 2011 Dunlop; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2.5.4 Defense mechanisms against organic solvents

Defense mechanisms against biofuel molecules have been studied at a molecular level and a systems level. As a result, genes and proteins involved in these defensive mechanisms have been extensively identified and studied. Dunlop summarized biofuel tolerance mechanisms in four categories; efflux pump, heat shock proteins, membrane modifications and general stress responses (Dunlop 2011). Several efflux pumps and their role in exporting biofuel molecules have been well characterized mainly in *E. coli*, *P. putida* and *P. aeruginosa*. Currently, heterologous expression of these efflux pumps in a desired strain is the most promising strategy to improve tolerance to biofuels. However, it was observed that expression of efflux pumps was not very effective for short chain alcohols. Moreover, expression level of these efflux pumps needed to be finely tuned not to disrupt production of biofuels as well as cell growth. On the other hand, overexpression of a few heat shock proteins improved the tolerance and production yield. In addition, overexpressing general stress response related genes, such as oxidative stress and iron transport related genes, was shown to improve tolerance against n-butanol in *E. coli*.

In addition to the above-mentioned strategies, there is a need for systematic engineering approach, which requires complicated analysis, sophisticated regulation and control of microbial cells. For example, if membranes are engineered to decrease permeability to biofuel molecules, it might increase viability of microbial host cells against biofuel molecules. However, decreased permeability of membranes might interrupt efficient export of biofuel molecules produced from the engineered microbial cells. Hence, modification of membrane permeability and transporters requires sophisticated regulations and controls. Moreover, biofuel production and stress responses are not only directly related to defensive mechanisms, but also affected by

cellular metabolisms and processes. Therefore, more systematic and rational engineering approaches will be required to control overall metabolic flux and phenotypes of the engineered strain.

In spite of the above-mentioned mechanisms and engineering attempts to increase tolerance against various organic solvents, it is notable that there were few studies on medium chain length alkanes, particularly isooctane, compared to other type of organic solvents. There was a report that a mutant strain KK-211 of *Saccharomyces cerevisiae* was more tolerant to isooctane compared to its parent strain DY-1 (Miura, Zou *et al.* 2000). The reported genes in this study, were related to cell stress response, trehalose uptake, glycogen production, and membrane maintenance (Miura, Zou *et al.* 2000). Another study showed that pleiotropic drug response elements (PDREs) were also differentially expressed in the mutant strain KK211 compared to its parent strain DY-1 (Matsui, Hirayama *et al.* 2006) and point mutation in Pdr1p, a transcription factor of PDREs, improved tolerance of the parental strain, DY-1 (Matsui, Teranishi *et al.* 2008). However, isooctane-related toxicity and tolerance was only reported in yeast strains, but not in *E. coli* strains.

CHAPTER 3 METHODOLOGY AND APPROACHES

Bioprocesses certainly benefit from advanced biotechnological tools that enable a biological system to perform heterologous metabolisms and to exhibit new phenotypes. *E. coli* has been the most extensively adopted microbial host due to easy genetic manipulation and simple culture methods. However, when heterologous genes and metabolic pathways are introduced in *E. coli*, a bioprocess is limited by innate differences between *E. coli* and the cells from which those heterologous genes and pathways originate. Furthermore, it has been shown that systems-level investigation of cellular responses is important to improve bioprocesses. In the literature review above, a particular emphasis has been placed on gene disruption in quorum sensing, oxidative stress responses under anaerobic condition and hydrocarbon solvent stress responses of *E. coli*.

As introduced in Chapter I, this thesis work aims to develop a systems engineering framework for rational strain optimization that leads to improved bioprocesses. This aim has been approached primarily through characterization and reconstruction of genetic regulatory networks underlying cellular responses to specific bioprocess-relevant conditions. With the above-mentioned objectives and scope, this chapter presents how systems-biology approach was adopted in this thesis to elucidate cellular responses and how bioinformatics tools were applied to reconstitute underlying genetic regulatory networks.

3.1 Systems biology

Systems biology is a field that aims to understand a biological system with a holistic point of view by integrating all information of metabolic pathways, gene regulatory

networks and interactions between biological components, rather than studying genes and proteins individually. While the function of many genes and proteins involved in various physiological traits of microorganisms has been studied for the last several decades, it is still inconclusive how these genes and proteins interact with each other and form an organism as an integrated result of the genes and proteins. Given that a biological system is under complex regulations of numerous genes and proteins, it is important to collect, analyze and present data at a systems level, not at a single gene or protein level.

Meeting the needs of systems biology, emergence of high-throughput technologies has enabled systems biology to obtain massive data from a biological system at once. These high-throughput technologies are i) DNA sequencing, ii) Chip-based microarrays for the quantification and qualification of mRNA (transcriptomics), protein-protein interactions, protein-nucleic acid interactions (e.g. chromatin immunoprecipitation, ChIP for genomic DNAs), iii) Gas or Liquid chromatography in combination with mass spectrometry (GC-MS or LC-MS) for proteomics and metabolomics. At the same time, development of bioinformatics tools supported by computational modeling, simulation, integrated database and visualization has enabled efficient analysis and meaningful representation of huge data set possible. In this context, the field of systems biology has been collectively formulated by two different aspects. One aspect is construction of bioinformatics databases that collect and standardize individual genetic information such as genes-proteins annotations, DNA-protein interactions, protein-protein interactions, small molecules-protein interactions, enzymatic reactions and metabolic pathways. The other aspect aims to describe and predict dynamic interactions between these individual components of biological systems. The former aspect has been achieved by development of high-throughput

technologies, such as microarray chips, that enable researchers to profile global changes in proteome and transcriptome of a biological system under various conditions and stimuli. On the other hand, the latter aspect has been investigated by development of computational algorithms and mathematical models that can quantitatively and qualitatively describe interactions between various molecular components. Regulatory networks are often constructed as a result of these systems-level studies that deliver and combine the abovementioned two aspects. For example, in the field of bioprocesses, the systems-level studies have been used to identify responses networks under isobutanol exposure in *E. coli* (Brynildsen and Liao 2009; Atsumi, Wu et al. 2010), outer membrane protein networks in responses to glucose concentrations (Yang, Wang et al. 2011) and metabolic network in metabolism-perturbed *E. coli* mutants (Kim, Kim et al. 2011). Identification of regulatory networks has facilitated understanding of a biological system, and tackling of problems and bottlenecks in medicine and bioprocesses. However, this approach is often limited to several model organisms due to lack of bioinformatics databases and difficulties in standardization between different species. Moreover, even for model organisms, annotation of genetic components has not been completed. Also, integrated regulatory networks, which ultimately aim to cover all levels of regulations involved in metabolites, RNA, DNA and proteins, are too complex to be modeled while providing practical and reliable insights in a biological system. Therefore, most regulatory genome scale models still remain as static models rather than dynamic models.

This thesis has chosen DNA microarray as the means of collecting data at the systems level, and the obtained transcriptomic data were analyzed to find statistically and biologically significant cellular functions and investigate regulatory relationships between transcription factors and genes under a specific condition. In the following

Section 3.2, experimental methods for microarray, fold change analysis and GO analysis are discussed. In particular, procedures for microarray experiments, including kits, chemicals, and instruments used in the experiments, are covered in detail in Section 3.2.1. Two subsequent Sections 3.2.2 and 3.2.3 introduce the analytical steps taken for computation of fold changes and gene ontology term analysis.

3.2 Microarray, fold change analysis and GO analysis

3.2.1 Microarray experimental methods

Total RNA samples for DNA microarray experiments were prepared by RNeasy Mini Kit (Qiagen, USA). In brief, adequate amount of cells were lysed by lysozyme (Egg yolk, sigma) prepared in Tris-EDTA buffer and total RNA was purified using a column provided in the kit. DNA was removed by incubating RNA samples on the column with DNase I recombinant (RNA-free, from bovine pancreas expressed in *Pichia pastoris*, Roche). Purified total RNA was quantified by Nano Drop (Thermo), and the integrity and quality of RNA samples were verified by Bioanalyzer using Agilent RNA 6000 Nano kit (Agilent, USA).

For coupling of nucleotide samples with fluorescent dyes, cDNA, which was reverse transcribed by iScript cDNA synthesis Kit (Bio-rad), was conjugated with Cy3 and Cy5 dyes (Amersham) by Fairplay Kit (Stratagene). One of the biological replicates was labeled with Cy3 while the other biological replicate was labeled with Cy5 to minimize bias resulting from different conjugation efficiency of the two dyes. After purification of labeled cDNA using a column provided in Fairplay Kit, equal amount of Cy3-labeled RNA and Cy5-labeled RNA was mixed and hybridized with probes on the *E. coli* gene expression microarrays slide (Agilent, G4813A, 8 microarrays), which contains 15,208 probes representing *E. coli* genome. After 16 hours of incubation in a

hybridization oven, unbound RNA samples were washed away from the microarray slide in two different gene expression wash buffers (Agilent, 5188-5327, USA with Triton-X). Absolute and relative intensity of two dyes in a spot which represents a probe was scanned by microarray scanner (Agilent, USA) and features of each spot was extracted by Feature extraction software for further analysis (Agilent, USA).

3.2.2 Fold Change analysis

Microarray arbitrarily quantifies the amount of mRNA for each gene present in total RNA extracted from either control samples or treated samples. The quantified mRNA for each gene was normalized by Quantile method after background signal subtraction to obtain relative mRNA amount in controls or treated samples. Fold changes of each gene expression level were calculated by a ratio of normalized mRNA quantity from treated samples to normalized mRNA quantity of control sample as described in the following equation:

$$\frac{\text{Normalized signal intensity of a gene in treated samples}}{\text{Normalized signal intensity of a gene in control samples}}$$

The microarray experiments generated two types of data, relative expression level of genes in a sample and relative expression fold change of a gene between multiple samples. In control-treatment experiments, it was aimed to investigate the differences or changes caused by the treatment in our experimental setting. Therefore, the microarray experiments were analyzed to find the differences or changes caused by the treatment in expression level of each gene. As a result, a list of significant genes (t-test $p < 0.05$ and fold change > 2) with their fold change values and p values of t-test were generated. Based on this list, genes of interest were selected for further characterization by studying the mechanism of action via over-expression or deletion.

This approach was based on an assumption that relatively induced genes or repressed genes were responsible for the changed phenotypes or parameters to be investigated in treatment sample. Given the fact that microbial cells grow fast and promptly respond to a new stimulus, prompt transcription of genes and simultaneous translation to proteins are necessary for microbial cells to achieve the desired behavior or phenotype in a limited time. Therefore, the fold change analysis of microarray data was ideal in generating a pool of good candidate genes for further investigation.

3.2.3 Gene Ontology (GO) Term analysis

Although fold change analysis provided a list of candidate genes that might be important in the treatment experiment, an individual gene often did not represent various cellular functions where the genes were involved. To describe cellular responses at the systems level, it was necessary to deduce cellular functions that were significantly affected under a specific condition of the treatments. Towards this goal, gene ontology (GO) term analysis, which statistically analyzed GO terms related to significantly induced or repressed genes from microarray data, was performed by GeneSpring GX (Agilent, USA) or Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang, Sherman *et al.* 2009). GO terms were proposed for standardized description of genes and proteins functions, which were categorized into three domains: cellular component, molecular function and biological process. The GO term analysis provided a list of GO terms with *p*-value and list of genes associated with each GO term. Therefore, each GO term suggested cellular functions of which transcription was significantly altered in the treatment sample. Increased or decreased expression trend of genes in each GO term implied whether the identified cellular function was beneficial or detrimental to microbial cells under stress conditions. In addition, GO terms analysis enriched interpretation of microarray data

by providing more insights to overall metabolic pathways of which changes often are not apparent during the experiments.

3.3 Transcriptional Regulations and its analysis

In the previous Section 3.2.2, it was mentioned that microarray quantified changes in mRNA amount of each gene in response to the treatments. It is worthwhile to note that quantity of mRNA is mainly determined by transcriptional regulations in response to dynamic environmental signals. In other words, bacterial cells induce or repress a group of genes to promptly cope with altered conditions. Therefore, microarray data provide not only expression levels of individual genes, but also additional insights to transcriptional regulators activated or repressed by the treatment given in the experiment. In addition, identification of transcriptional regulators that represent changed cellular functions under the treatment conditions might enable us to understand microbial systems as a biological network. As a background of these rationales, the following Sections introduce transcriptional regulatory components in bacterial cells. In the Section 3.3.1, how promoters determine transcriptional strength of operon is reviewed. Section 3.3.2 introduces transcription factors as important factors determining transcription level of genes or operons. In addition, Section 3.3.3 covers further hierarchical nature of transcriptional regulation mediated by transcription factors. Lastly, Section 3.3.4 explains Network Component Analysis, a computational algorithm used to identify essential transcription factors of which activity was likely to be changed under experimental conditions.

3.3.1 Promoter and RNA polymerase

Transcription level of genes and operons are regulated by multiple mechanisms, but intrinsically, gene expression level is first determined by strength and longevity of

RNA polymerase complex binding to a certain gene. In bacteria, RNA polymerase binds to a specific DNA sequence, and therefore gene transcription is initiated when a RNA polymerase complex recognize a specific region at the upstream of a gene or an operon. Figure 3-1 shows the structure of the promoter region including -35 and -10 regions and +1 transcription initiation site. The -35 and -10 regions of promoters have conserved nucleotide sequences among various genes and operons. Promoters without conserved nucleotide sequences among various genes and operons. Promoters without conserved sequence patterns or with mutations in its original sequence have significantly changed expression level of downstream genes. Although sequences of promoter regions affect transcription efficiency of genes, it is not condition-dependent because the sequence is a fixed variable, which does not dynamically change in responses to external stimuli. Given that microbial cells change their expression level in response to various intracellular and extracellular factors, such as temperature, pressure, nutrients availability, toxic chemicals and electron acceptors, it seems essential for microbial cells to be equipped with condition-dependent regulatory mechanisms.

Figure 3-1 Schematic diagram of the promoter region

3.3.2 Role and action of transcription factors in prokaryotes

In addition to promoter sequences and RNA polymerase, transcription factors regulate transcription units by facilitating or hindering the binding of RNA polymerase to promoters. Transcription rate of each promoter is highly dependent on transcription factors and their activities. Depending on the activity of transcription factors, expression level of genes is increased or repressed by inducers. In addition, there are other small molecules that are capable of increasing or decreasing transcription rate of a gene by binding to transcription complexes. Transcription factors can be activators, repressors or both depending on the type of interactions with other transcription factors and their target gene sequences. Binding strength of transcription factors to DNA sequences is determined by their structural conformation and interaction with other transcription factors. Most transcription factors change their structural conformation by physical binding of small molecules, chemical reactions with active molecules, dimerization with other subunits and phosphorylation by intracellular kinases.

For several decades since 1950s, many studies on transcription units have focused on identification of physicochemical binding of transcription factors to specific DNA fragment. Initially, individual genes of interest were tested to find a cis-DNA element sequence as well as identification of proteins that bound in this region by gel shift assay (Kadner 1974) and DNAs foot-printing (Galas and Schmitz 1978). In addition, interaction information of proteins and DNA has been enriched by gene expression analysis, human inference based on similarity to consensus sequences and site mutation analysis (Gama-Castro, Jimenez-Jacinto *et al.* 2008). With a development of high-throughput methodologies and bioinformatics, protein-DNA interaction data have been rapidly gathered and the collection has been maintained on web-based database.

The Regulon DB (Gama-Castro, Jimenez-Jacinto *et al.* 2008) is the one of the examples, which provides a list of genes and transcriptional regulators including sigma factors and transcription factors, promoter sequences of genes, transcription factor binding sequences and effects on transcription factor binding in *E. coli* K12.

3.3.3 Hierarchy of transcriptional regulation by TF

Currently, various methods are available to collect one-to-one interactions between genes and regulatory factors. However, transcriptional regulation of a gene or an operon *in vivo* is still complicated due to hierarchical nature of transcriptional regulatory networks in biological systems. Figure 3-2 shows hierarchy of transcription factor and genes. Transcription factor and genes, represented as red and green dots, are connected to each other in a complex manner. However, the efforts to represent interactions between target genes and transcription factors as a network have provided valuable insights. For example, it was suggested that one gene was regulated by more than two transcription factors on average and *E. coli* genes were dominantly regulated by less than 10 global transcription factors. Furthermore, expression level or activities of the global transcription factors were also regulated by many other sub-local transcription factors (Balleza, Lopez-Bojorquez *et al.* 2009). It is notable that most genes were connected to more than one transcription factor, of which activity and transcription level, in turn, was regulated in interactions with other transcription factors. In addition, degeneracy of transcription factors (TF) and transcription factor binding sites (TFBS) (Balleza, Lopez-Bojorquez *et al.* 2009) allow multiple transcription factors to crowd to one cis-DNA regulatory element of a coding sequence, simultaneously or condition-dependently. The involvement of multiple TFs in gene expression adds a layer of complexity to transcriptomic regulations.

Figure 3-2 Graphical presentations of transcriptional regulatory networks (Balaji, Babu *et al.* 2007)³

Red dots are transcription regulators and green dots are genes. Blue and dark green lines represent interactions between two edges in the networks.

³This figure has been reproduced by permission of Elsevier. This article can be found in this link: <http://www.sciencedirect.com/science/article/pii/S0022283607008959>

Due to the intrinsic characteristics of transcription unit interactions described above, transcription factors simultaneously activate or repress many different genes and operons. This property of transcriptional regulation explains how microorganisms quickly reconstitute genetic regulatory networks underlying cellular responses to new environments or intracellular metabolic states. Therefore, it seemed that manipulation of one transcription factor might make global changes to expressions of multiple genes, which alter physiological responses or phenotypes as a result. However, despite the important roles of transcription factors in *ad hoc* responses to treatments, the activity of transcription factors under specific condition has not been investigated adequately.

3.3.4 Network component analysis

To investigate cellular responses at the systems level, in addition to GO term analysis, a focus has been placed on identification of statistically and biologically significant transcription factors from microarray data. To do this, firstly, it was assumed that a different set of transcription factors acts on their target genes under different conditions in responses to environmental stimuli. Secondly, it was assumed that identification and engineering of these transcriptional regulators would be a good start point of the whole phenotype engineering at the systems level.

To identify essential regulatory factors which governed expression of target genes, Network Component Analysis (NCA) (Liao, Boscolo *et al.* 2003) has been used in this study. The NCA suggested that *in vivo* activity of transcription factors could be estimated by computing relative activities from microarray data. This is because expression level of transcription factors does not correlate well with their actual activities unlike other enzymes and cellular components. Furthermore, action of

transcription factors is dependent on whether they are in active form or deactivated forms. Transcription factors bind to particular DNA sequences when they are in active forms, and they are often released from the DNA strands when they are in deactivated forms allowing genes to be transcribed.

In NCA, microarray data were presented in a matrix E , of which row was a string of relative expression intensity of a gene under M different experimental conditions (e.g. treatment concentrations and times), and of which column was a string of relative expression intensities of N genes under one experimental condition. Therefore, the size of $[E]$ was N rows $\times$ M columns. Another matrix, called a connectivity matrix A (size: $N \times L$), showed regulatory relationships between genes (N) in the microarray data and transcription factors (L). The NCA algorithm, then, decomposed the matrix E into two other matrices, A and P , as described in Eq.1:

$$[E]=[A][P] \text{ (Eq. 1)}$$

where the matrix P was to derive relative activities of transcription factors (L) under the different experimental conditions (M). As a result, a smaller matrix $[P]$ was obtained while reducing dimensionality of matrix E . This result was attributed by the fact that *E. coli* had a much smaller number of transcription factors (L) than the number of genes (N).

3.4 Physiological traits analysis

Transcriptomic changes in microorganisms often result in physiological changes, particularly morphological changes in their size, shape, surface-membrane structures and accessory microstructures. These morphological responses to environment can be

observed by various microscopic techniques. Bright field microscopy and Phase contrast microscopy are most commonly used to investigate overall shape and size of the microorganisms because they allow examining a wet specimen without extensive sample preparation processes. However, to observe fine structures such as surface or flagellar filaments of bacterial cells, other microscopic techniques such as field emission scanning electron microscopy (FESEM), transmission electron microscopy (TEM) and atomic force microscopy (AFM) are required. While FESEM and TEM need samples to be pre-treated by fixation and coating processes, AFM specializes in scanning of surfaces without any chemico-physical process. Flagellar structures have been extensively visualized by electron microscopy or fluorescent microscopy. On the other hand, surface or membrane damage caused by toxic materials is often confirmed by AFM to show changes in depth of the surface. Therefore, important morphological changes followed by transcriptomic responses can be investigated by utilizing these microscopic tools.

CHAPTER 4 The absence of the *luxS* gene increases swimming motility and flagella synthesis in *Escherichia coli* K12⁴

4.1 Abstract

Despite the significant role of S-ribosylhomocysteinase (LuxS) in the activated methyl cycle pathway and quorum sensing, the connectivity between *luxS* and other cellular functions remains incomplete. Herein, it was shown that *luxS* deletion significantly increases swimming motility and flagella synthesis in *Escherichia coli* K12 using motility, transcriptome, and scanning electron microscopy assays. Furthermore, based on the transcriptome and network component analyses, and known regulatory relations, a conceptual genetic regulatory network underlying the increased flagella synthesis in response to *luxS* deletion was proposed.

4.2 Introduction

S-ribosylhomocysteinase (LuxS) plays an important role in the activated methyl cycle pathway (AMC) and quorum sensing (QS) in many bacterial species (Winzer, Hardie *et al.* 2002), and its corresponding phenotypic traits have been a topic of intense research. The previous research in this area has mostly been focused on phenotypic changes in *luxS* mutants that undergo AI-2-dependent QS deficiency and/or metabolic shifts associated with AMC pathway disruption (Table 2-1). Our current understanding, based on those traits of *luxS* mutants, includes the involvement of *luxS* in the regulation of cell growth, biofilm formation, QS, motility, virulence, and resistance. However, the actual effects of *luxS* deletion on the phenotypic traits have

⁴ The text, figures and tables in this chapter have been reproduced by permission of Elsevier. This article can be found in this link: <http://dx.doi.org/10.1016/j.bbrc.2010.09.080>

frequently been inconclusive (Table 2-1), implying that the effects may be strain- and/or condition-dependent.

Among the phenotypic traits with inconclusive linkage to LuxS is motility whereas flagella-derived motility is one of the most essential cellular functions in bacteria. Previous studies reported *luxS* mutants undergoing a change in motility (Table 2-1), suggesting that LuxS-based metabolism and/or QS might be linked to motility. In this light, this study aims to provide a more complete linkage between the *luxS* gene and essential cellular functions with a specific emphasis on motility in *E. coli* K12. Therefore, this thesis will (i) evaluate *luxS*-mediated relevant phenotypic trait changes; (ii) determine corresponding transcriptome profiles; and (iii) verify relevant cellular mechanisms using various assays including scanning electron microscopy. This study provides the first evidence that *luxS* deletion enhances motility and flagella synthesis in *E. coli* W3110. Further, the conceptual responses genetic networks will be proposed to give new insights into *luxS*-mediated cellular functions in *E. coli*.

