5-2021

Validation of Anti-Oxidative Stress Genes from Genome-Wide Screening of Escherichia coli

Carson Ercanbrack

Validation of Anti-Oxidative Stress Genes from Genome-Wide Screening of *Escherichia coli*

An Honors Thesis submitted in partial fulfillment of the requirements of Honors Studies in Chemistry with Biochemistry Concentration

By

Carson Ercanbrack

Spring 2021

Chemistry with Biochemistry Concentration

J. William Fulbright College of Arts and Sciences

The University of Arkansas

Acknowledgements

I want to express my gratitude to the University of Arkansas for providing me with such an expansive education, putting me in contact with outstanding mentors and peers, and providing an inclusive environment for my undergraduate. Additionally, the Fulbright Honors College has allowed me to become a more well-rounded individual and has inspired me to embody a life of continuous learning and critical thinking. The University of Arkansas has truly pushed me to grow and develop my skills and character so I can truly personify the best version of myself. Through this program, I have been able to realize my potential and fortify confidence in my own abilities.

I want to also thank Dr. Chenguang Fan for allowing me to work as an undergraduate researcher in his lab. Since the time I joined his lab as a sophomore, Dr. Fan has always been such an understanding and patient mentor. Coincidingly, he has vast understanding of biochemistry and other sciences and never hesitates to share that knowledge with others. Dr. Fan has always been very supportive of my future and my career aspirations. Also, thank you to Hao Chen, a graduate student that also works under Dr. Fan's guidance. Hao is always eager to help and has graciously taught me the numerous techniques used in our lab.

Finally, I would like to extend thanks to my family. They have always given me unconditional love and support. I am very grateful to have such an encouraging and nurturing environment that has enabled me to undergo the rigors of undergraduate and to pursue my goal whole-heartedly.

Table of Contents

Abstract

The primary purpose of this project is to evaluate the genes that play a role in the oxidative stress response in *Escherichia coli.* In doing so, the entire genome of *E. coli* was subject to throughput in which individual genes were determined to have a role in the bacteria's oxidative stress response. Moreover, this project focused on the validation of the genes that were able to pass the initial throughput stage. The genes were subject to two forms of validation. In the first validation technique, candidate genes were overexpressed and minimum inhibitory concentrations of hypochlorous acid were taken. Following, a second validation technique tested the minimum inhibitory concentrations of hypochlorous acid utilizing knock-out mutants. These experimental results will enable for better understanding of a common pathogen as well as ways to defend against it. Additionally, *E. coli* has often been dubbed as a model bacterium and is utilized and studied in many scientific investigations. Furthermore, this experiment may lead to a greater general understanding of the bacterium, which could assist in future studies.

Introduction

When an organism undergoes oxidative stress, the organism's cells experience an excess concentration of reactive oxygen intermediates, such as superoxide ($O_2^{\bullet -}$), hydroxyl radicals ($HO^{\bullet}$), or hydrogen peroxide (H_2O_2) [1]. These intermediates are capable of being produced naturally, primarily as a byproduct of electron transfer in the respiratory chain in the mitochondria[2]. Although these reactive oxygen species are produced endogenously within the cell, the reactive species can become detrimental to cells at both unusually high and unusually low concentrations. At high concentrations, reactive oxygen intermediates are capable of oxidizing a multitude of macromolecules, including proteins and DNA[2]. This ultimately may lead to the cessation of normal cellular function and subsequent cell death. Additionally, within humans, an excess of intermediates correlates with higher rates of cancer, cardiovascular, and neural diseases[4].

While high levels of reactive oxygen species are harmful to various organisms, low levels of endogenous reactive oxygen species can also prove disadvantageous. Abnormally low level of these intermediates can coincide with a lower levels of cell growth, development, and differentiation[3]. Thus, maintaining favorable levels of reactive oxygen species is vital for cells to maintain normal functions and efficiency.

To combat potential pathogens or unwanted bacteria, many pharmaceutical drugs and household cleaners attempt to target undesired organisms by introducing compounds, such as hydrogen peroxide and hypochlorous acid (HOCl), which attempt to generate high levels of reactive oxygen species within the pathogen or bacteria. In addition, bacteria are also subject to intermediates in other hostile environments. Plants and

animals regularly produce reactive oxygen intermediates in efforts to protect against pathogens as well[5].