4.3 Materials and methods

4.3.1 Bacterial strains, media, and growth conditions

E. coli K12 W3110 (wild type, WT) and its isogenic *luxS* mutants LW12 (W3110 $\Delta luxS::Kan$) (Wang, Li *et al.* 2005) were obtained from Dr. William Bentley at University of Maryland. The cells were grown at 30 °C with shaking at 250 rpm in Luria Broth (LB-Miller, BD: 10 g Tryptone, 5 g yeast extract and 10 g NaCl per liter of distilled water) medium plus 0.8 % (w/v) glucose. Overnight cultures of both WT and *luxS* mutant cells were diluted to $OD_{600} = 0.02$ and cultured until late exponential phase ($OD_{600} = 1.5$).

4.3.2 Motility assay

Cell culture in the late exponential phase was collected at the bottom of 50 ml tubes by centrifugation, and supernatant was discarded. Remaining cell pellets were washed with motility buffer (Hesse and Kim 2009). First, for swimming motility assay, cells were loaded onto the swim plates (Wang, Li *et al.* 2005), and incubated at 30 °C for 15 h. Second, for phase contrast microscopy, the cells were diluted in motility buffer with 0.1 M glucose, and loaded onto a microscope slide with a cover slip, and 10 frames per second for 13 seconds were acquired using Axiovert 200 fluorescent microscope (Carl Zeiss) with CoolSNAP HQ camera (Photometrics). Finally, for swarming motility assay, cells were loaded onto the swarm plates (LB-Miller (Difco) containing 0.5 % (w/v) Eiken agar (Eiken Chemical, Tokyo) and 0.8 % glucose (Sigma)). The swarming motility assay plate was incubated at 37 °C for 27 h and the image of the plates was taken by Molecular Imager® Gel Doc™ XR (Bio-Rad).

4.3.3 Microarray and real-time PCR assays

Total RNA was isolated from wild-type and *luxS* mutant cells at OD₆₀₀=1.0 using RNeasy Mini Kit (Qiagen). cDNA was synthesized and labeled using FairPlay II Microarray Labeling Kit (Stratagene). Cyanine 3 (Cy 3)- or cyanine 5 (Cy 5)-labeled cDNA was purified and hybridized onto the Agilent *E. coli* Microarrays (8 × 15K) (Agilent Technologies), according to the protocols for Agilent two-color microarray-based gene expression analysis. Finally, two-color scanning, data extraction and analysis were carried out using Agilent Microarray Scanner System, Agilent Feature Extraction Software v. 9.5.3.1, and GeneSpring GX 10. Statistically differentially expressed genes were selected based on their *p* values (< 0.05, *t*-test) and fold changes (> 4) in *luxS* mutants compared to WT. Microarray experiments were performed with two biological replicates.

To validate the microarray data, quantitative real-time PCR (qRT-PCR) was employed with the primers targeting the selected four genes, *fliC*, *flhC*, *flgK* and *rrsG* as a reference gene (Table 4-1). The qRT-PCR analysis was conducted in Bio-Rad iQ5 real-time PCR detection system with SYBR Green PCR Master Mix (Applied Biosystems).

Table 4-1 List of genes and specific primers used for RT-PCR

Gene	Sense primer sequence (5'-3')	Antisense sequence (5'-3')
<i>fliC</i>	CGCTTCAGACAGGTTGGTAGTG G	GGTTTGACTGCTGGTGGTGAGG
<i>flhC</i>	TCTCTGTTCATCCAGCAGTTGTG G	ATTGGTGCGGTTTGTGAAAGTG G
<i>flgK</i>	CAATGGTTACTCACTGGTTCAG G	GAGAACGGAATGTCAGAATGCC
<i>rrsG</i>	TCAAGGGCACAACCTCCAAGTC	GGTGTAGCGGTGAAATGCGTAG

4.3.4 Dot blotting of flagellin and flagella visualization

Cell pellets were collected by centrifugation and washed with motility buffer (10 mM potassium phosphate buffer at pH 7.0, 67 mM NaCl, and 0.1 mM EDTA, (Darnton and Berg 2008). Flagellin protein was extracted from wild-type and *luxS* mutant cells according to the method described by Ibrahim and colleagues (Ibrahim, Fleet *et al.* 1985). Extracted proteins were quantified by Bradford assay and equal amount of the protein was loaded onto PVDF membrane (Bio-Rad, CA, USA). To verify FliC expression, anti-FliC (BioLegend, CA, USA) and Immunoblot GAM-AP Kit (Bio-Rad, CA, USA) were used for blotting and color development according to the

manufacturers' instructions. Signal intensity was quantified by using Quantity One (Bio-Rad, CA, USA). For Field Emission Scanning Electron Microscopy (FESEM), cells were fixed with 1.25 % glutaraldehyde in 0.1 M potassium phosphate buffer at 4 °C for overnight, and washed with motility buffer and loaded onto polyethyleneimine (PEI)-coated silicon wafer. After air-drying the cells, the cells were dehydrated in acetone at gradually increasing concentrations from 30 to 100 %, and sequentially dried in acetone: hexamethyldisilazane (HMDS) (2:1), acetone:HMDS (1:2) and 100 % HMDS. The prepared samples were coated with platinum and examined using FESEM JSM-6700F (JEOL, Japan).

4.4 Results and discussion

To determine the effects of *luxS* deletion on the motility of *E. coli* K12, the swimming and the swarming motility of *luxS* mutants and WT was examined using the semi-solid agar assay and the phase contrast microscopy. First, the semi-solid agar assay indicated that *luxS* mutants had nine- to twelve-times higher swimming motility than WT (Figure 4-1A). Second, the swimming motility videos (Supplementary data is available at <http://www.sciencedirect.com/science/article/pii/S0006291X10017778>) indicated clear difference in the swimming motility speeds between WT and *luxS* mutants. Compared to WT, *luxS* mutants actively moved and changed the directions, likely to be propelled by the flagella. The increased swimming motility in *luxS*-deficient bacteria has not been reported before although induction of twitching motility was shown in *luxS*-deficient *Photorhabdus luminescens* (Table 2-1) (Krin, Chakroun *et al.* 2006). It is noteworthy that there was no difference in the growth rate between *luxS* mutants and WT. Figure 4-1B shows that WT and *luxS* mutants have a similar swarming motility. Harshey *et al.* (Harshey and Matsuyama 1994) reported

surface-induced differentiation of *E. coli* into swarming cells, which were marked by elongated, multi-nucleate, hyper-flagellate cellular characteristics. Therefore, there is a possibility that WT might be differentiated into the flagellated swarming cells that can perform flagellum-dependent swarming after the cells were loaded on the Eiken agar plate in our study, which requires further investigation.

Figure 4-1 Swimming and swarming motility of WT and *luxS* mutants

(A) Swimming motility assay. (B) Swarming motility assay. The values are means of triplicates with Standard Errors (SE). Solid and open bars represent WT and *luxS* mutants, respectively.

Our results have shown that *luxS* deletion increased the swimming motility of *E. coli* K12. To investigate cellular mechanisms associated with these changes, the genome-wide expression profiles of WT and *luxS* mutants were examined by whole-genome *E. coli* microarrays. Briefly, the transcriptome profiling analysis revealed that 58 genes showed statistically significant changes in transcript level in *luxS* mutants compared to WT (Table 4-2). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) with accession number GSE20606. Among these differentially expressed genes, 27 and 31 genes were up-regulated and down-regulated in *luxS* mutants compared to WT, respectively.

Table 4-2 Significantly up- or down-regulated genes in *luxS* mutants

Fold changes were calculated as ratio of gene expression level in *luxS* mutant compared to wild-type *E. coli* (WT) and presented in a log 2 scale.

B. No.	Gene	Fold change	p-value	Function description
Up-regulated genes				
b1890	<i>motA</i>	11.6	0.001	proton conductor component of flagella motor
b1882	<i>cheY</i>	10.4	0.024	chemotaxis regulator transmitting signal to flagellar motor component
b1888	<i>cheA</i>	10.2	0.037	fused chemotactic sensory histidine kinase in two-component regulatory system with CheB and CheY: sensory histidine kinase/signal sensing protein
b1887	<i>cheW</i>	9.6	0.042	purine-binding chemotaxis protein
b1885	<i>tap</i>	9.1	0.014	methyl-accepting protein IV
b1889	<i>motB</i>	9.1	0.047	protein that enables flagellar motor rotation
b1883	<i>cheB</i>	8.9	0.025	fused chemotaxis regulator: protein-glutamate methylesterase in two-component regulatory system with CheA
b1923	<i>fliC</i>	8.6	0.004	flagellar filament structural protein (flagellin)
b1891	<i>flhC</i>	8.1	0.018	DNA-binding transcriptional dual regulator with FlhD
b1886	<i>tar</i>	8.0	0.018	methyl-accepting chemotaxis protein II
b1892	<i>flhD</i>	7.7	0.009	DNA-binding transcriptional dual regulator with FlhC
b1076	<i>flgE</i>	6.7	0.038	flagellar hook protein
b1881	<i>cheZ</i>	6.6	0.012	chemotaxis regulator, protein phosphatase for CheY
b1879	<i>flhA</i>	6.5	0.001	predicted flagellar export pore protein
b1924	<i>fliD</i>	6.3	0.036	flagellar filament capping protein
b1077	<i>flgF</i>	6.2	0.036	flagellar component of cell-proximal portion of basal-body rod

b1074	<i>flgC</i>	6.2	0.033	flagellar component of cell-proximal portion of basal-body rod
b1880	<i>flhB</i>	6.1	0.03	predicted flagellar export pore protein
b3525	<i>yhjH</i>	6.1	0.000	EAL domain containing protein involved in flagellar function
b1082	<i>flgK</i>	6.0	0.032	flagellar hook-filament junction protein 1
b1194	<i>ycgR</i>	5.8	0.000	Rac prophage; predicted protein
b1071	<i>flgM</i>	5.5	0.045	anti-sigma factor for FliA (sigma 28)
b1070	<i>flgN</i>	5.0	0.049	export chaperone for FlgK and FlgL
b0929	<i>ompF</i>	3.3	0.039	outer membrane porin 1a (Ia;b;F)
b4244	<i>pyrI</i>	2.6	0.03	aspartate carbamoyltransferase, regulatory subunit
b0113	<i>pdhR</i>	2.5	0.003	DNA-binding transcriptional dual regulator
b4243	<i>yjgF</i>	2.3	0.012	Keto acid-binding protein
Down-regulated genes				
b1334	<i>fnr</i>	8.7	0.035	DNA-binding transcriptional dual regulator, global regulator of anaerobic growth
b1333	<i>uspE</i>	7.6	0.043	stress-induced protein
b1343	<i>dbpA</i>	6.7	0.008	ATP-dependent RNA helicase, specific for 23S rRNA
b1353	<i>sieB</i>	6.0	0.014	Rac prophage; phage super infection exclusion protein
b1332	<i>ynaJ</i>	5.2	0.032	predicted inner membrane protein
b1337	<i>abgB</i>	5.0	0.030	predicted peptidase, aminobenzoyl-glutamate utilization protein
b1224	<i>narG</i>	4.8	0.049	nitrate reductase 1, alpha subunit
b2579	<i>yfiD</i>	4.8	0.025	pyruvate formate lyase subunit
b1344	<i>ttcA</i>	4.6	0.006	predicted C32 tRNA thiolase
b1225	<i>narH</i>	4.4	0.034	nitrate reductase 1, beta (Fe-S) subunit
b1338	<i>abgA</i>	3.6	0.017	predicted peptidase, aminobenzoyl-glutamate utilization protein
b1226	<i>narJ</i>	3.5	0.032	molybdenum-cofactor-assembly chaperone subunit (delta subunit) of nitrate reductase 1

b1341	<i>ydaM</i>	3.5	0.000	predicted diguanylate cyclase, GGDEF domain signaling protein
b3509	<i>hdeB</i>	3.4	0.002	acid-resistance protein
b3516	<i>gadX</i>	3.2	0.031	DNA-binding transcriptional dual regulator
b3510	<i>hdeA</i>	2.9	0.019	stress response protein acid-resistance protein
b3511	<i>hdeD</i>	3.0	0.014	acid-resistance membrane protein
b3544	<i>dppA</i>	2.9	0.003	dipeptide transporter
b2997	<i>hybO</i>	2.8	0.038	hydrogenase 2, small subunit
b3517	<i>gadA</i>	2.7	0.008	glutamate decarboxylase A, PLP-dependent
b3408	<i>feoA</i>	2.7	0.021	ferrous iron transporter, protein A
b0904	<i>focA</i>	2.6	0.020	formate transporter
b1493	<i>gadB</i>	2.5	0.026	glutamate decarboxylase isozyme
b3512	<i>gadE</i>	2.5	0.039	DNA-binding transcriptional activator
b1492	<i>gadC</i>	2.3	0.026	acid sensitivity protein, putative transporter
b2343	<i>yfcZ</i>	2.3	0.037	conserved protein
b3922	<i>yiiS</i>	2.2	0.024	conserved protein
b3409	<i>feoB</i>	2.2	0.015	fused ferrous iron transporter, protein B: GTP-binding protein/membrane protein
b3506	<i>slp</i>	2.2	0.025	outer membrane lipoprotein
b1345	<i>intR</i>	2.1	0.015	Rac prophage; integrase
b1723	<i>pfkB</i>	2.1	0.000	6-phosphofructokinase II

To further examine the distribution of the genes with regard to their functions, the genes with transcript level changes were classified according to Comprehensive Microbial Resource (<http://cmr.jcvi.org/cgi-bin/CMR/shared/RoleList.cgi>) (Figure 4-2A). These genes primarily belonged to the functional classes of “cell envelope”, “cellular processes”, “central intermediary metabolism”, “energy metabolism”, “regulatory functions”, “transcription”, “transport and binding proteins” and “unknown function or hypothetical”.

In line with our earlier observation that *luxS* deletion led to increased motility, the transcriptome analysis showed that the transcription levels of the genes involved in flagella synthesis and assembly were significantly increased in *luxS* mutants compared to WT (Table 4-2). First, *flhDC* genes (Class I, master operon, Figure 4-2B) were induced in *luxS* mutants about 8-fold in a log 2 scale. Further, most genes in Class II operons (i.e. *flgE*, *flgF*, *flgC*, *flhB*, *flgK*, and *flgN*), required for the structural assembly of the hook and the basal body, were up-regulated more than 5-fold in a log 2 scale in *luxS* mutant.

It was also noted that the *flhDC* complex activates transcription of Class II operons (Apel and Surette 2008). Among the genes in Class III operons, the genes encoding flagellins (*fliC*), flagellar cap proteins (*fliD*), and motor proteins (*motA* and *motB*) were up-regulated 8.6, 6.3, 11.6 and 9.1-fold in a log 2 scale, respectively. Interestingly, *flgM* encoding anti-sigma factor for FliA was induced 5.5-fold in a log 2 scale in *luxS* mutants, while the transcription level of the *fliA* gene exhibited no significant change. FlgM inhibits Class III sigma factor FliA until the functional hook-basal body structures are completed. Once FlgM is secreted out from the hook-basal body structures, the released FliA triggers transcription of Class III operons and the two motor proteins (MotA and Mot B) (Hughes, Gillen *et al.* 1993).

Figure 4-2 Functional classification of differentially expressed genes in *luxS* mutants and a schematic representation of gene regulation and flagella structure

(A) Solid and open bars represent up- and down-regulated genes in *luxS* mutant compared to WT, respectively. The number in parenthesis represents the total number of genes within the genome in each functional class. (B) Classes, regulatory pathways and assembly structure of flagella genes. Genes related to chemotaxis are underlined. (C) A schematized flagellum (Terashima, Kojima *et al.* 2008) with genes related to each substructure of the flagellum. The gray area refers to the periplasmic region between inner and outer membranes.

Second, genes involved in the chemotaxis systems were induced in *luxS* mutants. As shown in Table 4-2, methyl-accepting chemotaxis protein genes (*tap* and *tar*) and *che* genes (*cheA*, *cheW*, *cheY*, and *cheB*) were up-regulated at least 8-fold in a log 2 scale. As previously reported, chemotaxis substances bind to the receptor (Tap and/or Tar) and the signals are transmitted into a cell and are transferred through the two-component phosphor-relay system via the Che proteins. Subsequently, the phosphorylated CheY binds to FliM in the flagellar switch and promotes flagellar rotation via Mot proteins (Muff, Foster *et al.* 2007; Terashima, Kojima *et al.* 2008). Since the above-mentioned genes represent most of the operons involved in flagella synthesis/assembly and the chemotaxis (Figure 4-2C), this result suggests that their increased transcription contributed to the increased motility in *luxS* mutants.

The microarray-based transcriptome analysis showed that many genes involved in flagella synthesis were significantly up-regulated in *luxS* mutants. To validate the transcript levels obtained from the microarray analysis, qRT-PCR was first employed for the flagella-related genes (*fliC*, *flhC*, and *flgK*). Results show that our microarray data were corroborated with qRT-PCR analyses of those flagella-related genes with the fold changes of 9.74 ± 0.09 (*fliC*), 8.1 ± 0.84 (*flhC*), and 8.08 ± 0.04 (*flgK*) in a log 2 scale (Table 4-2). To further verify the expression of flagellar protein (flagellin, also FliC), dot blotting was conducted. The dot blotting result shows that *luxS* deletion increased the flagellin expression by approximately 2-fold in a log 2 scale in *luxS* mutants compared to WT (data not shown), suggesting a possibility of more or longer flagella assembled on the cell surface of *luxS* mutants.

Our phenotype, microarray, qRT-PCR, and dot blotting analyses supports the hypothesis that more flagellar proteins might be synthesized and assembled in *luxS*

mutants compared to WT. To verify whether this hypothesis is valid, the flagella morphology of WT and *luxS* mutants were examined using Field Emission Scanning Electron Microscopy (FESEM). Figure 4-3 shows flagella around *luxS* mutants while flagella were not clearly observed around WT. This FESEM result suggests that *luxS* deficiency enhanced the synthesis of flagella assembly in *E. coli* K12, which is consistent with the microarray and qRT-PCR data.

Taken together, our results show that *luxS* deletion induced not only transcription of the flagella-related genes but also translation and relocating of the flagella-related proteins for a complete assembly and normal function of the flagella. As shown in Figure 4-2B, transcription-translation, transport and assembly of flagellar components are tightly controlled in a cascade fashion with a hierarchy to meet cells' immediate needs for motility and to prevent undesirable energy costs (Terashima, Kojima *et al.* 2008). Given this controlled, cascade regulation of flagella synthesis, it was hypothesized that *luxS* deletion affected the top or the bottom of the motility regulation hierarchy rather than the continuous processes in the middle. In this light, it was further investigated the hypothesis that *luxS* deletion would induce motility either by (i) increased *flhDC* transcription and/or (ii) the decreased concentration of cyclic diguanylate (c-di-GMP) that reduces motility by binding and changing conformation of a receptor protein (Tamayo, Pratt *et al.* 2007).

Figure 4-3 Flagella visualized by Field Emission Scanning Electron Microscopy (FESEM)

(a) WT; (c) *luxS* mutants; (b) and (d), magnified images of the corresponding framed areas in (a) and (c).

Figure 4-4 illustrates genetic regulatory networks potentially associated with the increased flagella synthesis in *luxS* mutants, also depicting the above-mentioned two hypotheses. The first hypothesis which suggests that *luxS* deletion induced motility by activating *flhDC* transcription was supported by the result which demonstrated an increased expression level of *flhDC* by 8-fold in a log 2 scale in *luxS* mutants (Figure 4-4 and Table 4-2). The *flhDC* encodes a master regulator of flagella synthesis, FlhDC, which forms heterotetramers (FlhD₂C₂) and positively interacts with a sigma factor 70 (σ^{70})-RNA polymerase complex allowing the expression of genes in Class II including a FliA, a sigma factor specific for the flagellar motility genes (Figure 4-2B and Figure 4-4). Interestingly, transcription of the *flhDC* operon is regulated by multiple transcription factors of which activities are controlled by such environmental signals as osmolarity (OmpR), temperature (DnaK), pH (H-NS), carbon sources (cAMP-CRP) and the quorum sensing (QseBC) (Soutourina, Kolb *et al.* 1999; Sperandio, Torres *et al.* 2002; Smith, Knapman *et al.* 2009). Therefore, there is a possibility that part of the regulators controlled by the above signals might contribute to the increased transcription of the *flhDC* operon. To explore this possibility, I conducted Network Component Analysis (NCA), which deduces relative activities of transcription factors from transcriptome data (Liao, Boscolo *et al.* 2003). Table 4-3 shows that FlhDC, FliA, H-NS and GadX were among the regulatory proteins whose activities were most perturbed by *luxS* deletion. The relative activities of FlhDC, FliA, H-NS and GadX in *luxS* mutants were changed by 7.13, 6.68, 3.93 and 1.06, respectively, with statistical significance. Considering that FlhDC, FliA, and H-NS all play a role in flagella synthesis and motility, this outcome suggests that those three may serve an important hubs that link *luxS* deletion to increased flagella synthesis and motility in *E. coli*.

Table 4-3 Relative transcription factor (TF) activities of *luxS* mutants (L) and *E. coli* K12 wild type (W)

Difference of two activities deduced by NCA (z-test $p < 10^{-3}$) between L and W was calculated as absolute value, and the numbers in parentheses refer to standard deviations.

TF	<i>luxS</i> mutants (L)		Wild type (W)		L-W
FlhDC	2.94	(0.26)	-4.19	(0.58)	7.13
FliA	2.50	(0.56)	-4.18	(1.53)	6.68
H-NS	1.50	(1.33)	-2.40	(1.08)	3.93
FNR	-0.12	(1.15)	2.27	(0.76)	2.39
IHF	-0.15	(0.46)	-2.37	(0.86)	2.21
GadX	1.18	(0.15)	2.23	(0.22)	1.06
Lrp	-0.13	(0.35)	-1.04	(0.32)	0.90

In line with this outcome, it was observed that the most genes regulated by GadX, a transcription factor whose activity was decreased according to NCA, which regulates cellular responses to pH and acid resistance, were down-regulated in *luxS* mutants. As shown in Figure 4-4, the GadX regulon includes *hdeAB* (acid resistance proteins), *rpoS* (an alternative sigma factor S, specific for general stress responses), *uspE* (a universal stress protein, especially involved in motility), and some of the regulon members (*gadA*, *gadx*, *hdeD*) repressed by H-NS. This result may imply that cellular mechanisms to pH change are differently regulated in *luxS* mutants, in accordance with our earlier outcome that H-NS, a pH-responsive transcription factor, showed increased activity. Consequently, the transcriptome, NCA, and known regulatory network analyses support the hypothesis that *luxS* deletion affects pH-responsive mechanisms, resulting in the increased transcription of *flhDC* (Figure 4-4).

The second hypothesis suggests that c-di-GMP might be associated with the increased flagella synthesis. This hypothesis is supported by the NCA, which shows that *luxS* mutants significantly increased activity of FliA. Moreover, transcriptome analysis indicates that the genes controlling the c-di-GMP level were significantly regulated in *luxS* mutants. As mentioned above, c-di-GMP is involved in a post-transcriptional regulation of flagellar motility in the bottom level of the regulatory hierarchy (Figure 4-2 and Figure 4-4). That is, c-di-GMP, an important secondary messenger when bacterial cells enter the biofilm-forming growth stage or a stationary phase, represses motility by binding to YcgR, which interacts with the proteins composing the flagellar basal structure (Pesavento, Becker *et al.* 2008). In our transcriptome and NCA data, FliA showed a much higher activity and the transcription level of YhjH, which has a c-di-GMP degrading activity, was increased while the transcription of *ydaM*, which

encodes diguanylate cyclase synthesizing c-di-GMP, was repressed. This may possibly result in a decreased level of c-di-GMP and thus, increased motility in *luxS* mutants.

Figure 4-4 Conceptual cellular response networks related to flagella synthesis in *luxS* mutants

A regulatory network of genes (italicized in circles) and proteins (hexagons) was constructed including related metabolites and chemicals based on the known regulatory relations publicly available in RegulonDB and the literature (Pesavento, Becker *et al.* 2008; Tavender, Halliday *et al.* 2008; Terashima, Kojima *et al.* 2008; Smith and Hoover 2009; Wood 2009). The number (in bold) below the gene refers to the fold changes (in a log 2 scale) of the respective gene in *luxS* mutants compared to the wild type (a minus sign, down-regulation; a plus sign, up-regulation). In the cytoplasm, genetic regulations related to *flhDC* operon (black), flagella synthesis and chemotaxis (blue), flagella activity (red) and pH-inducible genes under GadX regulon (green) were highlighted accordingly. The schematic diagram of flagella assembly was illustrated according to the gene regulation hierarchy of flagella synthesis and chemotaxis. To describe activated methyl cycle (AMC), metabolites involved in AMC were connected by arrows in a direction of their chemical reactions (SAM, S-adenosyl methionine; SAH, S-adenosyl homocysteine; SRH, S-ribosyl homoserine). Some relations were omitted to improve the clarity of the network.

4.5 Conclusions

Overall, this study provided the first evidence that *luxS* deletion enhances swimming motility and flagella synthesis in *E. coli* K12. Further, based on the transcriptome and NCA analyses, and previously known regulatory relations, it was proposed that the increased flagella synthesis and motility were contributed to by increased *flhDC* transcription level and/or decreased c-di-GMP concentration in *luxS*-deficient *E. coli*. I am currently investigating the *luxS*-mediated regulatory mechanisms underlying this connectivity.

CHAPTER 5 Systems-level oxidative stress response of *Escherichia coli* under anaerobic condition: identification of essential regulatory factors

5.1 Abstract

Despite of extensive studies on microbial response to oxidative stress, our understanding of anaerobic oxidative response mechanisms is still obscure. Here, this study reports systems-level response of anaerobically grown *Escherichia coli* to oxidative stress. Transcriptomic profiling revealed that H₂O₂ treatment under anaerobic condition induced OxyR and SoxRS-regulated defense mechanisms. By functional annotation analysis, it was found that, under anaerobic condition, H₂O₂ treatment induced transcription of genes involved in iron homeostasis, putrescine degradation and TCA cycle, but decreased expression level of genes related to nucleotide *de novo* biosynthesis, carbamoyl phosphate synthesis and anaerobic respiration. In relation to anaerobic respiration, ArcB and FNR were further identified as important transcription factors in oxidative stress responses of *E. coli* under anaerobic condition, of which deletion increased viability by 89-fold and decreased viability by 131-fold respectively. Moreover, it has been reported that oxidative stress caused by H₂O₂ significantly increases NADH: NAD⁺ ratio under anaerobic condition, which suggests that NADH: NAD⁺ ratio might play a role in determining cellular metabolisms during oxidative stress responses. Interestingly, supply of NADH increased tolerance of *E. coli* against H₂O₂ by 10-fold under aerobic condition, but not under anaerobic condition.

5.2 Introduction

Oxidative stress is a biological stress imposed by reactive oxygen species (ROS) such as H₂O₂, superoxide and hydroxyl radicals. ROS damage important macromolecules in microbial cells, mutate nucleic acids, and modify normal functions of proteins (Imlay 2008). To protect against ROS, microbial cells have cellular machineries that remove ROS and repair damaged cellular components. However, damaged nucleic acids and proteins significantly inhibit the growth and decrease the viability of microbial cells (Sedgwick, Ho *et al.* 1991), major limiting factors that drops the productivity of bioprocesses, such as anaerobic fermentations and microbe-mediated catalytic reactions (Zhu, Thakker *et al.* 2011; Kang and Chang 2012).

Biosynthesis of natural products and valuable chemicals has resolved some drawbacks of chemical synthesis, including a low reaction specificity and yield (Chang and Keasling 2006). Particularly, anaerobic metabolisms and process have been exploited for biological production of ethanol(Inui, Kawaguchi *et al.* 2004; Koskinen, Beck *et al.* 2008), isobutanol (Trinh, Li *et al.* 2011), hydrogen and methane gases using anaerobic reactor (Luo, Xie *et al.* 2011) and amino acids (Litsanov, Brocker *et al.* 2012). However, anaerobic biological processes frequently suffer from toxicity generating from oxidative stress because it is difficult to maintain strict anaerobic condition, and some anaerobic microbes are highly susceptible to oxygen and ROS (Imlay 2002). Despite prior reports on the deleterious effects of oxygen on anaerobic processes and anaerobes (Lin and Wang 1991; Upadhyaya, Clancy *et al.* 2012), few studies have reported systems-level microbial responses to oxidative stress under anaerobic condition and the effects of oxidative stress on anaerobic metabolic pathways. Certainly, improved anaerobic bioprocesses require a thorough

understanding of cellular responses to oxidative stress specifically under anaerobic condition because microbial cells utilize different respiratory pathways under aerobic and anaerobic conditions (Unden and Bongaerts 1997). Consequently, in this study, systems-level cellular responses to oxidative stress will be investigated first in *Escherichia coli* under anaerobic condition using transcriptome and bioinformatics analyses. Based on these analyses, I then identify cellular functions and important regulatory factors, including transcription factors and cofactors, involved in the tolerance of *E. coli* to oxidative stress under anaerobic condition. These regulatory factors will provide helpful insights into developing engineering strategies to improve microbial tolerance to oxidative stress under anaerobic condition.

5.3 Materials and Methods

5.3.1 Strains, growth condition and H₂O₂ treatment

Escherichia coli K12 wild type, W3110 (The Coli Genetic Stock Center #4474) was investigated in this study. Its isogenic mutants were constructed by P1 transduction from isogenic derivatives of *E. coli* K12 BW25113. Kanamycin resistance cassette insertion and deletion of each target gene was confirmed by PCR using primers specific to the kanamycin resistance cassette and primers that bind to flanking locus at the upstream or the downstream of a deleted gene. Each single colony was inoculated to 1 ml LB and incubated overnight at 37 °C with shaking at 250 rpm. The overnight culture was introduced to the anaerobic chamber (Coy, USA under 95 % N₂ and 5 % H₂) and diluted to OD₆₀₀ of 0.01 (more than 200 dilution times) in anaerobic M9 minimal medium (Na₂HPO₄ 33.9 g/L, KH₂PO₄ 15 g/L, NaCl 2.5 g/L and NH₄Cl 5 g/L) with supplements (0.4 % glucose, 0.4 % glycerol, 0.2 % casamino acid, 1mM thiamine, 2 mM MgSO₄, 0.1 mM CaCl₂, 0.025 % L-cysteine-HCl and 40 mM NaNO₃). A serum bottle containing anaerobic cell culture was sealed with a butyl

rubber stopper and an aluminum cap, and incubated at 37 °C without shaking. *E. coli* was treated with H₂O₂ at log phase of the growth (OD₆₀₀ = 0.1).

5.3.2 RNA extraction and microarray experiments

E. coli K12 W3110 was grown in the anaerobic M9 medium with same supplements mentioned above. At mid-exponential phase, half of the cell culture was incubated with 1 mM H₂O₂ for 10 min at 37 °C, and the rest was incubated without the H₂O₂ treatment under the same condition. After the treatment, the cells were collected on a 0.2 µm filter membrane using a syringe. To stabilize gene expression profiles before cells are exposed to the air, the filter membranes were submerged in 1.5 ml reagents solution (two volume of Bacteria RNA protect reagent (Qiagen, USA) and one volume of the medium). All above 4 steps, cell culture preparation, H₂O₂ treatment, cell harvesting and RNA profile stabilization, were performed in the anaerobic chamber. RNA was extracted by Rneasy Mini Kit (Qiagen), and RNA integrity was analyzed by Agilent bioanalyzer and prokaryote total RNA Nano 6000 kit. cDNA synthesis (Stratagene Fairplay III, USA), hybridization (Agilent microarray hybridization kit, USA) and microarray washing were performed according to manufacturers' instructions. A fluorescent signal from each spot was captured and recorded by Agilent Microarray Scanner and Agilent Feature Extraction Software v.9.5.3.1 respectively.

5.3.3 Data analysis

Microarray data was analyzed by GeneSpring GX11 (Agilent, USA). Signals from each gene spot were normalized by Quantile method after background signal subtraction. Significantly up- or down-regulated genes (T-test, |fold change| > 2, $p < 0.05$) were used for Table 5-2 and Table 5-3. The microarray data obtained from

anaerobic condition was further analyzed by two-way ANOVA together with the another data set, of which control had previously been used for a study (Ling, Kang *et al.* 2010) while treatments were exposed to 10 mM and 30 mM H₂O₂ for 30 min under aerobic condition. Expression intensities of significant genes (ANOVA $p < 0.01$ with multiple testing correction) were calculated as a ratio of signals of treated samples to signals of untreated samples, and genes of which fold change is bigger than 2 were selected by t-test in significance level $p < 0.05$. These significant genes were used for gene ontology term analysis ($p < 0.05$), DAVID functional annotation analysis (Huang, Sherman *et al.* 2009) and network component analysis (NCA) (Huang, Sherman *et al.* 2009). Only statistically significant genes were mentioned in the text, tables and figures.

5.3.4 NADH and NAD⁺ quantification

E. coli cell culture was prepared under same condition as described in the paragraph, ‘Strains, growth condition and H₂O₂ treatment’. For NADH and NAD⁺ quantification from the aerobic cultures, 5 ml of cell culture was collected by centrifugation and washed with 1ml of PBS, and 0.5 ml each cell culture was used for extraction of NADH and NAD⁺. Intracellular NADH and NAD⁺ was extracted by modified Acid-Base extraction method (Leonardo, Dailly *et al.* 1996). In brief, 0.4 M NaOH or 0.4 M HCl was added to cell pellets to extract NADH and NAD⁺ respectively. Acid- or base-treated cell pellets were heated at 100 °C for 10 min. Equal volume of 0.4 M HCl or 0.4 M NaOH was added to 0.4 M NaOH-treated sample and 0.4 M HCl-treated sample respectively to neutralize the samples. The neutralized samples were centrifuged at 16,200 g for 5 min at 4 °C to remove cell debris and the supernatant was used for quantification of NADH and NAD⁺, which was done by NADH/NAD⁺ quantification kit (Bio Vision, K337-100).

5.4 Results and discussion

5.4.1 Systems-level cellular response to oxidative stress under anaerobic condition

To understand cellular responses to anaerobic oxidative stress, transcriptome was analyzed by microarray upon exposure to H₂O₂ under anaerobic condition. Because aerobic oxidative response mechanisms are well documented, first the transcriptional changes were examined with focus on the genes reportedly involved in oxidative stress response under aerobic condition, mostly regulated by the transcription factor OxyR (Imlay 2008). Table 5-1 shows that these OxyR-regulated genes were also significantly up-regulated upon exposure to H₂O₂ under anaerobic condition. These genes encode proteins associated with H₂O₂ scavenging (*katG*), heme synthesis (*hemH*), iron-sulfur cluster assembly (*suf* regulon), iron homeostasis (*fur*), reduction of disulfide bond (*ahpF*) and redox homeostasis (*grxA*, *trxC* and *gor*). Other up-regulated genes include the genes encoding superoxide dismutase (*sodA*) and fumarate hydratase (*fumC*), regulated by the SoxRS two-component system, and the genes related to SOS response and in particular, DNA repair.

Table 5-1 List of genes regulated by OxyR, SoxRS or genes involved in SOS response, of which transcription was significantly changed by H₂O₂ under anaerobic condition ($p < 0.05$, |Fold change| > 2)

b. No.	Gene	Gene product name	<i>p</i> value [#]	Fold change [§]
OxyR				
b0475	<i>hemH</i>	ferrochelatase	0.005	12.7
b0606	<i>ahpF</i>	alkyl hydroperoxide reductase, F52a subunit, FAD/NAD(P)-binding	0.004	36.7
b0683	<i>fur</i>	DNA-binding transcriptional dual regulator of siderophore biosynthesis and transport	0.041	2.0
b0849	<i>grxA</i>	glutaredoxin 1, redox coenzyme for ribonucleotide reductase (RNR1a)	0.002	99.3
b1679	<i>sufE</i>	sulfur acceptor protein	0.000	112.7
b1680	<i>sufS</i>	selenocysteine lyase, PLP-dependent	0.000	117.3
b1681	<i>sufD</i>	component of SufBCD complex	0.000	249.5
b1682	<i>sufC</i>	component of SufBCD complex, ATP-binding component of ABC superfamily	0.001	256.9
b1683	<i>sufB</i>	component of SufBCD complex	0.001	394.3
b1684	<i>sufA</i>	Fe-S cluster assembly protein	0.002	665.0
b2392	<i>mntH</i>	manganese/divalent cation transporter	0.001	60.1
b2582	<i>trxC</i>	thioredoxin 2	0.002	66.8
b3500	<i>gor</i>	glutathione oxidoreductase	0.002	2.6
b3518	<i>yhjA</i>	predicted cytochrome C peroxidase	0.045	3.7
b3942	<i>katG</i>	catalase/hydroperoxidase HPI(I)	0.021	24.7
SoxRS				
b0462	<i>acrB</i>	multidrug efflux system protein	0.005	2.1
b0683	<i>fur</i>	DNA-binding transcriptional dual regulator of siderophore biosynthesis and transport	0.041	2.0
b0850	<i>ybjC</i>	predicted inner membrane protein	0.010	-2.6
b0851	<i>nfsA</i>	nitroreductase A, NADPH-dependent,	0.010	-2.8

		FMN-dependent		
b0852	<i>rimK</i>	ribosomal protein S6 modification protein	0.016	-3.0
b0853	<i>ybjN</i>	predicted oxidoreductase	0.003	-2.2
b0950	<i>pqiA</i>	paraquat-inducible membrane protein A	0.026	-4.1
b1530	<i>marR</i>	DNA-binding transcriptional repressor of multiple antibiotic resistance	0.008	3.4
b1531	<i>marA</i>	DNA-binding transcriptional dual activator of multiple antibiotic resistance	0.038	3.8
b1532	<i>marB</i>	predicted protein	0.017	3.2
b1611	<i>fumC</i>	fumarate hydratase (fumarase C), aerobic Class II	0.000	81.1
b2159	<i>nfo</i>	endonuclease IV with intrinsic 3'-5' exonuclease activity	0.001	10.5
b2895	<i>fldB</i>	flavodoxin 2	0.019	2.4
b3908	<i>sodA</i>	superoxide dismutase, Mn	0.007	102.5
b3924	<i>fpr</i>	ferredoxin-NADP reductase	0.002	7.3
b4062	<i>soxS</i>	DNA-binding transcriptional dual regulator	0.002	18.2
b4063	<i>soxR</i>	DNA-binding transcriptional dual regulator, Fe-S center for redox-sensing	0.008	10.1
<hr/> SOS response				
b0060	<i>polB</i>	DNA polymerase II	0.001	7.1
b0231	<i>dinB</i>	DNA polymerase IV	0.007	6.7
b0779	<i>uvrB</i>	excinulease of nucleotide excision repair, DNA damage recognition component	0.005	8.4
b0799	<i>dinG</i>	ATP-dependent DNA helicase	0.006	2.5
b0958	<i>sulA</i>	SOS cell division inhibitor	0.007	6.1
b1183	<i>umuD</i>		0.033	2.3
b1184	<i>umuC</i>		0.031	2.1
b1728	<i>ydjM</i>	predicted inner membrane protein regulated by LexA	0.001	5.2
b1848	<i>yebG</i>	conserved protein regulated by LexA	0.031	7.1

b1860	<i>ruvB</i>	ATP-dependent DNA helicase, component of RuvABC resolvosome	0.016	2.8
b2009	<i>sbmC</i>	DNA gyrase inhibitor	0.005	2.9
b2616	<i>recN</i>	recombination and repair protein	0.003	12.5
b2699	<i>recA</i>	DNA strand exchange and recombination protein with protease and nuclease activity	0.004	6.8
b3631	<i>rfaG</i>	glucosyltransferase I	0.041	2.4
b3813	<i>uvrD</i>	DNA-dependent ATPase I and helicase II	0.005	2.3
b4058	<i>uvrA</i>	ATPase and DNA damage recognition protein of nucleotide excision repair excinuclease UvrABC	0.009	7.2
b4347	<i>symE</i>	conserved protein	0.001	7.7

To better elucidate the roles of those OxyR-, SoxRS- and SOS-related genes in oxidative response under anaerobic condition, their expression levels were compared with those obtained under aerobic condition. Figure 5-1 shows that the abovementioned genes were mostly induced by H₂O₂ under both conditions but the expression fold changes were much larger under anaerobic condition. In particular, the expression level of the *suf* regulon (*sufA*, *sufB*, *sufC*, *sufD*, *sufS* and *sufE*) was most significantly induced under anaerobic condition, compared to the expression level obtained under aerobic condition. Further, the expression levels of the genes in Figure 5-1 were much lower prior to exposure to H₂O₂ under anaerobic condition than under aerobic condition. This result might be attributed to the fact that ROS are routinely generated by aerobic respiration under aerobic condition, where oxygen molecules and electron transfer flavoenzymes are involved (Imlay 2008).

Figure 5-1 Heat maps of expression levels of genes involved in defense mechanisms against oxidative stress

The genes in this figure include OxyR-regulons, SoxRS-regulons and SOS response obtained under aerobic condition and under anaerobic condition without (C, control) or with H₂O₂ treatment (T, treatment) with fold changes (FC).

Next, to identify cellular functions significantly perturbed by H₂O₂ under anaerobic condition, functional annotations were analyzed for the genes that showed significant expression level changes under anaerobic condition. Further, the expression levels of the genes that underwent the functional annotation analysis were compared with those under aerobic condition. These functional annotations originate from KEGG pathways and gene ontology (GO) terms that describe cellular components, molecular functions and biological processes. Table 5-2 shows statistically significant GO terms ($p < 0.05$, with multiple testing corrections) and 10 most statistically significant KEGG pathways ($p < 0.05$). The functional annotation analysis suggests that the cellular mechanisms perturbed by H₂O₂ under anaerobic condition include the following: “cellular respiration (GO:0045333)”, “iron ion transport (GO0006826)”, “cellular nitrogen compound biosynthetic process (GO:00044271)”, “nitrogen metabolism (ecj000910)”, “alanine, aspartate and glutamate metabolism (ecj000250)”, “citrate cycle (ecj00020)”, “purine metabolism (ecj00230)” and “pyrimidine metabolism (ecj00240)”.

Table 5-2 GO terms and KEGG pathways analyzed by GO analysis from GeneSpring GX and DAVID functional annotation analysis.

Accession Number	Annotation Terms	<i>p</i> -value
<i>GO terms</i>		
GO:0045333	cellular respiration	2.07E-04
GO:0009061	anaerobic respiration	7.24E-04
GO:0015980	energy derivation by oxidation of organic compounds	7.24E-04
GO:0006091	generation of precursor metabolites and energy	0.005289
GO:0006826		
GO:0015681	iron ion transport	0.007962
GO:0000041	transition metal ion transport	0.007962
GO:0015674	di-, tri-valent inorganic cation transport	0.007962
GO:0009055		
GO:0009053		
GO:0009054	electron carrier activity	0.012791
GO:0044271	cellular nitrogen compound biosynthetic process	0.017399
GO:0030001	metal ion transport	0.042526
<i>KEGG pathways</i>		
ecj02010	ABC transporters	1.33E-43
ecj00910	Nitrogen metabolism	4.78E-19
ecj02020	Two-component system	7.26E-19
ecj00250	Alanine, aspartate and glutamate metabolism	7.32E-19
ecj00020	Citrate cycle (TCA cycle)	6.67E-16
ecj00230	Purine metabolism	7.57E-15
ecj00630	Glyoxylate and dicarboxylate metabolism	2.32E-14
ecj00240	Pyrimidine metabolism	1.19E-12
ecj00520	Amino sugar and nucleotide sugar metabolism	1.62E-12
ecj00650	Butanoate metabolism	2.60E-10

5.4.1.1 Class of iron ion transport

Table 5-3 shows that most genes in the GO term “iron ion transport (GO:0015681)” were significantly up-regulated by H₂O₂ under anaerobic condition. These genes regulate iron homeostasis; their encoding proteins participate in storage (*bfd* and *bfr*), transport (*exbD*, *fecR*, *fepABCDG*, *fhuDE*, *fiu* and *entCERBA*) and release (*fes* and *bfd*) of iron in *E. coli*. It was noticed that the ferric iron uptake transcription repressor (Fur) is a transcription regulator of many genes in Table 5-3, of which expression levels were highly increased. It was further observed that the *fur*-deletion significantly decreased viability of *E. coli* upon exposure to H₂O₂ under anaerobic condition (data not shown), implying that the *fur* gene is essential for survival of *E. coli* against oxidative stress under anaerobic condition. The essentiality of *fur* gene is further supported by prior studies reporting that Fur controls various cellular functions such as oxidative stresses (Vassinova and Kozyrev 2000), tricarboxylic acid (TCA) cycle and iron homeostasis (Zheng, Wang *et al.* 2001).

Table 5-3 List of genes in “iron ion transport (GO:0006826 and GO:0015681)” with fold changes (FC) by H₂O₂ under anaerobic condition.

b. No.	Gene	Description	Fold change
b1102	<i>fhuE</i>	outer membrane receptor for ferric iron uptake	19.3
		outer membrane protein receptor for ferrichrome, colicin	
b0150	<i>fhuA</i>	M, and phages T1, T5, and phi80	21.8
		ATP-binding component of hydroxymate-dependent iron	
b0151	<i>fhuC</i>	transport	11.8
		outer membrane receptor for iron-regulated colicin I	
b2155	<i>cirA</i>	receptor; porin; requires <i>tonB</i> gene product	17.9
b4292	<i>fecR</i>	regulator for <i>fec</i> operon, periplasmic	5.0
b4293	<i>fecI</i>	probable RNA polymerase sigma factor	9.3
b0805	<i>fii</i>	putative outer membrane receptor for iron transport	14.2
		energy transducer; uptake of iron, cyanocobalamin;	
b1252	<i>tonB</i>	sensitivity to phages, colicins	2.8
b0584	<i>fepA</i>	outer membrane receptor for ferric enterobactin	4.8
b0585	<i>fes</i>	enterochelin esterase	6.4
b0586	<i>entF</i>	ATP-dependent serine activating enzyme	7.3
b0588	<i>fepC</i>	ATP-binding component of ferric enterobactin transport	12.9
b0589	<i>fepG</i>	ferric enterobactin transport protein	25.3
b0590	<i>fepD</i>	ferric enterobactin	14.5
b0592	<i>fepB</i>	ferric enterobactin	12.9
		isochorismate hydroxymutase 2, enterochelin	
b0593	<i>entC</i>	biosynthesis	35.4
b0594	<i>entE</i>	2,3-dihydroxybenzoate-AMP ligase	34.7
		2,3-dihydro-2,3-dihydroxybenzoate synthetase,	
b0595	<i>entB</i>	isochorismatase	15.7
		2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase,	
b0596	<i>entA</i>	enterochelin biosynthesis	10.3
b3336	<i>bfr</i>	bacterioferrin, an iron storage homoprotein	8.2
b3337	<i>bfd</i>	orf, hypothetical protein	41.4

5.4.1.2 Class of nitrogen metabolism and pyrimidine metabolism

Table 5-2 shows that the classes of “nitrogen metabolism (ecj000910)” and especially “alanine, aspartate and glutamate metabolism (ecj000250)” were significantly perturbed by H₂O₂ under anaerobic condition. Intriguingly, as shown in Figure 5-2A, the heat map analysis suggests that many genes involved in aspartate and glutamate metabolic pathways were differently regulated by H₂O₂ treatment between aerobic (a) and anaerobic (b) conditions. In particular, genes associated with putrescine degradation (*puuABCDE*) and carbamoyl phosphate synthesis reaction (*carAB*) (Figure 5-2B) showed contrasting expressional changes under aerobic and anaerobic conditions. That is, *puuABCDE* and *carAB* were significantly down- and up-regulated, respectively, under aerobic condition, whereas their regulation directions were opposite under anaerobic condition. The carbamoyl phosphate produced by carbamoyl-phosphate synthetase coded in *carAB* is an important metabolite in nitrogen metabolism, together with L-aspartate and L-glutamine, which accept a nitrogen atom from ammonia (NH₃). L-aspartate, L-glutamate and carbamoyl phosphate transfer amide groups to provide cells with various amino acids and essential intermediates, which connect nitrogen metabolism to carbon metabolism. Figure 5-2B also highlights that carbamoyl phosphate is a committed intermediate for arginine biosynthesis and pyrimidine biosynthesis.