Previously, it was thought that oxidative stress induced by reactive oxygen species had been well studied in *Escherichia coli*, the model bacterium. However, in more recent studies, it was found that these bacteria have other tools at its disposal to combat high concentrations of reactive oxygen species. More specifically, it was discovered that noncoding RNAs and small proteins, less than 50 amino acids, also played a role in the bacteria's oxidative stress response and have yet to be studied[6]. Prior studies, instead, tended to focus on *E. coli*'s use of enzymes to scavenge the reactive intermediates, including catalase, peroxidase, mutase[7]. The studies placed emphasis on regulatory genes, OxyR, SoxR and SoxS in particular, that are involved in the activation of detoxifying enzymes when a bacterium is undergoing oxidative stress as well. Some of these inducible detoxifying enzymes include alkyl hydroperoxide reductase and superoxide dismutase[8].

This evidence points to the fact that the knowledge of the oxidative stress response utilized by *E. coli* is still relatively limited. Therefore, this project's goal is to screen the entire genome of *E. coli* without reliance on previously established knowledge. Moreover, this project would be able to provide a complete set of genes that allow *E. coli* to survive toxic levels of reactive oxidative intermediates. By identifying all of the genes *E. coli* uses to resist oxidative stress, new ways to combat the bacteria may emerge and a greater understanding of the bacteria itself may be grasped. Since hypochlorous acid, more commonly known as bleach, is routinely used as an antiseptic, this project aimed to identify the genes in *E. coli* that allow for resistance to HOCl.

Prior to my experimentation, the growth effects of varying concentrations of HOCl on the AG1 cells containing the plasmid were measured by others in the Fan Lab. To limit the effects of deviation within the experiment, four duplicates of each plate were made. Candidates were only able to pass the initial testing if three of the four copies were able to effectively grow in 6 mM HOCl, which is three times the minimal inhibitory concentration of *E. coli.*

This earlier experimentation found that 23 genes passed initial screening. The 23 genes included: *prpE, frmB, rutC, sucC, ycbB, yccJ, hyaC, hyaE, tap, yoaE, motB, rnC, elaA, yfdY, exbD, ilvA, yhdJ, ytfI, ybhR, yedQ, marC, sanA,* and *rsuA*.

To ensure these 23 genes actually did play a role in the bacteria's stress response, two validation tests were conducted. The first form of validation was conducted by use of overexpression experiments. In this, *E. coli* cells were made to overexpress 1 of the 23 candidate genes listed previously. Thus, if the genes were truly able to assist in the bacteria's oxidative stress response, the minimum inhibitory concentration of HOCl would be predicted to increase. Conversely, the second form of validation was done so through knock-out experimentations. In doing so, minimum inhibitory concentrations of *E. coli* cells missing 1 of the 23 candidate genes were taken. As opposed to the overexpression test, the minimum inhibitory concentrations would be predicted to decrease if the candidate gene truly did allow for survival at toxic levels of HOCl.

Experimental Methods

1. Validation of Identified Genes Through Overexpression

Due to the selection pressure of hypochlorous acid, mutated genes may have occurred in *E. coli* cells that enabled cell growth despite the unfavorable conditions brought on by HOCl. Another potential problem may have resulted due to the fact that the AG1 strain of *E. coli* is not a wild-type strain. To remedy this, gene candidates that were able to withstand the initial round of screening were overexpressed in wild-type MG1655 *E. coli* K-12 strain. Additionally, the MG1655 *E. coli* K-12 is known to be highly accurate strain and is widely utilized in other *E. coli* physiology experiments[12,13]. Genes were overexpressed by employing a pCA24N plasmid vector.

Minimal inhibitory concentrations of HOCl for the various overexpressed MG1655 *E. coli* strains were determined and can be found in Figure 1. Figure 1 also displays the minimal inhibitory concentrations for a control in which a MG1655 *E. coli* strain was void of any overexpressed genes. The displayed genes mirror the genes that passed the first round of screenings.