Figure 5-2 Heat maps (A) of relative expression levels and metabolic pathways (B) of the genes in the GO term “alanine, aspartate and glutamate metabolism (ecj000250)”

Expression data and metabolic pathways were referred to the microarray data and KEGG pathways, respectively. Significantly increased or decreased expression level of each gene was depicted as red-up or blue-down arrows. Expression data were obtained under aerobic condition (a) or anaerobic condition (b), without (C, control) or with H_2O_2 treatment (T, treatment) with fold changes of T over C. Red areas indicate genes and pathways involved in putrescine degradation, while blue areas indicate genes and pathways involved in carbamoyl phosphate synthesis reaction.

Interestingly, pyrimidine biosynthesis genes showed a similar trend in expression level changes with *carA* and *carB*. Table 5-4 shows that most genes involved in *de novo* purine biosynthesis were down-regulated by H₂O₂ under both aerobic and anaerobic conditions (*purKETFCMNIDH*), while pyrimidine biosynthesis-specific genes (*carAB*, *cmk*, *pyrBCDGH*, *tdk*, *thyA*, and *udp*) were up-regulated under aerobic condition, but down-regulated under anaerobic condition. This result suggests that nitrogen metabolism including pyrimidine biosynthesis pathways were part of cellular response to H₂O₂, but the regulation directions were opposite between aerobic and anaerobic conditions.

Table 5-4 List of genes in “purine metabolism (ecj00230)” and “pyrimidine metabolism (ecj00240)” with fold changes (FC) by H₂O₂ under aerobic or anaerobic condition

b. No.	Gene	Description	Aerobic FC	Anaerobic FC
b0032	<i>carA</i>	carbamoyl-phosphate synthetase, glutamine	79.0	-51.6
b0033	<i>carB</i>	carbamoyl-phosphate synthase large subunit	80.0	-12.7
b0104	<i>guaC</i>	GMP reductase	1.3	-3.9
b0171	<i>pyrH</i>	uridylate kinase	2.6	-5.7
b0184	<i>dnaE</i>	DNA polymerase III, alpha subunit	-2.6	-4.2
		DNA polymerase III, tau and gamma		
b0470	<i>dnaX</i>	subunits; DNA elongation factor III	-3.1	-1.4
		adenylate kinase activity; pleiotropic effects		
		on glycerol-3-phosphate acyltransferase		
b0474	<i>adk</i>	activity	1.3	-5.7
b0512	<i>allB</i>	putative hydrolase	-4.0	2.9
		phosphoribosylaminoimidazole carboxylase		
b0522	<i>purK</i>	= AIR carboxylase, CO	-6.9	-6.9
		phosphoribosylaminoimidazole carboxylase		
b0523	<i>purE</i>	= AIR carboxylase, catalytic subunit	-7.6	-18.5
b0640	<i>holA</i>	DNA polymerase III, delta subunit	-3.0	-2.5
b0910	<i>cmk</i>	cytidylate kinase	2.1	-4.8
b0945	<i>pyrD</i>	dihydro-orotate dehydrogenase	1.9	-56.2
b1062	<i>pyrC</i>	dihydro-orotase	4.2	-12.1
b1131	<i>purB</i>	adenylosuccinate lyase	1.5	-4.2
b1238	<i>tdk</i>	thymidine kinase	1.4	-12.4
b1623	<i>add</i>	adenosine deaminase	3.7	2.0
b1676	<i>pykF</i>	pyruvate kinase I	2.0	-3.2
		phosphoribosylglycinamide		
b1849	<i>purT</i>	formyltransferase 2	-10.5	-4.8
b2143	<i>cdd</i>	cytidine/deoxycytidine deaminase	-2.5	1.9
		ribonucleoside-diphosphate reductase 1, beta		
b2235	<i>nrdB</i>	subunit, B2	1.7	1.8
		amidophosphoribosyltransferase = PRPP		
b2312	<i>purF</i>	amidotransferase	-4.3	-6.7
b2476	<i>purC</i>	phosphoribosylaminoimidazole-	-10.7	-9.5

		succinocarboxamide synthetase = SAICAR synthetase		
		phosphoribosylaminoimidazole synthetase =		
b2499	<i>purM</i>	AIR synthetase	-16.1	-18.0
		phosphoribosylglycinamide		
b2500	<i>purN</i>	formyltransferase 1	-5.6	-13.2
b2508	<i>guaB</i>	IMP dehydrogenase	-1.1	-10.0
b2518	<i>ndk</i>	nucleoside diphosphate kinase	3.5	-4.8
		phosphoribosylformyl-glycineamide		
b2557	<i>purL</i>	synthetase = FGAM synthetase	-7.0	-2.5
		ribonucleoside-diphosphate reductase 2,		
b2675	<i>nrdE</i>	alpha subunit	3.4	71.0
		ribonucleoside-diphosphate reductase 2, beta		
b2676	<i>nrdF</i>	chain, frag	2.1	24.3
b2744	<i>surE</i>	survival protein	1.6	-2.4
b2780	<i>pyrG</i>	CTP synthetase	3.1	-5.0
b2827	<i>thyA</i>	thymidylate synthetase	1.9	-2.0
b3831	<i>udp</i>	uridine phosphorylase	3.2	3.3
		phosphoribosylglycinamide synthetase =		
b4005	<i>purD</i>	GAR synthetase	-9.9	-5.2
		phosphoribosylaminoimidazolecarboxamidef		
		ormyltransferase = AICAR		
b4006	<i>purH</i>	formyltransferase; IMP cyclohydrolase	-13.7	-12.7
		anaerobic ribonucleoside-triphosphate		
b4238	<i>nrdD</i>	reductase	-2.7	-2.1
		aspartate carbamoyltransferase, catalytic		
b4245	<i>pyrB</i>	subunit	4.5	-19.7
b4374	<i>yjjG</i>	putative phosphatase	3.3	-3.3
b4383	<i>deoB</i>	phosphopentomutase	4.1	4.7
b4384	<i>deoD</i>	purine-nucleoside phosphorylase	3.6	2.7

5.4.1.3 Class of TCA cycle and cellular respiration

Figure 5-3 shows that most of the genes involved in TCA cycle were up-regulated by H₂O₂ under anaerobic condition, but not under aerobic condition. These genes encode pyruvate dehydrogenase (*aceF*; pyruvate to acetyl-CoA), citrate synthase (*gltA*; acetyl-CoA to citrate), iso-citrate hydrolyase (*acnAB*; citrate to iso-citrate), iso-citrate dehydrogenase (*icd*; iso-citrate to 2-oxoglutarate), malate dehydrogenase (*mdh*; oxaloacetate to malate) and fumarase C (*fumC*; malate to fumarate). In relation to this result, GO terms associated with respiration and energy generation pathways, including “cellular respiration (GO: 0045333)” and “energy derivation by oxidation of organic compounds (GO:0015980)”, were significantly affected by H₂O₂ under anaerobic condition. In particular, Table 5-5 shows that 52 out of 79 genes of “anaerobic respiration (GO:0009061)” were significantly down-regulated by H₂O₂ under anaerobic condition. Out of those 52 genes, 31 genes were down-regulated also under aerobic condition. These down-regulated genes are involved in electron transport in respiratory chains, encoding nitrate reductase (*narGH*, *narZ*, *napBDFH*), nitrite reductase (*nirBD*, *nrfABCG*), NADH:ubiquinone oxidoreductase (*nuoBE*), menaquinone biosynthesis (*menD* and *menF*) and hydrogenase (*hyaABCDEF*, *hybBO* and *hypCD*).

Figure 5-3 Heat maps of relative expression levels and metabolic pathways of genes involved in the TCA cycle

Expression data and metabolic pathways were referred to the microarray data and KEGG pathways, respectively. Significantly increased or decreased expression level of each gene was depicted as red-up or blue-down arrows, while empty arrows represented statistically insignificant changes. Expression data are obtained under aerobic condition (A) and anaerobic condition (B) without (C, control) or with H_2O_2 treatment (T, treatment) with fold changes (FC).

Table 5-5 List of genes in “Cellular respiration (GO: 0045333)” with fold changes (FC) by H₂O₂ under aerobic or anaerobic condition

B. No.	Gene	Description	Aerobic FC	Anaerobic FC
b0114	<i>aceE</i>	pyruvate dehydrogenase	1.9	3.2
b0115	<i>aceF</i>	pyruvate dehydrogenase	1.5	3.9
b0116	<i>lpd</i>	lipoamide dehydrogenase	2.1	3.1
b0118	<i>acnB</i>	aconitate hydratase B	3.9	4.3
b0720	<i>gltA</i>	citrate synthase	-1.2	6.3
		putative pyruvate formate-lyase 2		
b0824	<i>ybiY</i>	activating enzyme	-12.6	-2.4
b0851	<i>nfsA</i>	modulator of drug activity A	5.3	-2.9
		anaerobic dimethyl sulfoxide reductase		
b0894	<i>dmsA</i>	subunit A	-4.6	-2.6
		anaerobic dimethyl sulfoxide reductase		
b0895	<i>dmsB</i>	subunit B	-6.3	-2.0
b0972	<i>hyaA</i>	hydrogenase-1 small subunit	-3.9	-2.9
b0973	<i>hyaB</i>	hydrogenase-1 large subunit	-3.5	-2.2
		probable Ni/Fe-hydrogenase 1 b-type		
b0974	<i>hyaC</i>	cytochrome subunit	-2.7	-3.6
b0975	<i>hyaD</i>	processing of HyaA and HyaB proteins	-4.5	-3.3
b0976	<i>hyaE</i>	processing of HyaA and HyaB proteins	-2.6	-3.2
		nickel incorporation into hydrogenase-1		
b0977	<i>hyaF</i>	proteins	-7.9	-9.7
		response transcriptional regulator for		
b0995	<i>torR</i>	torA	-3.2	2.2
		trimethylamine N-oxide reductase,		
b0996	<i>torC</i>	cytochrome c-type subunit	4.6	1.4
b1109	<i>ndh</i>	respiratory NADH dehydrogenase	3.8	4.6
		isocitrate dehydrogenase, specific for		
b1136	<i>icd</i>	NADP+	-1.4	3.0
b1224	<i>narG</i>	nitrate reductase 1, alpha subunit	-14.5	-4.3
b1225	<i>narH</i>	nitrate reductase 1, beta subunit	-17.9	-3.3

		nitrate reductase 1, delta subunit,		
b1226	<i>narJ</i>	assembly function	-3.2	-1.5
b1227	<i>narI</i>	nitrate reductase 1, cytochrome b	-6.0	-1.4
b1276	<i>acnA</i>	aconitate hydratase 1	-1.8	23.6
		transcriptional regulation of aerobic,		
b1334	<i>fnr</i>	anaerobic respiration, osmotic balance	3.0	3.4
b1468	<i>narZ</i>	cryptic nitrate reductase 2, alpha subunit	-6.6	-2.5
		formate dehydrogenase-N, nitrate-		
b1474	<i>fdnG</i>	inducible, alpha subunit	1.1	-1.0
		formate dehydrogenase-N, nitrate-		
b1475	<i>fdnH</i>	inducible, iron-sulfur beta subunit	-2.8	1.5
		formate dehydrogenase-N, nitrate-		
b1476	<i>fdnI</i>	inducible, cytochrome B556	-1.1	1.6
b1478	<i>adhP</i>	alcohol dehydrogenase	2.8	-1.7
		putative DMSO reductase anchor		
b1590	<i>ynfH</i>	subunit	-2.8	1.1
		fumarase C= fumarate hydratase Class		
b1611	<i>fumC</i>	II; isozyme	1.2	100.1
b1676	<i>pykF</i>	pyruvate kinase I	2.0	-3.2
b2202	<i>napC</i>	cytochrome c-type protein	-6.0	-1.9
b2203	<i>napB</i>	cytochrome c-type protein	-4.0	-2.6
		ferredoxin-type protein: electron		
b2204	<i>napH</i>	transfer	-9.2	-3.0
b2206	<i>napA</i>	probable nitrate reductase 3	-5.1	-2.0
b2207	<i>napD</i>	orf, hypothetical protein	-3.2	-13.4
		ferredoxin-type protein: electron		
b2208	<i>napF</i>	transfer	-20.4	-5.2
b2242	<i>glpB</i>	sn-glycerol-3-phosphate dehydrogenase	1.6	2.1
b2260	<i>menE</i>	o-succinylbenzoate-CoA ligase	-2.3	-1.7
		2-oxoglutarate decarboxylase; SHCHC		
b2264	<i>menD</i>	synthase	-4.4	-3.6
		isochorismate hydroxymutase 2,		
b2265	<i>menF</i>	menaquinone biosynthesis	-2.7	-5.4

b2281	<i>nuoI</i>	NADH dehydrogenase I chain I	-4.4	-1.4
b2285	<i>nuoE</i>	NADH dehydrogenase I chain E	-2.7	-2.5
b2287	<i>nuoB</i>	NADH dehydrogenase I chain B	-2.8	-4.9
		sensor for nitrate reductase system,		
b2469	<i>narQ</i>	protein histidine kinase	-3.0	1.4
b2486	<i>hyfF</i>	hydrogenase 4 membrane subunit	-7.6	1.4
b2489	<i>hyfI</i>	hydrogenase 4 Fe-S subunit	-7.2	-1.2
b2712	<i>hypF</i>	transcriptional regulatory protein	-3.8	-2.0
		involved in electron transport from		
b2713	<i>hydN</i>	formate to hydrogen, Fe-S centers	-3.3	-1.5
		pleiotrophic effects on 3 hydrogenase		
b2726	<i>hypA</i>	isozymes	-5.6	1.5
		pleiotrophic effects on 3 hydrogenase		
b2728	<i>hypC</i>	isozymes	-8.5	-6.1
		pleiotrophic effects on 3 hydrogenase		
b2729	<i>hypD</i>	isozymes	-3.8	-4.0
		hydrogenase-2 operon protein: may		
		effect maturation of large subunit of		
b2990	<i>hybG</i>	hydrogenase-2	-3.5	-1.5
		probable processing element for		
b2993	<i>hybD</i>	hydrogenase-2	-2.0	-1.1
b2994	<i>hybC</i>	probable large subunit, hydrogenase-2	1.3	-1.4
		probable cytochrome Ni/Fe component		
b2995	<i>hybB</i>	of hydrogenase-2	-4.0	-2.4
b2997	<i>hybO</i>	putative hydrogenase subunit	-17.5	-4.2
b3236	<i>mdh</i>	malate dehydrogenase	-2.3	2.1
b3365	<i>nirB</i>	nitrite reductase	-5.1	-4.3
b3366	<i>nirD</i>	nitrite reductase	-4.0	-4.4
		sn-glycerol 3-phosphate transport		
b3453	<i>ugpB</i>	system; periplasmic binding protein	-1.6	3.5
		ATP-binding protein of nickel transport		
b3479	<i>nikD</i>	system	-7.7	-2.5
b3605	<i>lldD</i>	L-lactate dehydrogenase	-1.4	3.7

		formate dehydrogenase, cytochrome		
b3892	<i>fdoI</i>	B556	1.8	7.2
		formate dehydrogenase-O, iron-sulfur		
b3893	<i>fdoH</i>	subunit	-1.3	9.3
		formate dehydrogenase-O, major		
b3894	<i>fdoG</i>	subunit	2.2	8.9
		1,4-dihydroxy-2-naphthoate -->		
b3930	<i>menA</i>	dimethylmenaquinone	-1.6	-1.5
b4070	<i>nrfA</i>	periplasmic cytochrome c	-24.6	-6.7
		formate-dependent nitrite reductase; a		
b4071	<i>nrfB</i>	penta-haeme cytochrome c	-25.0	-5.6
		formate-dependent nitrite reductase; Fe-		
b4072	<i>nrfC</i>	S centers	-8.7	-5.6
		part of formate-dependent nitrite		
b4076	<i>nrfG</i>	reductase complex	-6.2	-5.1
b4125	<i>dcuS</i>	putative 2-component sensor protein	-2.2	2.5
		fumarate reductase, anaerobic,		
b4151	<i>frdD</i>	membrane anchor polypeptide	-2.4	1.2
		fumarate reductase, anaerobic,		
b4152	<i>frdC</i>	membrane anchor polypeptide	-1.7	1.1
		fumarate reductase, anaerobic, iron-		
b4153	<i>frdB</i>	sulfur protein subunit	-4.2	-1.1
		fumarate reductase, anaerobic,		
b4154	<i>frdA</i>	flavoprotein subunit	-3.6	-1.2
b4379	<i>yjjW</i>	putative activating enzyme	-11.9	-2.0

Then, to elucidate the effects of aerobic condition and H₂O₂ on the expression level of the abovementioned respiration-related genes, the expression levels of those genes were examined under aerobic and anaerobic conditions with and without H₂O₂ treatment. Genes that show relatively high expression levels under anaerobic condition (Figure 5-4A-(d); red dots) were comparably down-regulated by both H₂O₂ treatment (Figure 5-4A-(c) and (a)) and aerobic condition (Figure 5-4A-(b)). On the other hand, Figure 5-4B shows that the transcription pattern of the genes involved in cellular respirations (Table 5-5) were dependent on culture conditions. For example, genes in Cluster (a) of Figure 5-4B were down-regulated by H₂O₂ under both aerobic and anaerobic conditions, while genes in Cluster (b) of Figure 5-4B were down-regulated only under aerobic condition, but not under anaerobic conditions.

Taken together, the transcriptome analysis has so far shown that oxidative stress significantly changed the regulation profile of genes involved in a range of cellular mechanisms under anaerobic condition. Comparative analysis of the transcriptome profiles obtained under anaerobic and aerobic condition suggests that such cellular functions as iron homeostasis, nucleotide biosynthesis, nitrogen metabolisms and cellular respirations show different expression patterns upon H₂O₂ treatment under aerobic and anaerobic conditions.

Figure 5-4 Scatter plots and heat maps of genes involved in cellular respiration

(A) Scatter plots of genes with statistically significant changes in expression levels (ANOVA $p < 0.01$) under four different conditions: aerobic, H_2O_2 -treated (a); aerobic, control (b); anaerobic, H_2O_2 -treated (c); anaerobic, control (d). Y-axis represents relative expression level of the same set of genes under anaerobic condition without H_2O_2 treatment (d). (B) Heat maps of hierarchically clustered genes in the GO term “Cellular respiration (GO: 0045333)”. Expression data of each gene were obtained under aerobic condition or anaerobic condition without (C, control) or with H_2O_2 treatment (T, treatment). Two examples clusters were labeled and indicated as (a) with red bracket and (b) with blue bracket, respectively.

5.4.2 Transcription factors the activity of which was perturbed by oxidative stress under anaerobic condition

The abovementioned transcriptomic changes imply that the activity of some transcription factors, which regulate their regulon members, might be perturbed by H₂O₂ under anaerobic condition. To identify transcriptional factors of which activities were significantly perturbed by H₂O₂, the transcriptomic data were further analyzed by the Network Component Analysis (NCA). The NCA is an algorithm that calculates relative activities of transcription factors from gene expression data (Kang and Chang 2012). The NCA suggests that the activities of the following five transcription factors were significantly changed upon exposure to oxidative stress under anaerobic condition: OxyR, Fur, PurR, ArcAB, and FNR. The activity differences between the control and H₂O₂-treated samples were calculated by subtracting the relative activity under anaerobic condition from that under aerobic condition, and the absolute values of the activity differences were 6.66, 2.66, 2.45, 1.04 and 0.65 for OxyR, Fur, PurR, ArcAB and FNR, respectively.

To investigate the essentiality of those transcription factors for oxidative stress tolerance under anaerobic condition, the viability of the knock-out mutants lacking each of the abovementioned transcription factor genes was tested under oxidative stress. Figure 5-5 shows that the deletion of *arcA*, *arcB* and *fnr* significantly increased or decreased the viability of *E. coli* upon exposure to H₂O₂, whereas their growth rates remained comparable to that of the wild type in the absence of H₂O₂. It is worth to note that OxyR- and Fur-deleted mutants could not survive under the same H₂O₂ treatment condition, while PurR-deletion did not significantly change the viability of *E. coli* (data not shown). Among these transcription factors, it was noticed that ArcAB and FNR are transcriptional regulators essential for cellular respiration (Compan and

Touati 1994; Gunsalus and Park 1994). In particular, ArcAB was previously reported as essential regulators for oxidative stress tolerance under aerobic condition (Loui, Chang *et al.* 2009). Intriguingly, Figure 5-5A shows that under anaerobic condition, the *arcB* deletion significantly increased the viability of the cells by 89-fold, while the deletion of its cognate regulator, *arcA*, had no significant effects on the viability. On the other hand, Figure 5-5B shows that under aerobic condition, the *arcA* deletion increased the viability, but the *arcB* deletion showed no significant effects. In addition to ArcAB, the *fnr* deletion significantly decreased the viability of *E. coli* by 131-fold under anaerobic condition (Figure 5-5A).

Figure 5-5 Relative viability of the wild type (W3110, WT) and *arcA*-, *arcB*-, *fnr*-deleted isogenic mutants upon H_2O_2 treatment under anaerobic (A) or aerobic condition (B)

Growth curves of the wild type and three mutants were depicted as follows: black circle, wild type; black rectangle, *arcB*-deleted mutants; empty circle, *arcA*-deleted mutants; black triangle, *fnr*-deleted mutants.

This result suggests that FNR is an essential transcription factor for oxidative stress tolerance under anaerobic condition, but not under aerobic condition. FNR regulates its activity by conformational changes that enable monomers to form a tetramer to recognize and bind to the specific palindromic DNA sequence (Gerasimova, Rodionov *et al.* 2001; Moore and Kiley 2001) in the absence of oxygen molecules. When oxygen molecules are present, active FNR loses the ability to bind to the target sequences because a cofactor, a [4 Fe-4 S] cluster, is oxidized by molecular oxygen and superoxide (Jervis, Crack *et al.* 2009). Therefore, this result implies that the activated FNR conferred *E. coli* viability against H₂O₂ under anaerobic condition by transcriptional regulation of its regulon members.

While the activity of FNR is directly regulated by oxygen molecules, ArcAB is a two-component regulatory system. In this system, the ArcB is a cognate redox-sensor kinase, of which activity is regulated by auto-phosphorylation in response to cellular redox states, extent of aerobiosis (Rolfe, Ter Beek *et al.* 2011) and/or relative compositions of quinones (Georgellis, Lynch *et al.* 1997; Malpica, Sandoval *et al.* 2006; Bekker, Alexeeva *et al.* 2010). Under anaerobic condition, a cognate kinase ArcB phosphorylates ArcA, and phosphorylated ArcA subsequently activates or represses genes involved in respiratory metabolisms (Lin and Iuchi 1991). On the other hand, under aerobic condition, activity of ArcB is reduced by oxidized quinones (Georgellis, Kwon *et al.* 2001), and oxidation of quinones leads to dephosphorylation of ArcA. Although the mechanisms by which the deletions of *arcA* and *arcB* confer improved viability under aerobic and anaerobic conditions, respectively, are still obscure, this result suggests that ArcAB two-component regulatory system plays a major role in oxidative stress responses under both culture conditions.

5.4.3 NADH: NAD⁺ ratio and oxidative stress

Oxidative stress induces a range of oxidation-reduction reactions that contribute to cellular protection and repair. These oxidation-reduction reactions are, for example, to scavenge oxidizing molecules (oxidants) or to reduce oxidized biochemical molecules back to normal redox states. It is worth to note that the latter reactions require reduced cofactors such as NADH (Imlay 2008) or NADPH (Ritz and Beckwith 2001). From the microarray data, it was found that H₂O₂ significantly changed the transcriptomic pattern of nucleotide biosynthesis pathways (Table 5-4), nitrogen metabolisms (Figure 5-2), TCA cycle (Figure 5-3) and cellular respirations (Figure 5-4A and 5-4B), which are all associated with producing and consuming the reduced cofactor NADH. In this light, it was hypothesized that oxidative stress responses of *E. coli* against H₂O₂ alter redox homeostasis of NADH and NAD⁺.

To verify the hypothesis, the relative amounts of NADH and NAD after H₂O₂ treatment were quantified under anaerobic and aerobic conditions in comparison to the amounts of NADH and NAD without H₂O₂ treatment. Figure 5-6A shows that the NADH: NAD⁺ ratio is significantly increased by H₂O₂ treatment under both anaerobic and aerobic conditions, while the ratio was much higher under anaerobic condition. Moreover, Figure 5-6A shows that the anaerobic growth of *E. coli* resulted in significantly higher NADH: NAD⁺ ratio compared to that of the aerobic condition even prior to the H₂O₂ treatment.

Figure 5-6 NADH: NAD⁺ ratio without (empty) or with H_2O_2 (filled) under aerobic and anaerobic condition (A). Effect of FDH1 expression on relative viability without (empty) or with H_2O_2 (filled) under aerobic (B) or anaerobic condition (C)

This result suggests that the intracellular NADH: NAD⁺ ratios in *E. coli* are different between anaerobic and aerobic conditions, and the ratio is increased by oxidative stress responses and/or directly H₂O₂ treatment. It is noteworthy that NADH: NAD⁺ ratio is involved in a broad spectrum of cellular functions such as enzyme activities, gene expressions, oxidation level of substrates, pool of available electron acceptors and therefore, metabolic fluxes (Riondet, Cachon *et al.* 2000). Therefore, this result implies that the increased NADH: NAD⁺ ratio, in turn, might play a role in determining transcriptomic profiles and metabolic pathways of *E. coli* during oxidative stress responses.

Based on these results, I generated another hypothesis whether the increased NADH: NAD⁺ ratio would improve *E. coli*'s tolerance to oxidative stress. To test this hypothesis, NAD⁺-dependent formate dehydrogenase (FDH1) coded in the *fdh1* gene of *Pichia pastoris* was expressed in *E. coli* as a NADH regeneration system. FDH1 converts NAD⁺ to its reduced form, NADH, which increases NADH availability (i.e. moles of NADH available for reduced product formation per mole of glucose consumed) (Berrios-Rivera, Bennett *et al.* 2002). Interestingly, Figure 5-6B shows that the overexpression of *fdh1* significantly increased H₂O₂ tolerance 10-fold under aerobic condition, but not under anaerobic condition (Figure 5-6C). Given that the NADH: NAD⁺ ratio was already higher under anaerobic condition (Figure 5-6A), the effect of FDH1 expression was likely more significant under aerobic condition, where lower NADH: NAD⁺ ratio was formed, and as a result, FDH1 expression increased the viability of *E. coli* against H₂O₂ under aerobic condition.

5.5 Conclusions

This study has investigated systems-level responses to oxidative stress in *E. coli* under anaerobic condition in comparison with the responses under aerobic condition. Our transcriptome and bioinformatics analyses suggest the following. First, oxidative stress significantly induced genes in the OxyR and SoxRS regulons and genes related to iron homeostasis and TCA cycle under anaerobic condition. Second, oxidative stress repressed genes involved in nucleotide *de novo* biosynthesis and cellular respiration, and in particular, anaerobic respiration. Third, genes involved in nitrogen metabolism, putrescine degradation and carbamoyl phosphate synthesis showed opposite regulation directions between aerobic and anaerobic conditions. Based on the NCA analysis of the transcriptome data, potential transcription factors, the activity of which was likely significantly perturbed by oxidative stress under anaerobic condition, have been identified. Through gene knock-out assays, it has been demonstrated that FNR was a transcription factor essential for viability and *arcB* knock-out increased tolerance to oxidative stress 89-fold under anaerobic condition. Moreover, this study suggested that NADH: NAD⁺ ratio was an important regulatory factor in oxidative stress responses. That is, the NADH: NAD⁺ ratio was significantly increased by oxidative stress under both aerobic and anaerobic conditions, while increased NADH availability significantly improved tolerance against oxidative stress under aerobic condition. This study provides insights into systems-level regulatory mechanisms underlying oxidative stress responses under anaerobic condition, which can be exploited for rational strain optimization that leads to improved tolerance against oxidative stress.

CHAPTER 6 Identification and reconstitution of genetic regulatory networks for improved microbial tolerance to isooctane⁵

6.1 Abstract

Microbial tolerance to hydrocarbons has been studied as an effort to improve the productivity of biochemical processes and to enhance the efficiency of hydrocarbon bioremediation. Despite these works, few attempts have been made to design rational strategies to improve microbial tolerance to hydrocarbons. Herein, this study presents an engineering framework that enables us to harness our understanding of genetic regulatory networks to improve hydrocarbon tolerance. In this study, isooctane was used as a representative hydrocarbon due to its use in petroleum refining and in biochemical processes. To increase isooctane tolerance, I first identified essential transcriptional determinants and genetic regulatory networks underlying cellular response to isooctane in *Escherichia coli* using genome-wide microarray analysis. Based on functional transcriptome and bioinformatics analysis, a range of combinations of transcription factors whose activity was predictably perturbed by isooctane were knocked out and overexpressed to reconstitute the regulatory networks. It was demonstrated that the reconstitution of the regulatory networks led to a significant improvement in isooctane tolerance, and especially, engineered *E. coli* strains lacking and overexpressing some of the perturbed transcription factors showed 3- to 5-fold improvement. This microbe with high tolerance to isooctane can be harnessed for biochemical processes, fuel oil bioremediation and metabolic engineering for biofuel production. Furthermore, the engineering framework employed

⁵ The text, figures and tables in this chapter have been reproduced by permission of The Royal Society of Chemistry. This article can be found in this link: <http://dx.doi.org/10.1039/C2MB05441H>

to improve the tolerance in this study can be exploited for developing other microbes with desired phenotypes.

6.2 Introduction

As the world is running short of the fossil fuels, many alternative energy sources have been investigated and developed, including biological production of energy molecules such as alcohols and hydrocarbons. In particular, the microbial production of hydrocarbons, including alkanes, has recently drawn much attention because they are major components of the fossil fuels. In nature, alkanes are found in microorganisms, insects and plants (Ladygina, Dedyukhina *et al.* 2006) and can directly be converted from fatty acid metabolites, for example, by decarbonylation of fatty aldehydes (Cheesbrough and Kolattukudy 1984). Therefore, much effort has been made to identify and characterize metabolic pathways that can be introduced into microorganisms for alkane synthesis. For instance, alkane biosynthesis pathways from cyanobacteria were recently re-constructed in *Escherichia coli* (Schirmer, Rude *et al.* 2010). Despite this promising potential of microbial hydrocarbon production, a thorough understanding of interactions between host cells and produced hydrocarbons, including cellular toxicity, is a prerequisite for the development of hydrocarbon synthesis bioprocesses.

Hydrocarbon toxicity and corresponding microbial tolerance have been studied as an effort to improve the productivity of biochemical processes, such as whole-cell catalyzed reactions performed in hydrophobic organic solvents, and to enhance the efficiency of hydrocarbon bioremediation (Heipieper, Weber *et al.* 1994). Those studies have so far suggested the followings. First, hydrophobic organic solvents disrupt microbial membranes by direct insertion, and the cytotoxicity of these

chemicals shows a strong inverse correlation with the logarithm of their partition coefficients ($\log P_{ow}$) (Aono and Kobayashi 1997; Rutherford, Dahl *et al.* 2010). Second, organic solvent tolerance might be related to lipopolysaccharides modification (Pinkart, Wolfram *et al.* 1996) and induction of efflux systems (Aono, Kobayashi *et al.* 1995; Nakajima, Kobayashi *et al.* 1995; Isken and de Bont 1996) in *E. coli*, *Pseudomonas putida* and *Pseudomonas aeruginosa*. Third, organic solvents affect the chain length and the composition of fatty acids (Ingram 1977; Heipieper, Weber *et al.* 1994; Sikkema, de Bont *et al.* 1995), outer membrane compositions (Ingram 1977; Weber, Isken *et al.* 1994; Pinkart, Wolfram *et al.* 1996), and cell surface hydrophobicity (Asako, Nakajima *et al.* 1997). Despite these findings, which are mostly focused on cell physiology and physicochemical properties of the cell membrane, there is a lack of understanding of cellular mechanisms underlying hydrocarbon response at the systems level. Moreover, few attempt have been so far made to harness our understanding of hydrocarbon response mechanisms to improve hydrocarbon tolerance.

Consequently, this study aimed to design rational strategies to improve the tolerance of *E. coli* against hydrocarbons by reconstitution of genetic regulatory networks underlying response mechanisms. Here, isooctane was used as a representative hydrocarbon because isooctane (2, 2, 4-trimethylpentane, C_8H_{18}) has been widely utilized in petroleum refining as an important additive of gasoline to prevent automobile engine knocking. Further, isooctane is frequently used as an organic solvent in biochemical processes because it improves enzyme-mediated catalytic reactions (Klibanov 2001). One such example is stereoselective microbial reduction in two liquid phase bioreactors (Miura, Zou *et al.* 2000; Gong and Xu 2005). Briefly, genetic regulatory networks and essential transcriptional determinants were identified

based on transcriptome analysis. Then, it was demonstrated that the reconstitution of the networks via knock-out and overexpression of perturbed transcription factors significantly increased isooctane tolerance. This microbe with high tolerance to isooctane can be harnessed for biochemical processes, fuel oil bioremediation and metabolic engineering for biofuel production. Furthermore, the engineering framework employed to improve the tolerance in this study can be exploited for developing other microbes with desired phenotypes.

6.3 Materials and Methods

6.3.1 Strains, growth conditions and isooctane treatment

Escherichia coli K12 BW25113 was used as a wild-type *E. coli* in this study. Isogenic mutants and their parental strain BW25113 were obtained from the Coli Genetic Stock Center (CGSC). For culture preparation, an overnight culture was diluted to optical density of 0.02 at 600 nm ($OD_{600}=0.02$) in M9 minimal medium (Na_2HPO_4 33.9 g/L, KH_2PO_4 15 g/L, NaCl 2.5 g/L and NH_4Cl 5 g/L) with supplements (0.8% glycerol, 0.2% casamino acid, 1 mM thiamine, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$) (Saeidi, Wong *et al.* 2011), and the diluted culture was incubated at 37°C with shaking at 250 rpm. When the OD_{600} of the cell culture reached 0.1 (early exponential phase), 5%-20% isooctane (v/v) was added in the cell culture, and the isooctane-containing cell culture was continuously incubated under the same conditions. All strains and plasmids used in this study are listed in Table 6-1.

Table 6-1 A list of bacterial strains, plasmids and primers used in Chapter 6.

Strains/ plasmids	Relevant genotype and/or property	Reference
<i>E. coli</i> K12		
BW25113	F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC# 7636
	F- mcrA D(mrr-hsdRMS-mcrBC) w80lacZDM15	
TOP10	DlacX74 deoR nupG recA1 araD139 D(ara- eu)7697 galE15 galK16 rpsL(StrR) endA1 l	invitrogen
JW0389-1	BW25113 ΔphoB763::kan	CGSC# 8565
JW0494-1	BW25113 ΔallR789::kan	CGSC# 8639
JW0669-2	BW25113 Δfur-731::kan	CGSC# 8758
JW0872-2	BW25113 Δlrp-787::kan	CGSC# 8901
JW0895-3	BW25113 ΔihfB735::kan	CGSC# 8917
JW1316-1	BW25113 ΔtyrR760::kan	CGSC# 9179
JW1702-1	BW25113 ΔihfA786::kan	CGSC# 9441
JW1935-1	BW25113 ΔrcsA726::kan	CGSC# 9610
JW2205-2	BW25113 ΔrcsB770::kan	CGSC# 9783
JW2521-1	BW25113 ΔhcaR782::kan	CGSC# 10007
JW3100-1	BW25113 ΔagaR745::kan	CGSC# 10362
JW3229-1	BW25113 Δfis-779::kan	CGSC# 10443
JW3484-1	BW25113 ΔgadX771::kan	CGSC# 10580
JW3883-1	BW25113 ΔcpXR772::kan	CGSC# 10800
JW4050-1	BW25113 ΔrpiR784::kan	CGSC# 10915
JW4356-2	BW25113 ΔtrpR789::kan	CGSC# 11110
JW5096-1	BW25113 ΔkdpE743::kan	CGSC# 11171
JW5249-1	BW25113 ΔmarA752::kan	CGSC# 11269
JW5437-1	BW25113 ΔrpoS746::kan	CGSC# 11387
JW5702-4	BW25113 Δcrp-765::kan	CGSC# 11596
agar-fis	BW25113 ΔagaR745, Δfis-779::kan	In this study

agar-fur	BW25113 Δ agaR745, Δ fur-731::kan	In this study
agar-ihfa	BW25113 Δ agaR745, Δ ihfA786::kan	In this study
agar-ihfb	BW25113 Δ agaR745, Δ ihfB735::kan	In this study
agar-rpos	BW25113 Δ agaR745, Δ rpoS746::kan	In this study
fis-fur	BW25113 Δ fis-779, Δ fur-731::kan	In this study
fis-ihfa	BW25113 Δ fis-779, Δ ihfA786::kan	In this study
fis-ihfb	BW25113 Δ fis-779, Δ ihfB735::kan	In this study
fur-rpos	BW25113 Δ fur-731, Δ rpoS746::kan	In this study
gadx-rpos	BW25113 Δ gadX771, Δ rpoS746::kan	In this study
hcar-agar	BW25113 Δ hcaR782, Δ agaR745::kan	In this study
hcar-fur	BW25113 Δ hcaR782, Δ fur-731::kan	In this study
hcar-ihfa	BW25113 Δ hcaR782, Δ ihfA786::kan	In this study
hcar-ihfb	BW25113 Δ hcaR782, Δ ihfB735::kan	In this study
hcar-rpos	BW25113 Δ hcaR782, Δ rpoS746::kan	In this study
rcsa-rcsb	BW25113 Δ rcsA726, Δ rcsB770::kan	In this study
rcsb-crp	BW25113 Δ rcsB770, Δ crp-765::kan	In this study
rpos-ihfa	BW25113 Δ rpoS746, Δ ihfA786::kan	In this study
rpos-ihfb	BW25113 Δ rpoS746, Δ ihfB735::kan	In this study

6.3.2 Knock-out mutant construction, gene overexpression, and viability testing

Knock-out mutants were constructed by P1 phage transduction method (Lennox 1955). Briefly, a donor mutant strain (A) was infected by P1 phages, which contain no selection marker. After cell lysis, P1 phage lysate, which contains deletion of a target gene A, was collected from the supernatant after centrifugation. Marker-free recipient mutant strains were prepared as previously described (Datsenko and Wanner 2000). After removing the kanamycin resistance marker, the recipient mutant strain (B) was infected by the P1 phage lysate obtained from the donor mutant strain (A), and the deletion of the target gene A was transferred to the recipient mutant strain (B) by homologous recombination. A positive double mutant strain was screened on an agar plate containing 30 µg/ml Kanamycin, and the deletion of the target gene A was confirmed by PCR using two primers which bind to intact upstream and downstream sequence of the target gene A.

Each coding sequence of the selected genes was cloned from the genomic DNA of *E. coli* BW25113 to the arabinose-inducible plasmid pBAD30 (Guzman, Belin *et al.* 1995), and each plasmid was transformed into the corresponding mutant strain (Leong and Chen 2008). To test viability of each mutant strain that overexpressed the deleted gene, overnight culture of each mutant strain was diluted 100 times in the M9 minimal medium containing 0.2 % arabinose. The mutant strain culture was incubated with 20% (v/v) isooctane for 24 hours at 37 °C and colony forming unit (CFU) of each mutant strain was counted on LB agar plates with two biological replicates (Zhou, Qi *et al.* 2010). CFU of each strain treated with 20% isooctane was normalized to CFU of each strain without isooctane treatment, and relative CFU of each mutant strain was calculated as a ratio of normalized CFU of each mutant strain to normalized CFU of the wild type (WT).

6.3.3 Total RNA extraction and microarray analysis

E. coli K12 BW25113 was grown as described above. At the early log phase ($OD_{600}=0.1$), *E. coli* cell culture was exposed to isooctane. The cells were harvested after 10 min-, 30 min- and 60 min-incubation by centrifugation at 6000 rpm for 10 min. To protect RNA profiles, the harvested cells were immediately preserved in 1 ml RNeasy Lysis Buffer (Qiagen, USA). The samples were prepared with two biological replicates. The cells in 1 ml RNeasy Lysis Buffer were collected by centrifugation at 6000 rpm for 10 min. The *E. coli* cell pellets were incubated in RNeasy lysis buffer containing lysozyme (1 mg/ml, 10 mM Tris-HCl, 1 mM EDTA, pH 8.0) to disrupt membranes. Total RNA was extracted by RNeasy Mini Kit (Qiagen, USA). The integrity of the RNA was evaluated by Agilent Bioanalyzer using the RNA Nano 6000 Kit (Jang, Chang *et al.* 2008). cDNA was prepared and labeled using Fairplay III Kit (Stratagen, USA). The same amount of cDNA prepared from each sample was hybridized on the *E. coli* Gene expression microarray (8x15K, Agilent, USA) for 16 hours at 65°C using Agilent Hybridization Kit. The unbound cDNA samples were washed out before scanning (Agilent Microarray Scanner, Agilent). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE32131.

6.3.4 Transcriptome data processing and analysis

Raw fluorescent signals were extracted by Agilent Feature Extraction Software v. 9.5.3.1 and data files were analyzed in GeneSpring GX 11 (Agilent, USA) (Chang, Toghrol *et al.* 2007). The signal from each spot was normalized and scaled to median of all samples after background signal subtraction. One-way ANOVA was subsequently performed with Benjamini-Hochberg multiple testing correction ($p<0.05$) for significance analysis. Fold changes of the signals were calculated as a

ratio of signal intensity at time points 10, 30 and 60 min to signal intensity at time point 0. The genes that showed significantly different expression levels (fold change >2, t-test $p < 0.05$) were used for DAVID functional annotation analysis and network component analysis (NCA) (Huang, Sherman *et al.* 2009). Only these statistically significant genes were also used for further analyses and shown in the tables and the figures.

6.3.5 Transmission electron microscopy (TEM)

For TEM, *E. coli* cells were incubated with 20% isooctane for 24 hours while the control samples were incubated with 20% PBS under same condition. The cells were collected by centrifugation at 3,220 g for 10 min and the cell pellets were washed with PBS. The cell pellets were fixed by 2.5 % glutaraldehyde (GTA) prepared in 0.1 M sodium cacodylate for 2 hours at 4°C, and washed with PBS. TEM grids were submerged in the sample for 30 min. TEM grids were dehydrated in ethanol solutions at 37, 67, 95 and 100%, and sequentially dried in acetone: hexamethyldisilazane (HMDS) (2:1), acetone: HMDS (1:2) and 100 % HMDS (Ling, Kang *et al.* 2010). The prepared samples were examined using TEM (JEOL-3010).

6.4 Results and discussion

6.4.1 Essential genetic determinants underlying cellular response to isooctane

A colony forming unit (CFU) counting assay showed that toxicity of isooctane was not apparent in *E. coli* during the exponential growth phase, but it became significant at the stationary phase (data not shown). Therefore, to elucidate cellular mechanisms and genetic determinants underlying *E. coli* response to isooctane, transcriptional profile changes were monitored by DNA microarrays. The genes with statistically significant changes in transcription level were selected based on their p -value (< 0.05) and fold

change (>2) at 10 min or 60 min compared to 0 min. Among the most significantly induced genes were those involved in the following three stress responses: heat stress, iron transport and acid stress (Table 6-2 and Table 6-3).

Table 6-2 Top 10 genes out of 225 significantly up-regulated ($p < 0.05$, fold change > 2) genes at 10min after isooctane treatment

: 17 out of 23 genes related to heat shock or iron transport are among the most highly induced genes.