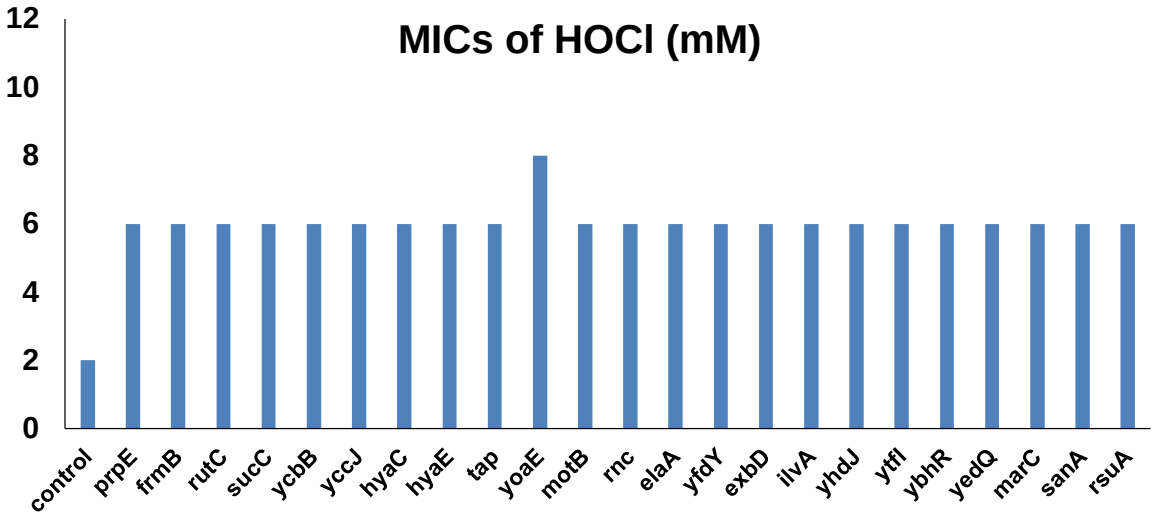

Figure 1. Minimum inhibitory concentrations (MIC) of HOCl in overexpressed gene MG1655 mutants.

2. Validation of Identified Genes Through Knock-Out Mutations

As priorly stated, the *E. coli* cells may be subject to random mutation or selective pressure due to HOCl. Thus, the genes that were able to pass the screening were subject to an additional form of validation. A second test was used to further verify the results of the previous overexpression test and initial throughput screening. Instead of the ASKA Library, the Keio Collection was used in this section of experimentation. The Keio Collection contains a set of single-gene, knockout mutants for all nonessential genes in *E. coli*[14,15]. Thus, the Keio Collection allows for robust analysis of *E. coli* cells with loss of specific gene function.

Much like the previous validation test, minimal inhibitory concentrations of HOCl were determined. These results are presented in Figure 2. Figure 2 also exhibits the minimal inhibitory concentration of the BW25113 strain, the wild-type control of the Keio Collection. The results of Figure 2 confirm a higher sensitivity to HOCl when

compared to the wild-type BW25113 strain. Furthermore, the removal of the gene candidates from the overexpression portion of testing resulted in a high sensitivity to HOCl as well. This data substantiates the idea that the genes identified do play a role in the bacteria's defense against oxidative stress.