B. No	Gene	<i>P</i> value	Fold Change	Gene description	Stress Response
b3686	<i>ibpB</i>	0.02	144.1	heat shock protein	Heat Shock
b3687	<i>ibpA</i>	0.02	93.5	heat shock protein	Heat Shock
b0593	<i>entC</i>	0.01	19.5	isochorismate hydroxymutase 2, enterochelin biosynthesis	Iron transport
b0473	<i>htpG</i>	0.02	19.2	chaperone Hsp90, heat shock protein C 62.5	Heat Shock
b2155	<i>cirA</i>	0.02	18.7	outer membrane receptor for iron-regulated colicin I receptor; porin; requires tonB gene product	Iron transport
b4511	<i>ybdZ</i>	0.01	18.1	conserved protein	
b0594	<i>entE</i>	0.01	17.9	2,3-dihydroxybenzoate-AMP ligase	Iron transport
b2150	<i>mglB</i>	0.01	16.5	galactose-binding transport protein; receptor for galactose taxis	
b4367	<i>fhuF</i>	0.02	16.2	orf, hypothetical protein	Iron transport
b2659	<i>csiD</i>	0.01	15.5	orf, hypothetical protein	

Table 6-3 Top 10 genes out of 696 significantly up-regulated ($p < 0.05$, fold change > 2) genes at 60 min after isooctane treatment

B. No	Gene	<i>p</i> value	Fold Change	Gene description	Stress Response
b1493	<i>gadB</i>	0.03	50.7	glutamate decarboxylase isozyme	Acid stress
b3517	<i>gadA</i>	0.03	44.8	glutamate decarboxylase isozyme	Acid stress
b1492	<i>gadC</i>	0.03	31.1	acid sensitivity protein, putative transporter	Acid stress
b1616	<i>uidB</i>	0.04	18.1	glucuronide permease	
b4365	<i>yjjQ</i>	0.04	16.1	putative regulator	
b1264	<i>trpE</i>	0.03	15.2	anthranilate synthase component I	
b2851	<i>ygeG</i>	0.03	13.3	orf, hypothetical protein	
b3512	<i>gadE</i>	0.03	11.6	orf, hypothetical protein	Acid stress
b4067	<i>actP</i>	0.04	10.9	putative transport protein	
b3506	<i>slp</i>	0.03	10.7	outer membrane protein induced after carbon starvation	

As shown in Figure 6-1a, the heat stress-related genes generally underwent a sharp increase and a subsequent decrease in their transcription levels at 10 and 30 min, respectively. In particular, *ibpA* and *ibpB*, which code for small heat shock proteins, showed 93.5- and 144.1-fold increase in the transcription level at 10 min. Small heat shock proteins (sHsps) are protein chaperons that bind to unfolded proteins and prevent protein aggregation (Kuczyńska-Wisnik, Kedzierska *et al.* 2002). In relation to this finding, it was also observed that other chaperone genes including *grpE* and *hchA* (Hsp31) were up-regulated at 10 min (Figure 6-1a). These changes in the transcription levels of the protein chaperon genes may suggest that proper protein folding might be

important for maintaining cellular functions in the presence of isooctane. Figure 6-1b shows that the differently expressed genes involved in iron transport were significantly up-regulated by isooctane treatment. Given that regulation of iron transport is altered in response to oxidative stress (Imlay 2008), low pH (King, Lucchini *et al.* 2010) and isobutanol stress (Brynildsen and Liao 2009), iron transport and homeostasis may be part of common stress responses. Figure 6-1c shows that the differently expressed genes involved in acid stress response continued to increase up to 60 min after isooctane treatment. *gadA*, *gadB*, *gadC* and *gadE* are the most significantly up-regulated genes showing 44.8-, 50.7-, 31.1- and 11.6-fold changes, respectively, after 60 min. GadE is a transcriptional regulator of *gadA* and *gadBC* (glutamate decarboxylase), which are induced at low pH (Tucker, Tucker *et al.* 2002), also part of the glutamate dependent acid resistance system (GAD system). In addition to the abovementioned three stress responses, genes involved in oxidative stress responses were examined because oxidative stress responses are reportedly related to various stresses such as organic solvent stress (Brynildsen and Liao 2009; Rutherford, Dahl *et al.* 2010), osmotic stress (Gunasekera, Csonka *et al.* 2008), acid stress (Maurer, Yohannes *et al.* 2005) and heat stress (Matuszewska, Kwiatkowska *et al.* 2008). Figure 6-1d shows that the genes involved in oxidative stress response, related to OxyR regulon, were mostly repressed (*fldA*, *fldB*, *ribA*, *rimK*, *trxA*, *grxC*), while *fumC*, *sodA* and *katG*, which encode fumarate C superoxide dismutase and catalase, respectively, were up-regulated.

Figure 6-1 Temporal transcriptome responses of *E. coli* to isooctane

Normalized expression levels are presented at the times 0, 10, 30 and 60 min for genes related to (a) heat stress (b) iron transport (c) oxidative stress and (d) acid stress.

Based on these results, to determine whether any of the four stress response mechanisms above were essentially associated with isooctane tolerance, I tested the viability of *E. coli* knock-out mutants lacking corresponding key genes, in particular the genes encoding transcription factors essential for the abovementioned stress responses, in the presence of isooctane. As indicated in Figure 6-2, those knocked out genes were *ibpA* (heat shock stress), *oxyR* and *soxS* (oxidative stress), *gadE* and *gadX* (acid stress) and *fur* (iron homeostasis). Figure 6-2 shows that the knock-out of *fur*, *gadX* and *ibpA* decreased the viability (t-test: $p < 0.05$), whereas the knock-out of *oxyR*, *soxS* and *gadE* mostly caused no change in the viability in presence of isooctane. The finding that the *fur* deletion decreased the viability most significantly by over 80% implies that Fur may be among essential regulatory determinants in cellular response to isooctane. It is noteworthy that Fur regulates various cellular functions such as respiration and energy metabolism (McHugh, Rodriguez-Quinones *et al.* 2003; Zhang, Gosset *et al.* 2005), flagellum chemotaxis (Stojiljkovic, Baumler *et al.* 1994), and acid and oxidative stress (Kozyrev and Vassinova 2000). Overall, the results above suggest that cellular mechanisms involved in heat stress responses, acid stress responses and iron homeostasis, mediated by IbpA, GadX and Fur, respectively, may serve as essential entities in the regulatory network of isooctane responses.

Figure 6-2 Essentiality of various stress mechanisms in isooctane-tolerance of *E. coli*.

Relative viability of *fur*-, *gadE*-, *gadX*-, *ibpA*-, *oxyR*- and *soxS*-deleted mutants was measured in the presence of isooctane. Relative viability was determined as a ratio of the viability of the mutant to that of the wild type (WT). The relative viabilities shown are statistically significant for WT and all mutants.

6.4.2 Functional transcriptome analysis and corresponding physiological and metabolic responses

To further elucidate isooctane response-relevant cellular mechanisms from the transcriptome profiles, the differently expressed genes during 60 min exposure were clustered using hierarchical clustering (Figure 6-3a), and the clusters were analyzed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang, Sherman *et al.* 2009). Cluster 0 contains genes of which expression level was decreased by isooctane (Figure 6-3b), whereas Cluster 1 includes genes with increased expression level during 60 min of incubation (Figure 6-3c). Gene ontology (GO) analysis on the genes of these clusters indicates that many of those genes are related to the membranes and the envelopes of bacterial cells (Figure 6-3d). In particular, Figure 6-3d shows that two GO terms, “periplasmic space” and “pilus”, were significant in Cluster 1, where the expression of the genes was continuously increased in general.

Figure 6-3 Clustering and functional annotation analysis of significantly induced or repressed genes

Significantly induced or repressed genes at the time point 60 min (t-test $p < 0.05$, fold change > 2) were analyzed by hierarchical clustering (a). Cluster 0 (b) included genes of which expression continued to be repressed, and Cluster 1 (c) contained genes of which expression continued to be induced by isooctane. A functional annotation analysis performed in DAVID suggested a distribution of genes in each significant GO term ($p < 0.05$) associated with cellular components (d: solid bar, Cluster 0; open bar, Cluster 1).

Next, the annotations of the genes in each cluster were further clustered by DAVID's annotation clustering. As shown in Table 6-2, the annotation clustering analysis results in the identification of physiological functions likely affected by the changed expression of genes during isooctane responses. First, the results suggest that part of central metabolic pathways were affected by isooctane. Table 6-4 shows that the carbohydrate biosynthetic process and the lipopolysaccharides biosynthetic process were down-regulated at the transcription level, while genes involved in carbohydrate catabolic processes such as galactose metabolism and fatty acid oxidation were up-regulated by isooctane. Second, transport mechanisms might be affected by isooctane given that annotation clusters included organelle envelope (GO:0031967), envelope (GO:0031975), iron ion transport (GO:0006826) and ABC transporters (eco02010)

Table 6-4 Annotation clusters and associated terms of genes of Clusters 0 and 1.

Enrichment score (E. score) was calculated for each annotation cluster from Gene Ontology and KEGG pathway databases, and the most significantly associated annotation clusters, of which E. score is above 10, were used for further analysis and are shown here.

Annotation clusters	E. Score	Database	Accession Number	Terms	No. of genes
Cluster 0					
C0-1	15.5	BP_FAT	GO:0016051	Carbohydrate biosynthetic process	41
C0-2	13.4	BP_FAT	GO:0009103	Lipopolysaccharide biosynthetic process	23
Cluster 1					
C1-1	30.5	CC_FAT	GO:0031967	Organelle envelope	84
C1-2	24.4	CC_FAT	GO:0031975	Envelope	123
C1-3	18.3	KEGG_P	eco02010	ABC transporters	38
C1-4	13.8	KEGG_P	ecj00052	Galactose metabolism	16
C1-5	13.0	BP_FAT	GO:0019395	Fatty acid oxidation	15
C1-6	10.2	BP_FAT	GO:0006826	Iron ion transport	19

*BP_FAT: GO term for biological processes; CC_FAT: GO term for cellular components; KEGG_P: KEGG pathways

In addition, this functional annotation analysis shows that isooctane changed transcription regulation of genes associated with the surface and outer membrane properties of *E. coli*. As shown in Table 6-5 these genes are related to Type 1 fimbriae (*fimEDFG*), curlin production (*csgBAC*), colanic acid biosynthesis (*wcaL*, *wcaK*, *wzcC*, *cpsG*, *cpsB*, *wcaIFEDBA-wzc*) and cryptic genes (*ycbRT*, *yadKLM-htrE-ecpD*, *sfmACHF*, *yraHI*, *yfcPQSV* and *yehB*), which were recently reported to enhance *E. coli* K12's adhesion to various surfaces including abiotic surfaces and epithelial cells surfaces (Korea, Badouraly *et al.* 2010). On the other hand, genes involved in lipopolysaccharide biosynthesis were down-regulated (*wbbKJI-rfc-glf-rfbXCA*, *rfaFCYI*, *wzzE-rffEDGC-wzyE-rffM*) (Table 6-5).

Table 6-5 Significantly up- or down-regulated genes related to membrane and surface structure and properties

Fold changes were calculated as a ratio of the signal intensities at the time points 10, 30 and 60 min to those at the time point 0.

b. No.	Gene	Gene Name	p-value	10min /0min	30min /0min	60min /0min
b0136	<i>yadK</i>	predicted fimbrial-like adhesin protein	0.04	1.4	3.0	3.5
b0137	<i>yadL</i>	predicted fimbrial-like adhesin protein	0.02	1.2	2.4	2.4
b0138	<i>yadM</i>	predicted fimbrial-like adhesin protein	0.03	1.9	3.8	4.4
b0139	<i>htrE</i>	predicted outer membrane usher protein	0.03	2.2	3.7	4.6
b0140	<i>ecpD</i>	predicted periplasmic pilin chaperone	0.03	1.7	3.6	4.5
b0530	<i>sfmA</i>	predicted fimbrial-like adhesin protein	0.02	1.1	2.8	2.7
b0531	<i>sfmC</i>	pilin chaperone, periplasmic	0.03	1.3	3.5	3.9
b0533	<i>sfmH</i>	predicted fimbrial-like adhesin protein	0.03	1.5	4.2	4.2
b0534	<i>sfmF</i>	predicted fimbrial-like adhesin protein	0.03	1.7	4.4	4.8
b0939	<i>ycbR</i>	predicted periplasmic pilin chaperone	0.03	1.1	2.4	2.7
b0941	<i>ycbT</i>	predicted fimbrial-like adhesin protein	0.04	1.1	2.4	2.9
b1041	<i>csgB</i>	curlin nucleator protein, minor subunit in curli complex	0.03	2.4	4.8	3.6

b1042	<i>csgA</i>	cryptic curlin major subunit	0.03	1.6	3.6	3.2
b1043	<i>csgC</i>	predicted curli production protein	0.03	1.2	2.8	3.3
b2032	<i>wbbK</i>	lipopolysaccharide biosynthesis protein	0.03	-3.1	-4.1	-4.8
b2033	<i>wbbJ</i>	predicted acyl transferase	0.03	-3.7	-4.0	-4.7
b2034	<i>wbbI</i>	conserved protein	0.04	-3.4	-3.6	-4.5
b2035	<i>rfe</i>	O-antigen polymerase	0.04	-4.8	-6.2	-7.5
b2036	<i>glf</i>	UDP-galactopyranose mutase, FAD/NAD(P)-binding	0.04	-3.1	-4.5	-5.6
b2037	<i>rfbX</i>	predicted polisoprenol-linked O-antigen transporter	0.04	-2.5	-3.8	-4.3
b2038	<i>rfbC</i>	dTDP-4-deoxyrhamnose-3,5-epimerase	0.04	-3.6	-3.6	-5.3
b2039	<i>rfbA</i>	glucose-1-phosphate thymidyltransferase	0.05	-2.2	-2.7	-3.2
b2044	<i>wcaL</i>	predicted glycosyl transferase	0.03	1.1	2.5	2.9
b2045	<i>wcaK</i>	predicted pyruvyl transferase	0.03	1.4	3.3	3.2
b2046	<i>wzxC</i>	colanic acid exporter	0.03	1.1	3.1	4.0
b2048	<i>cpsG</i>	phosphomannomutase	0.03	1.3	3.0	3.4
b2049	<i>cpsB</i>	mannose-1-phosphate guanylttransferase	0.03	1.3	2.9	3.2
b2050	<i>wcaI</i>	predicted glycosyl transferase	0.03	1.1	2.6	3.0
b2054	<i>wcaF</i>	predicted acyl transferase	0.03	1.2	2.7	3.5
b2055	<i>wcaE</i>	predicted glycosyl transferase	0.04	-1.0	2.7	2.8
b2056	<i>wcaD</i>	predicted colanic acid polymerase	0.03	1.2	2.7	3.5
b2058	<i>wcaB</i>	predicted acyl transferase	0.04	1.4	3.5	4.7
b2059	<i>wcaA</i>	predicted glycosyl transferase	0.03	1.5	4.0	4.7
b2060	<i>wzc</i>	protein-tyrosine kinase	0.03	1.6	3.4	4.2
b2109	<i>yehB</i>	predicted outer membrane protein	0.03	2.0	4.0	5.1
b2333	<i>yfcP</i>	predicted fimbrial-like adhesin protein	0.03	1.4	3.6	4.1
b2334	<i>yfcQ</i>	predicted fimbrial-like adhesin protein	0.03	1.3	3.4	3.7
b2336	<i>yfcS</i>	predicted periplasmic pilus chaperone	0.03	1.4	3.5	3.9
b2339	<i>yfcV</i>	predicted fimbrial-like adhesin protein	0.03	1.6	3.7	4.9
b3142	<i>yraH</i>	predicted fimbrial-like adhesin protein	0.03	-1.1	1.8	2.3
b3143	<i>yraI</i>	predicted periplasmic pilin chaperone	0.03	1.8	4.0	5.9
b3620	<i>rfaF</i>	ADP-heptose:LPS heptosyltransferase II	0.03	-1.5	-2.6	-2.4
b3621	<i>rfaC</i>	ADP-heptose:LPS heptosyl transferase	0.03	-2.1	-2.4	-3.5

		I				
b3625	<i>rfaY</i>	lipopolysaccharide core biosynthesis protein	0.04	-2.2	-2.3	-3.0
b3627	<i>rfaI</i>	UDP-D-galactose:(glucosyl)lipopolysaccharide-alpha-1,3-D-galactosyltransferase	0.03	-2.4	-2.6	-3.6
b3785	<i>wzzE</i>	Entobacterial Common Antigen (ECA) polysaccharide chain length modulation protein	0.03	-1.9	-2.9	-3.0
b3786	<i>rffE</i>	UDP-N-acetyl glucosamine-2-epimerase	0.03	-2.4	-2.8	-3.5
b3787	<i>rffD</i>	UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase	0.03	-2.5	-3.1	-3.3
b3788	<i>rffG</i>	dTDP-glucose 4,6-dehydratase	0.04	-2.4	-2.8	-3.0
b3790	<i>rffC</i>	TDP-fucosamine acetyltransferase	0.03	-1.7	-1.9	-2.2
b3793	<i>wzyE</i>	predicted Wzy protein involved in ECA polysaccharide chain elongation	0.04	-1.3	-2.1	-2.2
b3794	<i>rffM</i>	UDP-N-acetyl-D-mannosaminuronic acid transferase	0.03	-1.5	-1.6	-2.2
b4313	<i>fimE</i>	tyrosine recombinase/inversion of on/off regulator of fimA	0.03	1.8	2.8	3.5
b4317	<i>fimD</i>	outer membrane usher protein, type 1 fimbrial synthesis	0.03	1.1	2.4	2.9
b4318	<i>fimF</i>	minor component of type 1 fimbriae	0.03	1.2	2.6	3.3
b4319	<i>fimG</i>	minor component of type 1 fimbriae	0.02	1.2	2.2	2.5

Based on these results, it was hypothesized that the transcriptional changes of the abovementioned surface/membrane-related genes might lead to morphological changes. As shown in Figure 6-4, isooctane-treated *E. coli* developed fimbriae- and pilus-like structures and showed smoother surface edges (Figure 6-4c and d), while the control sample had clearly contrasted cellular boundaries (Figure 6-4a and b). These results indicate that the properties of cell surfaces and membrane were changed and the filamentary structures were developed in response to isooctane.

In relation to this finding, it was observed that *E. coli* produced unknown materials to form an insoluble layer between two liquid phases. Figure 6-4e shows that many small isooctane drops were entangled with the layer, which prevents isooctane drops from emerging. Interestingly, Matrix-Assisted Laser Desorption/Ionization-Time of Flight/Time of Flight (MALDI-TOF/TOF) analysis suggested that the layer contained dipeptide binding protein (DppA) and oligopeptide binding protein (OppA). In accordance with this result, the expression level of *dppA* was highly induced by isooctane treatment. DppA, a periplasmic component of dipeptide transport system, binds to dipeptides and enables them to be transported by the ABC transporter. Given that DppA seems to have a broad range of substrate specificity (Letoffe, Delepelaire *et al.* 2006), this result may suggest that *E. coli* transports isooctane into the cytoplasm using this transporting system.

Figure 6-4 Transmission electron microscopy images of *E. coli* incubated without (a and b) or with (c and d) isooctane

Arrows indicate pilus and fimbrial structures. Phase contrast microscopy images of the insoluble layer between the water phase and the isooctane phase of *E. coli* cell culture after 24-hr incubation with isooctane (IO) (e) and at the time 0 (f).

6.4.3 Reconstitution of genetic regulatory networks of isooctane response by knocking out and overexpressing related transcription regulators

Based on the abovementioned changes in cellular mechanisms in response to isooctane, to identify transcription factors associated with these changes, the microarray data were further analyzed using network component analysis (NCA) (Liao, Boscolo *et al.* 2003). As indicated in Figure 6-5, the NCA analysis suggested that 19 transcription factors underwent a significant change in their regulatory activities in responses to isooctane. First, to characterize the involvement of these 19 transcriptional factors in isooctane tolerance, each of the 19 transcription factor genes was individually knocked out and to examine the viability of the knock-out mutants in the presence of isooctane. As shown in Figure 6-5, the deletion of the genes encoding the following 6 proteins significantly decreased the viability: IHF (integration host factor), RpoS (sigma factor S), AgaR (N-acetylgalactosamine repressor), HcaR (hydrocinnamic acid regulator), GadX (a regulator for glutamic acid decarboxylase) and Fur (a transcription factor for ferric uptake regulation). Note that the growths of those knock-out mutants and the wild type remained comparable in the absence of isooctane. This result implies that these transcription factors are essential entities in genetic regulatory networks of *E. coli*'s stress responses to isooctane. Interestingly, out of these 6 regulators, GadX, Fur and RpoS are directly related to stress responses. That is, GadX and Fur are important regulators in acid stress responses and iron homeostasis respectively, while Sigma factor S (RpoS) is an alternative sigma factor for RNA polymerase, which is activated under various stress conditions in *E. coli*. Intriguingly, the deletion of some of the transcription factors perturbed by isooctane significantly increased the tolerance level. In Figure 6-5, the knock-out mutants lacking *rpiR*, *trpR*, *phoB*, *lrp*, *kdpE*, and *crp* show increased tolerance against isooctane. In particular, the knock-out of *phoB*, *lrp*, and *kdpE* increased the tolerance

3.6-, 3.8-, and 5.0-fold, respectively. Second, to test whether the overexpression of the transcription factors perturbed by isooctane would increase the tolerance, an overexpression vector containing each of the 9 transcription factors indicated in Figure 6-6 was constructed. Figure 6-6 shows that the expression of *gadX* and *ibpA* increased the viability of *E. coli* 2.9- and 3.5-fold. Note that the growths of the overexpression strains and the wild type remained comparable in the absence of isooctane.

Figure 6-5 Relative viability of single transcription factor-deleted mutants in the presence of isooctane.

Each of the transcription factors whose activity was perturbed by isooctane, predicted by NCA, was deleted in *E. coli*. Relative viability was determined as a ratio of the viability of the mutant to that of the wild type (WT). The relative viabilities shown are statistically significant for WT and all mutants.

Figure 6-6 Effects of overexpression of the transcription factors whose activity was perturbed on *E. coli* viability in the presence of isooctane

Each of the genes encoding the perturbed transcription factors was overexpressed in pBAD30 plasmid (+) in the corresponding knock-out mutant. Relative viability was determined as follows. The viability of the treated cells was normalized to that of the treated cells containing the vector without the gene insert, and then to that of the wild type (WT). WT contained pBAD30 without any gene insert (-).

This result suggests that the increased expression of *gadX* and *ibpA* improved the tolerance of *E. coli*, and defense mechanisms related to acid stress and heat stress might be important in tolerance against isooctane.

To further verify the essential role of transcription factors in cellular tolerance to isooctane, and to investigate possible interactions between two transcription factors, a range of double knock-out mutants lacking two perturbed transcription factors were constructed. Figure 6-7 shows that the deletion of *fis/fur*, *fis/ihfA*, and *fis/ihfB* significantly decreased the viability of *E. coli* compared to the wild type and the corresponding single knock-out mutants. This result suggests that Fis and Fur, or Fis and IHF might play cooperative or complementary roles in tolerance against isooctane. It is noteworthy that Fis is a small protein which maintains DNA structure (Schneider, Lurz *et al.* 2001) and also a global transcriptional factor which regulates expression of genes involved in various cellular functions including metabolisms, cell morphology and stress responses (Bradley, Beach *et al.* 2007). IHF is a general DNA-binding transcriptional regulator associated with DNA supercoiling and DNA duplex destabilization. Also, IHF is related to the stabilization of nucleoproteins involved in replication, recombination and transcription of many genes (Dhavan, Crothers *et al.* 2002; Swinger and Rice 2004). Considering possible interactive actions between Fis and IHF in maintenance of DNA structures, the decreased tolerance of the *fis/ihf* knock-out mutant might be due to destabilized DNA structures.

Lastly, part of the double knock-out mutants conferred increased tolerance against isooctane. Figure 6-7 indicates that the double knock-out mutants lacking *agaR/ihfA* and *agaR/fis* increased tolerance against isooctane 3.7- and 4.4-fold, respectively. Note that the growths of these knock-out mutants and the wild type remained comparable in

the absence of isooctane. AgaR is a repressor of *aga* gene cluster related to transport and metabolism of N-acetylgalactosamine and d-galactosamine (Ray and Larson 2004). Of another interest is that the double knock-out of *fis*, *ihfA*, and *ihfB* with *agaR* resulted in increased tolerance, while the single knock-out of each of *fis*, *ihfA*, and *ihfB* and the double knock-out of *fis/ihfA* and *fis/ihfB* led to a decrease in isooctane tolerance (Figure 6-5 and Figure 6-7, respectively). The results above suggest that isooctane tolerance can be significantly improved by altering the regulatory networks of transcriptional factors whose activity is perturbed by isooctane stress. Based on the above-mentioned results, a regulatory network of transcription factors, genes involved in cellular responses to isooctane in *E. coli* was proposed in Figure 6-8. Transcriptomic, physiological changes are summarized with proposed toxic effects of isooctane.

Figure 6-7 Relative viability of *E. coli* lacking two transcription factor genes in the presence of isooctane.

A range of combinations of transcription factors whose activity was perturbed, predicted by NCA, were deleted in *E. coli*. Relative viability was determined as a ratio of the viability of the mutant to that of the wild type (WT). The relative viabilities shown are statistically significant for WT and all mutants.

6.5 Conclusions

This chapter presented an engineering framework that harnesses systems biology-derived regulatory information to improve microbial phenotypes. In particular, it was demonstrated that desired phenotypes of microbes could be achieved by reconstitution of underlying genetic regulatory networks. I first identified genetic regulatory networks and essential transcriptional determinants using transcriptome analysis. Then, the regulatory networks were reconstituted by knocking out and overexpressing a range of combinations of transcription factors, whose activity was predictably perturbed during isooctane response. It was shown that this approach significantly improved isooctane tolerance and especially, engineered strains lacking and overexpressing some of the perturbed transcription factors underwent 3- to 5-fold increase in isooctane tolerance. This microbe with improved tolerance to isooctane can be exploited for biochemical processes, fuel oil bioremediation and metabolic engineering for biofuel production. I envision that the engineering framework employed in this study can contribute to developing other microbes with desired phenotypes.

Figure 6-8 Conceptual genetic regulatory networks in isooctane stress responses

CHAPTER 7 **Conclusions and Future considerations**

Both chemical and biological reactions produce substances used in daily basis. Traditionally, it is food products that were mainly manufactured by biological reactions of microorganisms such as yeast and lactic acid bacteria. However, due to recent advances in genetic engineering techniques, and floods of genetic information available, valuable chemicals, drugs and medicinal proteins other than food products can now be produced through biochemical processes in the industry. Biological systems, particularly microbial hosts, have been used and manipulated to perform naturally existing or heterologous biochemical reactions more efficiently. However, some chemicals of which production in microbial hosts are economically promising have been found to be not efficient as much as previously expected. Although genetics and metabolisms of numerous microorganisms have been extensively studied by microbiologists and others, more intensive and condition-specific understanding of microbial systems has recently gained much significance, particularly to overcome the limitation of industrial bioprocesses. Beyond the elucidation of cellular responses under specific conditions, this thesis intended to harness rational engineering approaches to optimize microbial strains.

Therefore, the work in this thesis aimed to develop a systems-level cellular engineering framework for rational strain optimization through characterization and reconstitution of relevant regulatory systems. Towards improved bioprocesses, *E. coli* K12 was used throughout this work as a model organism (Chapter 2). Comprehensive understanding of cellular responses in *E. coli* was achieved first by transcriptomic profiling of gene expressions and then further by analyses based on the bioinformatics algorithms, such as GO analysis and NCA (Chapter 3). Identified genes, transcription

factors and cellular functions were deleted or disturbed to verify their involvement in cellular responses under the three experimental conditions. This systems engineering framework was applied to the following three bioprocess-relevant conditions; gene disruption in quorum sensing (Chapter 4), oxidative stress under anaerobic condition (Chapter 5) and isooctane solvent stress (Chapter 6).

Chapter 4 discussed cellular responses to knockout of *luxS* gene from the genome. This study revealed that deletion of *luxS* gene in *E. coli* K12 increased swimming motility and flagella formation. The microarray data analysis showed that genes involved in the swimming motility were highly expressed. In addition, active swimming motility and flagella formation was confirmed by microscopy. Since connectivity between LuxS or quorum sensing and swimming motility has not previously been demonstrated, one objective of this study was to identify molecular components that regulate increased swimming motility in the absence of *luxS* gene. It is noteworthy that this is the first report that suggests a relationship between absence of AI-2 synthesizing enzyme, LuxS, and the identified regulatory factors, FlhDC, FliA, H-NS, GadX and c-di-GMP. Since the regulatory roles of FlhDC, FliA and H-NS have been well documented regarding motility and flagellar synthesis, this result strongly supported that increased swimming motility was regulated through these transcription factors in *luxS*-deleted mutants. On the other hand, the other two factors, GadX and c-di-GMP, have been reported for their roles in acid stress resistance mechanisms and in biofilm formation processes, respectively. GadX and c-di-GMP have previously not been investigated for their potential connectivity with quorum sensing or with swimming motility. Therefore, identification of GadX and c-di-GMP provided a new platform that physiological roles of *luxS* gene would be further investigated towards improved bioprocess that utilizes quorum sensing as a feedback

regulatory mechanism. In addition to bioprocesses, roles of *luxS* gene can be further investigated in relation to bacterial infections that accompany biofilm formation mediated by quorum sensing of bacterial populations.

While transcriptomic responses and regulatory factors that are specific to genetic disturbance in *luxS* gene were studied in Chapter 4, genetic regulatory networks for oxidative stress and isooctane stress were characterized in Chapters 5 and 6, respectively. Oxidative stress is a term that describes biological conditions developed by harmful oxidants such as ROS. Oxidative stress has been a topic of extensive researches because oxidative stress is an important phenomenon in defense mechanisms of plants and mammalian cells against bacterial infection, and in the course of aerobiosis. Unlike the previous works, this study exploited the fact that cellular responses to oxidative stress had not been studied under anaerobic culture conditions, and this study highlighted the findings in relation to fermentation process. It was interesting that transcriptomic responses of *E. coli* against oxidative stress showed different patterns under aerobic and anaerobic conditions. The different patterns were found in genes involved in anaerobic respiration, nucleotide de novo synthesis, nitrogen metabolism and TCA cycle. This result suggests that cellular responses to oxidative stress reflect transcriptional regulations of not only directly related defense mechanisms, but also irrelevant elementary metabolisms and cellular functions.

With the emphasis on the transcriptional regulation of anaerobic or aerobic respiration under oxidative stress, FNR was identified as an essential transcription factor, of which deletion abolished resistance of *E. coli* to oxidative stress under anaerobic condition. In addition, it was found that deletion of *arcB* and altered NADH availability increased tolerance of *E. coli* against oxidative stress under anaerobic and

aerobic condition, respectively. Based on the fact that the above factors regulate either transcription of a group of genes or metabolic fluxes, it has been rationalized that deletion of a transcription factor would reconstitute the genetic regulatory networks developed in response to specific conditions used in this thesis. Supporting this rationale, it has been discussed that disturbance in above-mentioned regulatory factors did not abolish growth ability of *E. coli*. Therefore, this thesis suggested that functions of a regulatory factor could be complemented by other regulatory factors. In other words, *E. coli* developed alternative regulatory networks in response to disturbance in *fnr*, *arcB* and NADH availability. Consequently, the re-organized regulatory networks resulted in improved resistance to oxidative stress under anaerobic and aerobic conditions in *E. coli*. Therefore, this thesis proposed a strategy that optimizes host strains against oxidative stress that microbial cells experience during anaerobic or fermentative bioprocesses.

Because low microbial tolerance to solvents and potential biofuel molecules has limited biochemical processes, stress responses of *E. coli* to hydrocarbon isooctane have been characterized in Chapter 6. It was interesting that the cellular responses to isooctane were accompanied with previously reported defense mechanisms against heat shock, oxidative stress, iron homeostasis and acid stress. In addition, this study showed that isooctane changed morphology of *E. coli* including the development of filaments and flagella-like structures on the surface. On the other hand, through systematic analysis of transcriptomic responses of *E. coli* to isooctane, essential transcription factors that determined genetic regulatory networks in the presence of isooctane were identified. By using the systems engineering framework developed, the genetic regulatory networks were reconstituted by genetic manipulation of transcription factors. It was further found that deletion of such transcription factors as

RpiR, TrpR, PhoB, Lrp, KdpE and CRP, increased tolerance of *E. coli* to isooctane. Subsequently, it was demonstrated that isooctane tolerance could be increased or decreased by deleting different combinations of transcription factors, which re-established the genetic regulatory networks for tolerance to isooctane. Through this study, it was shown that the optimized microbial strains with higher tolerance could be harnessed for biochemical processes particularly related to bioremediation of organic solvents contaminants and metabolic engineering for bioproduction of fuel molecules.

Although this thesis provides a foundation for rational strain optimization, there are several aspects that require further studies for future applications. One of the aspects is that interpretation of microarray data is limited to size of database available to research communities. For example, functional analyses of microarray data can only be performed based on available functional annotations such GO terms and KEGG pathways. If a gene is not characterized in detail and thus lacks information on its functions, location in cellular compartment, functional annotations, this gene will be excluded even if the gene might play a significant role under experimental conditions. Similarly, NCA is restrained by known connectivity between genes and transcription factors. If information on transcriptional regulation of a gene is not sufficient, activity changes of transcription factors regulating the gene cannot be identified regardless of their importance. Therefore, it is essential for researchers to build public databases that have sufficient information on genetic components to characterize cellular responses at the systems level and to identify essential genetic regulatory networks in a microbial host. For further research, on the other hand, profiles of proteins and metabolites can be monitored together with mRNA profiles to enrich bioinformatics analysis. This approach requires development of computer-based algorithms or models that combine and analyze three different data sets quantitatively and qualitatively. Lastly, the

performance of optimized strains with improved tolerance under stress conditions need to be tested for a desired bioprocess. This requires the engineering of microbial hosts that perform desired biological reactions as well as strain optimization. With the aforementioned future works taken into consideration, the systems engineering framework presented in this thesis has great potential to be readily extended to a wider range of applications.

Moreover, this thesis provides fundamental platforms and interesting hypotheses that can be used to further explore cellular responses characterized in Chapters 4, 5 and 6 at a molecular level. For example, Chapter 4, which addressed *luxS*-mediated quorum sensing and swimming motility, can be extended to a research of molecular interactions between identified transcription factors and genes related to chemotaxis and flagella biosynthesis genes. Swimming motility test and transcriptomic profiling of transcription factor-deleted double mutants provide more insights into linkages between *luxS*-mediated genetic regulation and swimming motility. In addition, Chapter 5 and Chapter 6 can be enriched not only by engineering tolerance against oxidative stress or isooctane, but also by elucidating molecular mechanisms underlying the improved tolerance of *E. coli*. Particularly, it will be interesting to re-characterize systems-level responses against stress exposure in the engineered strains, where the regulatory networks are reconstituted by deleting or over-expressing the identified regulatory factors. These future works will provide deeper understanding of microbial cells and more efficient target mechanisms that can be used for advanced strain optimization.

Overall, this thesis presents the development of a systems-level strain optimization framework towards the rational engineering of microbial cells for improved

bioprocesses. Besides systems-level characterization of cellular responses to environmental conditions, molecular targets to engineer were systematically identified from high-throughput transcriptomic data. This thesis further demonstrates that the phenotypes of microbial cells can be improved through re-wiring of underlying genetic regulatory networks. Lastly, the systems-level cellular engineering framework developed in this work is readily applicable to designing a wide range of rational strain optimization strategies for improved bioprocesses, and to developing synthetic microbes that exhibit other desired phenotypic characteristics.

BIBLIOGRAPHY

- Ahmed, N. A., F. C. Petersen, et al. (2007). "AI-2 quorum sensing affects antibiotic susceptibility in *Streptococcus anginosus*." *J Antimicrob Chemother* **60**(1): 49-53.
- Angov, E. (2011). "Codon usage: nature's roadmap to expression and folding of proteins." *Biotechnology journal* **6**(6): 650-659.
- Aono, R. and H. Kobayashi (1997). "Cell surface properties of organic solvent-tolerant mutants of *Escherichia coli* K-12." *Appl Environ Microbiol* **63**(9): 3637-3642.
- Aono, R., M. Kobayashi, et al. (1995). "A close correlation between improvement of organic solvent tolerance levels and alteration of resistance toward low levels of multiple antibiotics in *Escherichia coli*." *Biosci Biotechnol Biochem* **59**(2): 213-218.
- Apel, D. and M. G. Surette (2008). "Bringing order to a complex molecular machine: the assembly of the bacterial flagella." *Biochim Biophys Acta* **1778**(9): 1851-1858.
- Asako, H., H. Nakajima, et al. (1997). "Organic solvent tolerance and antibiotic resistance increased by overexpression of *marA* in *Escherichia coli*." *Appl Environ Microbiol* **63**(4): 1428-1433.
- Atsumi, S., T. Y. Wu, et al. (2010). "Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*." *Molecular systems biology* **6**: 449.
- Balaji, S., M. M. Babu, et al. (2007). "Interplay Between Network Structures, Regulatory Modes and Sensing Mechanisms of Transcription Factors in the Transcriptional Regulatory Network of *E. coli*." *Journal of Molecular Biology* **372**(4): 1108-1122.
- Balestrino, D., J. A. Haagensen, et al. (2005). "Characterization of type 2 quorum sensing in *Klebsiella pneumoniae* and relationship with biofilm formation." *J Bacteriol* **187**(8): 2870-2880.
- Balleza, E., L. N. Lopez-Bojorquez, et al. (2009). "Regulation by transcription factors in bacteria: beyond description." *FEMS Microbiol Rev* **33**(1): 133-151.
- Balleza, E., L. N. Lopez-Bojorquez, et al. (2009). "Regulation by transcription factors in bacteria: beyond description." *Fems Microbiology Reviews* **33**(1): 133-151.
- Barras, F., L. Loiseau, et al. (2005). "How *Escherichia coli* and *Saccharomyces cerevisiae* build Fe/S proteins." *Advances in microbial physiology* **50**: 41-101.
- Bassler, B. L. and R. Losick (2006). "Bacterially speaking." *Cell* **125**(2): 237-246.
- Bekker, M., S. Alexeeva, et al. (2010). "The *ArcBA* two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool." *J Bacteriol* **192**(3): 746-754.

- Berrios-Rivera, S. J., G. N. Bennett, et al. (2002). "Metabolic engineering of *Escherichia coli*: increase of NADH availability by overexpressing an NAD(+)-dependent formate dehydrogenase." Metab Eng **4**(3): 217-229.
- Blattner, F. R., G. Plunkett, 3rd, et al. (1997). "The complete genome sequence of *Escherichia coli* K-12." Science **277**(5331): 1453-1462.
- Borden, J. R. and E. T. Papoutsakis (2007). "Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum*." Appl Environ Microbiol **73**(9): 3061-3068.
- Bradley, M. D., M. B. Beach, et al. (2007). "Effects of Fis on *Escherichia coli* gene expression during different growth stages." Microbiology **153**(Pt 9): 2922-2940.
- Brynildsen, M. P. and J. C. Liao (2009). "An integrated network approach identifies the isobutanol response network of *Escherichia coli*." Mol Syst Biol **5**: 277.
- Burgess, N. A., D. F. Kirke, et al. (2002). "LuxS-dependent quorum sensing in *Porphyromonas gingivalis* modulates protease and haemagglutinin activities but is not essential for virulence." Microbiology **148**(Pt 3): 763-772.
- Chang, M. C. and J. D. Keasling (2006). "Production of isoprenoid pharmaceuticals by engineered microbes." Nature chemical biology **2**(12): 674-681.
- Chang, M. W., F. Toghrol, et al. (2007). "Toxicogenomic response to chlorination includes induction of major virulence genes in *Staphylococcus aureus*." Environ Sci Technol **41**(21): 7570-7575.
- Chang, W., D. A. Small, et al. (2006). "Global transcriptome analysis of *Staphylococcus aureus* response to hydrogen peroxide." J Bacteriol **188**(4): 1648-1659.
- Cheesbrough, T. M. and P. E. Kolattukudy (1984). "Alkane Biosynthesis by Decarbonylation of Aldehydes Catalyzed by a Particulate Preparation from *Pisum-Sativum*." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **81**(21): 6613-6617.
- Choudhary, S. and C. Schmidt-Dannert (2010). "Applications of quorum sensing in biotechnology." Applied Microbiology and Biotechnology **86**(5): 1267-1279.
- Compan, I. and D. Touati (1994). "Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA." Molecular Microbiology **11**(5): 955-964.
- Coulthurst, S. J., C. L. Kurz, et al. (2004). "luxS mutants of *Serratia* defective in autoinducer-2-dependent 'quorum sensing' show strain-dependent impacts on virulence and production of carbapenem and prodigiosin." Microbiology **150**(Pt 6): 1901-1910.

- Cunningham, L., M. J. Gruer, et al. (1997). "Transcriptional regulation of the aconitase genes (acnA and acnB) of Escherichia coli." Microbiology **143** (Pt 12): 3795-3805.
- Cunningham, R. P., S. M. Saporito, et al. (1986). "Endonuclease IV (nfo) mutant of Escherichia coli." Journal of Bacteriology **168**(3): 1120-1127.
- Darnton, N. C. and H. C. Berg (2008). "Bacterial flagella are firmly anchored." J Bacteriol **190**(24): 8223-8224.
- Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." Proceedings of the National Academy of Sciences of the United States of America **97**(12): 6640-6645.
- de Marco, A. (2009). "Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli." Microbial cell factories **8**: 26.
- Defoirdt, T., P. Bossier, et al. (2005). "The impact of mutations in the quorum sensing systems of Aeromonas hydrophila, Vibrio anguillarum and Vibrio harveyi on their virulence towards gnotobiotically cultured Artemia franciscana." Environ Microbiol **7**(8): 1239-1247.
- Dhavan, G. M., D. M. Crothers, et al. (2002). "Concerted binding and bending of DNA by Escherichia coli integration host factor." J Mol Biol **315**(5): 1027-1037.
- Doherty, N., M. T. Holden, et al. (2006). "Functional analysis of luxS in Staphylococcus aureus reveals a role in metabolism but not quorum sensing." J Bacteriol **188**(8): 2885-2897.
- Dunlop, M. J. (2011). "Engineering microbes for tolerance to next-generation biofuels." Biotechnology for biofuels **4**: 32.
- Elvers, K. T. and S. F. Park (2002). "Quorum sensing in Campylobacter jejuni: detection of a luxS encoded signalling molecule." Microbiology **148**(Pt 5): 1475-1481.
- Engelberg-Kulka, H., S. Amitai, et al. (2006). "Bacterial programmed cell death and multicellular behavior in bacteria." PLoS genetics **2**(10): e135.
- Galas, D. J. and A. Schmitz (1978). "DNAse footprinting: a simple method for the detection of protein-DNA binding specificity." Nucleic acids research **5**(9): 3157-3170.
- Gama-Castro, S., V. Jimenez-Jacinto, et al. (2008). "RegulonDB (version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation." Nucleic Acids Res **36**(Database issue): D120-124.

- Gao, Y., J. Song, et al. (2009). "The luxS gene is involved in AI-2 production, pathogenicity, and some phenotypes in *Erwinia amylovora*." Curr Microbiol **58**(1): 1-10.
- Georgellis, D., O. Kwon, et al. (2001). "Quinones as the redox signal for the arc two-component system of bacteria." Science **292**(5525): 2314-2316.
- Georgellis, D., A. S. Lynch, et al. (1997). "In vitro phosphorylation study of the arc two-component signal transduction system of *Escherichia coli*." J Bacteriol **179**(17): 5429-5435.
- Gerasimova, A. V., D. A. Rodionov, et al. (2001). "[Computer analysis of regulatory signals in bacterial genomes. Fnr binding segments]." Molekuliarnaia biologiya **35**(6): 1001-1009.
- Gething, M. J. and J. Sambrook (1992). "Protein folding in the cell." Nature **355**(6355): 33-45.
- Gong, P.-F. and J.-H. Xu (2005). "Bio-resolution of a chiral epoxide using whole cells of *Bacillus megaterium* ECU1001 in a biphasic system." Enzyme and Microbial Technology **36**(2-3): 252-257.
- Gong, P. F. and J. H. Xu (2005). "Bio-resolution of a chiral epoxide using whole cells of *Bacillus megaterium* ECU1001 in a biphasic system." Enzyme and Microbial Technology **36**(2-3): 252-257.
- Gunasekera, T. S., L. N. Csonka, et al. (2008). "Genome-wide transcriptional responses of *Escherichia coli* K-12 to continuous osmotic and heat stresses." Journal of Bacteriology **190**(10): 3712-3720.
- Gunsalus, R. P. and S. J. Park (1994). "Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons." Research in microbiology **145**(5-6): 437-450.
- Gustafsson, C., S. Govindarajan, et al. (2004). "Codon bias and heterologous protein expression." Trends in Biotechnology **22**(7): 346-353.
- Guzman, L. M., D. Belin, et al. (1995). "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter." Journal of Bacteriology **177**(14): 4121-4130.
- Hardie, K. R. and K. Heurlier (2008). "Establishing bacterial communities by 'word of mouth': LuxS and autoinducer 2 in biofilm development." Nat Rev Microbiol **6**(8): 635-643.
- Harshey, R. M. (2003). "Bacterial motility on a surface: many ways to a common goal." Annual review of microbiology **57**: 249-273.
- Harshey, R. M. and T. Matsuyama (1994). "Dimorphic transition in *Escherichia coli* and *Salmonella typhimurium*: surface-induced differentiation into hyperflagellate swarmer cells." Proc Natl Acad Sci U S A **91**(18): 8631-8635.

- He, Y., J. G. Frye, et al. (2008). "Analysis of AI-2/LuxS-dependent transcription in *Campylobacter jejuni* strain 81-176." Foodborne Pathog Dis **5**(4): 399-415.
- Heipieper, H. J., F. J. Weber, et al. (1994). "Mechanisms of Resistance of Whole Cells to Toxic Organic-Solvents." Trends in Biotechnology **12**(10): 409-415.
- Henle, E. S., Z. Han, et al. (1999). "Sequence-specific DNA cleavage by Fe²⁺-mediated fenton reactions has possible biological implications." The Journal of biological chemistry **274**(2): 962-971.
- Hesse, W. R. and M. J. Kim (2009). "Visualization of flagellar interactions on bacterial carpets." J Microsc **233**(2): 302-308.
- Huang, D. W., B. T. Sherman, et al. (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nature Protocols **4**(1): 44-57.
- Hughes, K. T., K. L. Gillen, et al. (1993). "Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator." Science **262**(5137): 1277-1280.
- Ibrahim, G. F., G. H. Fleet, et al. (1985). "Method for the isolation of highly purified *Salmonella* flagellins." J Clin Microbiol **22**(6): 1040-1044.
- Iinoya, K., T. Kotani, et al. (2009). "Engineering of the yeast antioxidant enzyme Mpr1 for enhanced activity and stability." Biotechnol Bioeng **103**(2): 341-352.
- Imlay, J. A. (1995). "A metabolic enzyme that rapidly produces superoxide, fumarate reductase of *Escherichia coli*." The Journal of biological chemistry **270**(34): 19767-19777.
- Imlay, J. A. (2002). "How oxygen damages microbes: oxygen tolerance and obligate anaerobiosis." Advances in microbial physiology **46**: 111-153.
- Imlay, J. A. (2003). "Pathways of oxidative damage." Annu Rev Microbiol **57**: 395-418.
- Imlay, J. A. (2008). "Cellular defenses against superoxide and hydrogen peroxide." Annu Rev Biochem **77**: 755-776.
- Imlay, J. A. and S. Linn (1987). "Mutagenesis and stress responses induced in *Escherichia coli* by hydrogen peroxide." Journal of Bacteriology **169**(7): 2967-2976.
- Ingledeu, W. J. and R. K. Poole (1984). "The Respiratory Chains of *Escherichia-Coli*." Microbiological Reviews **48**(3): 222-271.
- Ingram, L. O. (1977). "Changes in lipid composition of *Escherichia coli* resulting from growth with organic solvents and with food additives." Appl Environ Microbiol **33**(5): 1233-1236.