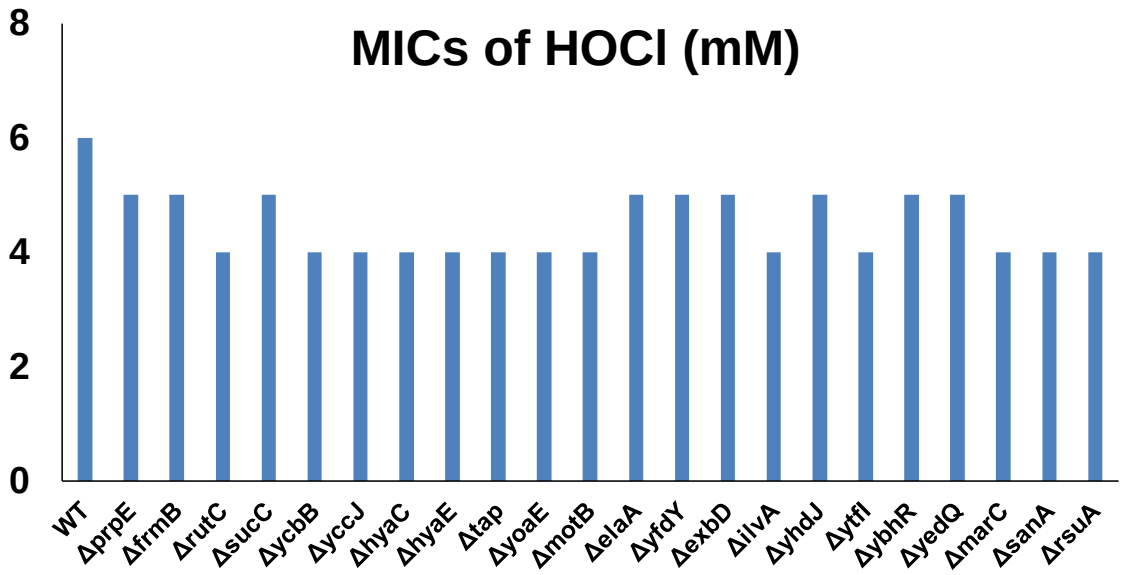

Figure 2. Minimum inhibitory concentrations (MIC) of HOCl in gene knockout BW25113 mutants.

Discussion

From Table 1, which displays the tested genes' functions and subcellular location, one can see that most of the genes that played a role in *E. coli*'s oxidative stress defense are predominately located in the cytosol or within the cellular membrane. More specifically, 12 of the 23 gene products that passed through screening and validation are found in the cytosol. Contrastingly, the remaining 11 gene products are observed in the bacteria's cell membrane. However, many of these genes have not been the subject of any studies in regard to oxidative stress defense. The localization and function of candidate genes were determined using the UniProt-GOA Database, STRING Database, and EcoCyc Database.

Table 1. Brief description of identified genes for HOCl resistance.

Genes	Encoded Protein Function	Location
prpE	Propionate-CoA ligase; Catalyzes the synthesis of propionyl-CoA from propionate and CoA.	cytosol
frmB	S-formylglutathione hydrolase; Hydrolyzes S-formylglutathione to glutathione and formate.	cytosol
rutC	Aminoacrylate peracid reductase RutC.	cytosol
sucC	Succinate-CoA synthetase; Succinyl-CoA synthetase functions in the citric acid cycle.	cytosol
ycbB	L, D-transpeptidase; Involved in peptidoglycan biosynthesis and exhibits transferase activity.	plasma membrane
yccJ	PF13993 family protein YccJ.	cytosol

hyaC	Probable nickel-dependent hydrogenase b-type cytochrome subunit.	inner membrane
hyaE	Hydrogenase-1 operon protein HyaE; appears to be a non-essential chaperone that is specific for the beta subunit of hydrogenase isoenzyme 1.	cytosol
taP	Chemotaxis methyl-accepting receptor.	inner membrane
yoaE	Integral membrane protein TerC; Involved in flavin adenine dinucleotide (FAD) binding.	inner membrane
motB	Motility protein B; MotA and MotB comprise the stator element of the flagellar motor complex which is required for rotation of the flagella.	inner membrane
rnC	Ribonuclease 3; Involved rRNA processing and digestion of double-stranded RNA formed within single-stranded substrates.	cytosol
elaA	N-acetyltransferase ElaA; Involved in transference of acyl groups other than amino-acyl groups.	cytosol
yfdY	Protein YfdY; Involved in colanic acid synthesis and biofilm development.	inner membrane
exbD	Biopolymer transport protein ExbD; ExbD is used in the TonB energy system[26].	inner membrane
ilvA	Threonine deaminase; Catalyzes the anaerobic formation of alpha-ketobutyrate and ammonia from threonine in a two-step reaction.	cytosol

yhdJ	DNA-methylase; Methylates the second adenine of the NsiI recognition sequence.	cytosol
ytfI	Uncharacterized membrane protein YtfI.	inner membrane
ybhR	Encodes probable multidrug ABC transporter permease.	inner membrane
yedQ	Catalyzes the alpha, beta-elimination reaction of D-cysteine and of several D-cysteine derivatives; May be used as a defense mechanism against D-cysteine[27].	inner membrane
marC	MarC is an inner membrane protein with six predicted transmembrane domains.	inner membrane
sanA	Multi-copy expression of sanA can complement permeability defects to vancomycin-susceptible mutants[28].	inner membrane
rsuA	Pseudouridine synthase A that is responsible for pseudouridylation of 16S rRNA at position 516[29].	cytosol

Among the genes brough forth by this study, *ycbB, yedQ, sanA, marC, ybhR. exbD, yfdY, motB, yoaE, hyaC,* and *taP*, encode proteins that are vital for the maintenance of the membrane system and its functions. *ycbB,* a gene that encodes for L, D-transpeptidase and aids in peptidoglycan synthesis, was discovered to be allow for resistance to beta-lactam antibiotics, which targets the outer membrane of the bacteria[16]. The gene, *yfdY,* has been shown to be involved in the synthesis of colonic acid, an extracellular polysaccharide that aids in *E. coli*'s ability to form biofilms[17]. These biofilms can provide the bacteria added protection against antibiotics as well as give the bacteria significant adhesive properties to hard surfaces and other bacteria that allow

them to potentially evade host defense mechanisms[18]. Additionally, *taP* and *motB* were found to encode proteins that assist in *E. coli*'s ability to move, with *taP* mediating chemotaxis and *motB* partially encoding a protein that is necessary for flagellar motion[18]. Genes, *hyaC* and *yoaE,* play roles in bioenergetics within the cell. Additionally, the genes assist in the complex transport system that exists within the cell, specifically *exbD* and *ybhR*. The remaining of the suggested genes located within the plasma membrane, *ytfI, yedQ, marC,* and *sanA*, participate in various barrier functions of the plasma membrane.

Furthermore, Figure 1 demonstrates that the greatest increase of minimum inhibitory concentration coincided with overexpression of the gene *yoaE*. Although not studied within *E. coli,* the gene *yoaE* has been more closely analyzed inside a different bacterium, *Salmonella enterica*. *S. enterica* and *E. coli* are closely related as they are both facultative anaerobes, Gram-negative, and rod-shaped[19]. Moreover, it was observed that *yoaE,* a gene that codes for a membrane protein, was upregulated by CpxR when *S. enterica* were put into chicken egg whites[20]. An egg white presents a very complex and unique environment that aides in a chicken's defense against pathogens, including its high alkalinity, the large number of lysozymes present, and its high viscosity[21,22]. Huang et. al concluded that, although it is unknown which stress factor the upregulation is in response to, *yoaE* was necessary for the bacteria's survival within the egg white environment. Therefore, due to the two bacteria's structural and genetic comparability, *E. coli* may function similarly and require upregulation in similar hostile environments.

Moving from genes localized in the membrane, the remaining suggested genes, *prpE*, *frmB, rutC, sucC, yccJ, hyaE*, *rnC, elaA, ilvA, yhdJ,* and *rsuA,* are localized within the cytosol. These genes encode proteins with a wide variety of functions. Among these

many functions included proteins involved in rRNA production (*rnC*) as well as DNA and RNA modification (*yhdJ*). Furthermore, *ilvA* contributes to isoleucine synthesis and *rsuA* is responsible for pseudouridine synthesis. Pseudouridine, one of the most common RNA modifications, has been exhibited to have important roles in various aspects of gene regulation, such spliceosomal small nuclear ribonucleoprotein biogenesis and translation fidelity[23]. Finally, genes *prpE* and *sucC* play roles in cellular metabolism, specifically propionate synthesis and succinayl-CoA synthesis, respectively.

From the study, we have found that there is still much to learn about *E. coli*'s ability to combat oxidative stress generated by HOCl, a common household disinfectant. While many of the genes have been previously studied, they were not done so in regard to their ability to allow for the bacteria to survive toxic levels of reactive oxygen species. Thus, future studies may focus the ways two or more genes interact with one another when the genes are overexpressed or deleted. Additionally, further experimentation may be necessary to determine the bacteria's response to other defense mechanisms that seek to generate excess reactive oxygen species, such as superoxide which is commonly produced by mammalian phagocytes in efforts to destroy potential pathogenic cells.

References

[1]Storz, G.; Imlayt, J. A. Oxidative Stress. Current Opinion in Microbiology 1999, 2 (2), 188–194.