- Inui, M., H. Kawaguchi, et al. (2004). "Metabolic engineering of *Corynebacterium glutamicum* for fuel ethanol production under oxygen-deprivation conditions." Journal of molecular microbiology and biotechnology **8**(4): 243-254.
- Isken, S. and J. A. de Bont (1996). "Active efflux of toluene in a solvent-resistant bacterium." J Bacteriol **178**(20): 6056-6058.
- Jang, H. J., M. W. Chang, et al. (2008). "Microarray analysis of toxicogenomic effects of triclosan on *Staphylococcus aureus*." Appl Microbiol Biotechnol **78**(4): 695-707.
- Jeon, B. and K. Itoh (2007). "Production of shiga toxin by a luxS mutant of *Escherichia coli* O157:H7 in vivo and in vitro." Microbiol Immunol **51**(4): 391-396.
- Jeon, B., K. Itoh, et al. (2003). "Effects of quorum sensing on flaA transcription and autoagglutination in *Campylobacter jejuni*." Microbiol Immunol **47**(11): 833-839.
- Jervis, A. J., J. C. Crack, et al. (2009). "The O₂ sensitivity of the transcription factor FNR is controlled by Ser24 modulating the kinetics of [4Fe-4S] to [2Fe-2S] conversion." Proc Natl Acad Sci U S A **106**(12): 4659-4664.
- Johnson, D. L., S. A. Middleton, et al. (1996). "Refolding, purification, and characterization of human erythropoietin binding protein produced in *Escherichia coli*." Protein expression and purification **7**(1): 104-113.
- Johnston, S., S. Lin, et al. (2009). "A genomic island of the sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough promotes survival under stress conditions while decreasing the efficiency of anaerobic growth." Environ Microbiol **11**(4): 981-991.
- Jordan, D. M., V. Sperandio, et al. (2005). "Colonization of gnotobiotic piglets by a luxS mutant strain of *Escherichia coli* O157:H7." Infect Immun **73**(2): 1214-1216.
- Kadner, R. J. (1974). "Transport systems for L-methionine in *Escherichia coli*." Journal of bacteriology **117**(1): 232-241.
- Kanda, T., N. Miyata, et al. (1998). "Doubly entrapped baker's yeast survives during the long-term stereoselective reduction of ethyl 3-oxobutanoate in an organic solvent." Applied Microbiology and Biotechnology **49**(4): 377-381.
- Kane, J. F. (1995). "Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*." Current opinion in biotechnology **6**(5): 494-500.
- Kang, A. and M. W. Chang (2012). "Identification and reconstitution of genetic regulatory networks for improved microbial tolerance to isooctane." Molecular bioSystems **8**(4): 1350-1358.

- Keseler, I. M., J. Collado-Vides, et al. (2011). "EcoCyc: a comprehensive database of Escherichia coli biology." Nucleic acids research **39**(Database issue): D583-590.
- Kim, H. U., T. Y. Kim, et al. (2011). "Framework for network modularization and Bayesian network analysis to investigate the perturbed metabolic network." BMC systems biology **5 Suppl 2**: S14.
- King, T., S. Lucchini, et al. (2010). "Transcriptomic analysis of Escherichia coli O157:H7 and K-12 cultures exposed to inorganic and organic acids in stationary phase reveals acidulant- and strain-specific acid tolerance responses." Appl Environ Microbiol **76**(19): 6514-6528.
- Klibanov, A. M. (2001). "Improving enzymes by using them in organic solvents." Nature **409**(6817): 241-246.
- Kohanski, M. A., D. J. Dwyer, et al. (2007). "A common mechanism of cellular death induced by bactericidal antibiotics." Cell **130**(5): 797-810.
- Korea, C. G., R. Badouraly, et al. (2010). "Escherichia coli K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities." Environ Microbiol **12**(7): 1957-1977.
- Korem, M., Y. Gov, et al. (2009). "Global gene expression in Staphylococcus aureus following exposure to alcohol." Microb Pathog.
- Korshunov, S. and J. A. Imlay (2006). "Detection and quantification of superoxide formed within the periplasm of Escherichia coli." Journal of Bacteriology **188**(17): 6326-6334.
- Korshunov, S. and J. A. Imlay (2010). "Two sources of endogenous hydrogen peroxide in Escherichia coli." Molecular Microbiology **75**(6): 1389-1401.
- Koskinen, P. E., S. R. Beck, et al. (2008). "Ethanol and hydrogen production by two thermophilic, anaerobic bacteria isolated from Icelandic geothermal areas." Biotechnology and Bioengineering **101**(4): 679-690.
- Kozlova, E. V., V. L. Popov, et al. (2008). "Mutation in the S-ribosylhomocysteinase (*luxS*) gene involved in quorum sensing affects biofilm formation and virulence in a clinical isolate of *Aeromonas hydrophila*." Microb Pathog **45**(5-6): 343-354.
- Kozyrev, D. and N. Vassinova (2000). "A method for direct cloning of Fur-regulated genes: identification of seven new Fur-regulated loci in Escherichia coli." Microbiology-Uk **146**: 3171-3182.
- Krin, E., N. Chakroun, et al. (2006). "Pleiotropic role of quorum-sensing autoinducer 2 in *Photobacterium luminescens*." Appl Environ Microbiol **72**(10): 6439-6451.
- Kuczynska-Wisnik, D., S. Kedzierska, et al. (2002). "The Escherichia coli small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins

- denatured in vivo during extreme heat shock." Microbiology **148**(Pt 6): 1757-1765.
- Laasik, E., L. Andresen, et al. (2006). "Type II quorum sensing regulates virulence in *Erwinia carotovora* ssp. *carotovora*." FEMS Microbiol Lett **258**(2): 227-234.
- Ladygina, N., E. G. Dedyukhina, et al. (2006). "A review on microbial synthesis of hydrocarbons." Process Biochemistry **41**(5): 1001-1014.
- Lennox, E. S. (1955). "Transduction of linked genetic characters of the host by bacteriophage P1." Virology **1**(2): 190-206.
- Leonardo, M. R., Y. Dailly, et al. (1996). "Role of NAD in regulating the adhE gene of *Escherichia coli*." J Bacteriol **178**(20): 6013-6018.
- Leong, S. S. and W. N. Chen (2008). "Preparing recombinant single chain antibodies " Chem Eng Sci **63**(6): 1401-1414.
- Letoffe, S., P. Delepelaire, et al. (2006). "The housekeeping dipeptide permease is the *Escherichia coli* heme transporter and functions with two optional peptide binding proteins." Proc Natl Acad Sci U S A **103**(34): 12891-12896.
- Lewis, K. (2007). "Persister cells, dormancy and infectious disease." Nat Rev Microbiol **5**(1): 48-56.
- Li, Z. and B. Dimple (1994). "SoxS, an activator of superoxide stress genes in *Escherichia coli*. Purification and interaction with DNA." The Journal of biological chemistry **269**(28): 18371-18377.
- Liao, J. C., R. Boscolo, et al. (2003). "Network component analysis: reconstruction of regulatory signals in biological systems." Proc Natl Acad Sci U S A **100**(26): 15522-15527.
- Lin, E. C. and S. Iuchi (1991). "Regulation of gene expression in fermentative and respiratory systems in *Escherichia coli* and related bacteria." Annual review of genetics **25**: 361-387.
- Lin, M. S. and H. H. Wang (1991). "Anaerobic growth and oxygen toxicity of *Rhizopus* cultures isolated from starters made by solid state fermentation." Zhonghua Minguo wei sheng wu ji mian yi xue za zhi = Chinese journal of microbiology and immunology **24**(2): 229-239.
- Ling, H., A. Kang, et al. (2010). "The absence of the luxS gene increases swimming motility and flagella synthesis in *Escherichia coli* K12." Biochem Biophys Res Commun **401**(4): 521-526.
- Liochev, S. I. and I. Fridovich (1992). "Fumarase C, the stable fumarase of *Escherichia coli*, is controlled by the soxRS regulon." Proceedings of the National Academy of Sciences of the United States of America **89**(13): 5892-5896.

- Litsanov, B., M. Brocker, et al. (2012). "Toward homosuccinate fermentation: metabolic engineering of *Corynebacterium glutamicum* for anaerobic production of succinate from glucose and formate." Applied and environmental microbiology **78**(9): 3325-3337.
- Loui, C., A. C. Chang, et al. (2009). "Role of the ArcAB two-component system in the resistance of *Escherichia coli* to reactive oxygen stress." BMC Microbiol **9**(1): 183.
- Luo, G., L. Xie, et al. (2011). "Enhancement of bioenergy production from organic wastes by two-stage anaerobic hydrogen and methane production process." Bioresource technology **102**(18): 8700-8706.
- Lyon, W. R., J. C. Madden, et al. (2001). "Mutation of luxS affects growth and virulence factor expression in *Streptococcus pyogenes*." Mol Microbiol **42**(1): 145-157.
- Malpica, R., G. R. Sandoval, et al. (2006). "Signaling by the arc two-component system provides a link between the redox state of the quinone pool and gene expression." Antioxid Redox Signal **8**(5-6): 781-795.
- Matsui, K., T. Hirayama, et al. (2006). "Screening for candidate genes involved in tolerance to organic solvents in yeast." Applied Microbiology and Biotechnology **71**(1): 75-79.
- Matsui, K., S. Teranishi, et al. (2008). "Discovery of a modified transcription factor endowing yeasts with organic-solvent tolerance and reconstruction of an organic-solvent-tolerant *Saccharomyces cerevisiae* strain." Applied and environmental microbiology **74**(13): 4222-4225.
- Matuszewska, E., J. Kwiatkowska, et al. (2008). "Escherichia coli heat-shock proteins IbpA/B are involved in resistance to oxidative stress induced by copper." Microbiology **154**(Pt 6): 1739-1747.
- Maurer, L. M., E. Yohannes, et al. (2005). "pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12." Journal of Bacteriology **187**(1): 304-319.
- McHugh, J. P., F. Rodriguez-Quinones, et al. (2003). "Global iron-dependent gene regulation in *Escherichia coli*. A new mechanism for iron homeostasis." J Biol Chem **278**(32): 29478-29486.
- Messner, K. R. and J. A. Imlay (1999). "The identification of primary sites of superoxide and hydrogen peroxide formation in the aerobic respiratory chain and sulfite reductase complex of *Escherichia coli*." The Journal of biological chemistry **274**(15): 10119-10128.
- Messner, K. R. and J. A. Imlay (2002). "Mechanism of superoxide and hydrogen peroxide formation by fumarate reductase, succinate dehydrogenase, and aspartate oxidase." The Journal of biological chemistry **277**(45): 42563-42571.

- Miura, S., W. Zou, et al. (2000). "Screening of genes involved in isooctane tolerance in *Saccharomyces cerevisiae* by using mRNA differential display." Appl Environ Microbiol **66**(11): 4883-4889.
- Moore, L. J. and P. J. Kiley (2001). "Characterization of the dimerization domain in the FNR transcription factor." The Journal of biological chemistry **276**(49): 45744-45750.
- Muff, T. J., R. M. Foster, et al. (2007). "CheX in the three-phosphatase system of bacterial chemotaxis." J Bacteriol **189**(19): 7007-7013.
- Nakajima, H., K. Kobayashi, et al. (1995). "Overexpression of the *robA* gene increases organic solvent tolerance and multiple antibiotic and heavy metal ion resistance in *Escherichia coli*." Appl Environ Microbiol **61**(6): 2302-2307.
- Nakamura, C. E. and G. M. Whited (2003). "Metabolic engineering for the microbial production of 1,3-propanediol." Current opinion in biotechnology **14**(5): 454-459.
- Ng, W. L. and B. L. Bassler (2009). "Bacterial quorum-sensing network architectures." Annu Rev Genet **43**: 197-222.
- Osaki, T., T. Hanawa, et al. (2006). "Mutation of *luxS* affects motility and infectivity of *Helicobacter pylori* in gastric mucosa of a Mongolian gerbil model." J Med Microbiol **55**(Pt 11): 1477-1485.
- Park, J. H. and S. Y. Lee (2008). "Towards systems metabolic engineering of microorganisms for amino acid production." Current opinion in biotechnology **19**(5): 454-460.
- Partridge, J. D., G. Sanguinetti, et al. (2007). "Transition of *Escherichia coli* from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components." Journal of Biological Chemistry **282**(15): 11230-11237.
- Partridge, J. D., C. Scott, et al. (2006). "*Escherichia coli* transcriptome dynamics during the transition from anaerobic to aerobic conditions." J Biol Chem **281**(38): 27806-27815.
- Patton, T. G., K. C. Rice, et al. (2005). "The *Staphylococcus aureus* *cidC* gene encodes a pyruvate oxidase that affects acetate metabolism and cell death in stationary phase." Molecular Microbiology **56**(6): 1664-1674.
- Pesavento, C., G. Becker, et al. (2008). "Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*." Genes Dev **22**(17): 2434-2446.
- Pierucci, O. (1978). "Dimensions of *Escherichia coli* at various growth rates: model for envelope growth." Journal of Bacteriology **135**(2): 559-574.
- Pinkart, H. C., J. W. Wolfram, et al. (1996). "Cell Envelope Changes in Solvent-Tolerant and Solvent-Sensitive *Pseudomonas putida* Strains following Exposure to o-Xylene." Appl Environ Microbiol **62**(3): 1129-1132.

- Rader, B. A., S. R. Campagna, et al. (2007). "The quorum-sensing molecule autoinducer 2 regulates motility and flagellar morphogenesis in *Helicobacter pylori*." J Bacteriol **189**(17): 6109-6117.
- Ray, W. K. and T. J. Larson (2004). "Application of AgaR repressor and dominant repressor variants for verification of a gene cluster involved in N-acetylgalactosamine metabolism in *Escherichia coli* K-12." Molecular microbiology **51**(3): 813-826.
- Regev-Yochay, G., K. Trzcinski, et al. (2007). "SpxB is a suicide gene of *Streptococcus pneumoniae* and confers a selective advantage in an in vivo competitive colonization model." Journal of Bacteriology **189**(18): 6532-6539.
- Rezzonico, F. and B. Duffy (2007). "The role of luxS in the fire blight pathogen *Erwinia amylovora* is limited to metabolism and does not involve quorum sensing." Mol Plant Microbe Interact **20**(10): 1284-1297.
- Rice, K. C. and K. W. Bayles (2008). "Molecular control of bacterial death and lysis." Microbiol Mol Biol Rev **72**(1): 85-109, table of contents.
- Riondet, C., R. Cachon, et al. (2000). "Extracellular oxidoreduction potential modifies carbon and electron flow in *Escherichia coli*." J Bacteriol **182**(3): 620-626.
- Ritz, D. and J. Beckwith (2001). "Roles of thiol-redox pathways in bacteria." Annual review of microbiology **55**: 21-48.
- Roifman, C. M., G. B. Mills, et al. (1985). "Functional comparison of recombinant interleukin 2 (IL-2) with IL-2-containing preparations derived from cultured cells." Cellular immunology **95**(1): 146-156.
- Rolfe, M. D., A. Ter Beek, et al. (2011). "Transcript profiling and inference of *Escherichia coli* K-12 ArcA activity across the range of physiologically relevant oxygen concentrations." J Biol Chem.
- Rutherford, B. J., R. Dahl, et al. (2010). "Functional Genomic Study of Exogenous n-Butanol Stress in *Escherichia coli*." Appl Environ Microbiol.
- Saeidi, N., C. K. Wong, et al. (2011). "Engineering microbes to sense and eradicate *Pseudomonas aeruginosa*, a human pathogen." Molecular Systems Biology **7**: 521.
- Schirmer, A., M. A. Rude, et al. (2010). "Microbial Biosynthesis of Alkanes." Science **329**(5991): 559-562.
- Schneider, R., R. Lurz, et al. (2001). "An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA." Nucleic Acids Res **29**(24): 5107-5114.
- Sedgwick, S. G., C. Ho, et al. (1991). "Mutagenic DNA repair in enterobacteria." Journal of Bacteriology **173**(18): 5604-5611.

- Sikkema, J., J. A. de Bont, et al. (1995). "Mechanisms of membrane toxicity of hydrocarbons." Microbiol Rev **59**(2): 201-222.
- Smith, D. P., T. W. Knapman, et al. (2009). "Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies." Eur J Mass Spectrom (Chichester, Eng) **15**(2): 113-130.
- Soni, K., P. Jesudhasan, et al. (2007). "Proteomic analysis to identify the role of LuxS/AI-2 mediated protein expression in Escherichia coli O157:H7." Foodborne Pathog Dis **4**(4): 463-471.
- Soutourina, O., A. Kolb, et al. (1999). "Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon." J Bacteriol **181**(24): 7500-7508.
- Sperandio, V., A. G. Torres, et al. (2001). "Quorum sensing is a global regulatory mechanism in enterohemorrhagic Escherichia coli O157:H7." J Bacteriol **183**(17): 5187-5197.
- Sperandio, V., A. G. Torres, et al. (2002). "Quorum sensing Escherichia coli regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in E. coli." Mol Microbiol **43**(3): 809-821.
- Steinbuechel, A. (2005). "Non-biodegradable biopolymers from renewable resources: perspectives and impacts." Current opinion in biotechnology **16**(6): 607-613.
- Stojiljkovic, I., A. J. Baumler, et al. (1994). "Fur regulon in gram-negative bacteria. Identification and characterization of new iron-regulated Escherichia coli genes by a fur titration assay." J Mol Biol **236**(2): 531-545.
- Surette, M. G. and B. L. Bassler (1998). "Quorum sensing in Escherichia coli and Salmonella typhimurium." Proc Natl Acad Sci U S A **95**(12): 7046-7050.
- Swinger, K. K. and P. A. Rice (2004). "IHF and HU: flexible architects of bent DNA." Curr Opin Struct Biol **14**(1): 28-35.
- Sztajer, H., A. Lemme, et al. (2008). "Autoinducer-2-regulated genes in Streptococcus mutans UA159 and global metabolic effect of the luxS mutation." J Bacteriol **190**(1): 401-415.
- Tamayo, R., J. T. Pratt, et al. (2007). "Roles of cyclic diguanylate in the regulation of bacterial pathogenesis." Annu Rev Microbiol **61**: 131-148.
- Tchou, J., H. Kasai, et al. (1991). "8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity." Proceedings of the National Academy of Sciences of the United States of America **88**(11): 4690-4694.
- Terashima, H., S. Kojima, et al. (2008). "Flagellar motility in bacteria structure and function of flagellar motor." Int Rev Cell Mol Biol **270**: 39-85.

- Tian, Y., Q. Wang, et al. (2008). "Involvement of LuxS in the regulation of motility and flagella biogenesis in *Vibrio alginolyticus*." Biosci Biotechnol Biochem **72**(4): 1063-1071.
- Trinh, C. T., J. Li, et al. (2011). "Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol." Applied and environmental microbiology **77**(14): 4894-4904.
- Tucker, D. L., N. Tucker, et al. (2002). "Gene expression profiling of the pH response in *Escherichia coli*." J Bacteriol **184**(23): 6551-6558.
- Ukibe, K., K. Hashida, et al. (2009). "Metabolic engineering of *Saccharomyces cerevisiae* for astaxanthin production and oxidative stress tolerance." Appl Environ Microbiol **75**(22): 7205-7211.
- Unden, G. and J. Bongaerts (1997). "Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors." Biochimica et biophysica acta **1320**(3): 217-234.
- Upadhyaya, G., T. M. Clancy, et al. (2012). "Effect of air-assisted backwashing on the performance of an anaerobic fixed-bed bioreactor that simultaneously removes nitrate and arsenic from drinking water sources." Water research **46**(4): 1309-1317.
- Vassinova, N. and D. Kozyrev (2000). "A method for direct cloning of fur-regulated genes: identification of seven new fur-regulated loci in *Escherichia coli*." Microbiology **146 Pt 12**: 3171-3182.
- Vendeville, A., K. Winzer, et al. (2005). "Making 'sense' of metabolism: autoinducer-2, LuxS and pathogenic bacteria." Nat Rev Microbiol **3**(5): 383-396.
- Wacker, M., D. Linton, et al. (2002). "N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*." Science **298**(5599): 1790-1793.
- Walters, M., M. P. Sircili, et al. (2006). "AI-3 synthesis is not dependent on luxS in *Escherichia coli*." J Bacteriol **188**(16): 5668-5681.
- Wang, L., J. Li, et al. (2005). "luxS-dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling." J Bacteriol **187**(24): 8350-8360.
- Weber, F. J., S. Isken, et al. (1994). "Cis/Trans Isomerization of Fatty-Acids as a Defense-Mechanism of *Pseudomonas*-Putida Strains to Toxic Concentrations of Toluene." Microbiology-Uk **140**: 2013-2017.
- Wen, Z. T. and R. A. Burne (2004). "LuxS-mediated signaling in *Streptococcus mutans* is involved in regulation of acid and oxidative stress tolerance and biofilm formation." J Bacteriol **186**(9): 2682-2691.
- Winzer, K., K. R. Hardie, et al. (2002). "Bacterial cell-to-cell communication: sorry, can't talk now - gone to lunch!" Curr Opin Microbiol **5**(2): 216-222.

- Wong, S. M., K. R. Alugupalli, et al. (2007). "The ArcA regulon and oxidative stress resistance in *Haemophilus influenzae*." Mol Microbiol **64**(5): 1375-1390.
- Xavier, K. B. and B. L. Bassler (2003). "LuxS quorum sensing: more than just a numbers game." Curr Opin Microbiol **6**(2): 191-197.
- Yang, J. N., C. Wang, et al. (2011). "Outer membrane proteome and its regulation networks in response to glucose concentration changes in *Escherichia coli*." Molecular bioSystems **7**(11): 3087-3093.
- Yuan, L., J. D. Hillman, et al. (2005). "Microarray analysis of quorum-sensing-regulated genes in *Porphyromonas gingivalis*." Infect Immun **73**(7): 4146-4154.
- Zhang, Z., G. Gosset, et al. (2005). "Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*." J Bacteriol **187**(3): 980-990.
- Zheng, M., X. Wang, et al. (2001). "DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide." J Bacteriol **183**(15): 4562-4570.
- Zhou, C., X. Qi, et al. (2010). "High potency and broad-spectrum antimicrobial peptides synthesized via ring-opening polymerization of alpha-aminoacid-N-carboxyanhydrides." Biomacromolecules **11**(1): 60-67.
- Zhu, J., C. Thakker, et al. (2011). "Effect of culture operating conditions on succinate production in a multiphase fed-batch bioreactor using an engineered *Escherichia coli* strain." Applied Microbiology and Biotechnology **92**(3): 499-508.