[2]Tamarit, J.; Cabiscol, E.; Ros, J. Identification of the Major Oxidatively Damaged Proteins InEscherichia Coli Cells Exposed to Oxidative Stress. J. Biol. Chem. 1998, 273 (5), 3027–3032.

[3]Dahl, J.-U.; Gray, M. J.; Jakob, U. Protein Quality Control under Oxidative Stress Conditions. Journal of Molecular Biology 2015, 427 (7), 1549–1563.

[4]Jomova, K.; Jenisova, Z.; Feszterova, M.; Baros, S.; Liska, J.; Hudecova, D.; Rhodes, C. J.; Valko, M. Arsenic: Toxicity, Oxidative Stress and Human Disease. Journal of Applied Toxicology 2011, 31 (2), 95–107.

[5]Nur, I.; Munna, M. S.; Noor, R. Study of Exogenous Oxidative Stress Response in Escherichia Coli, Pseudomonas Spp., Bacillus Spp., and Salmonella Spp. Turkish Journal of Biology 2014, 38, 502–509.

[6]Storz, G. New Perspectives: Insights into Oxidative Stress from Bacterial Studies. Archives of Biochemistry and Biophysics 2016, 595, 25–27.

[7]Imlay, J. A. The Molecular Mechanisms and Physiological Consequences of Oxidative Stress: Lessons from a Model Bacterium. Nature Reviews. Microbiology; London 2013, 11 (7), 443–454.

[8]Burbank, L.; Roper, M. C. OxyR and SoxR Modulate the Inducible Oxidative Stress Response and Are Implicated During Different Stages of Infection for the Bacterial Phytopathogen Pantoea Stewartii Subsp. Stewartii. MPMI 2014, 27 (5), 479–490.

[10]Imlay, J. A. Diagnosing Oxidative Stress in Bacteria: Not as Easy as You Might Think. Current Opinion in Microbiology 2015, 24, 124–131.

[11]Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of Escherichia Coli K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection. Molecular Systems Biology 2006, 2 (1), 2006.0008.

[12]Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. The complete genome sequence of Escherichia coli K-12. Science. 1997 Sep 5;277(5331):1453-62.

[13]Soupene, E.; van Heeswijk, W. C.; Plumbridge, J.; Stewart, V.; Bertenthal, D.; Lee, H.; Prasad, G.; Paliy, O.; Charernnoppakul, P.; Kustu, S. Physiological Studies of Escherichia Coli Strain MG1655: Growth Defects and Apparent Cross-Regulation of Gene Expression. *Journal of Bacteriology* **2003**, *185* (18), 5611–5626.

[14]Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006;2:2006.0008. Epub 2006 Feb 21.

[15]Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000 Jun 6;97(12):6640-5.

[16]Caveney, N. A.; Caballero, G.; Voedts, H.; Niciforovic, A.; Worrall, L. J.; Vuckovic, M.; Fonvielle, M.; Hugonnet, J.-E.; Arthur, M.; Strynadka, N. C. Structural Insight into YcbB-Mediated Beta-Lactam Resistance in Escherichia Coli. *Nature Communications* **2019**, *10* (1).

[17]Hanna, A.; Berg, M.; Stout, V.; Razatos, A. Role of Capsular Colanic Acid in Adhesion of Uropathogenic Escherichia Coli. *Applied and Environmental Microbiology* **2003**, *69* (8), 4474–4481.

[18]Vestby, L. K.; Grønseth, T.; Simm, R.; Nesse, L. L. Bacterial Biofilm and Its Role in the Pathogenesis of Disease. *Antibiotics* **2020**, *9* (2), 59.

[19]Fabrega, A.; Vila, J. Salmonella Enterica Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *Clinical Microbiology Reviews* **2013**, *26* (2), 308–341.

[20]Huang, X.; Hu, M.; Zhou, X.; Liu, Y.; Shi, C.; Shi, X. Role of YoaE Gene Regulated by CpxR in the Survival of Salmonella Enterica Serovar Enteritidis in Antibacterial Egg White. *mSphere* **2020**, *5* (1).

[21]Baron, F.; Nau, F.; Guérin-Dubiard, C.; Bonnassie, S.; Gautier, M.; Andrews, S. C.; Jan, S. Egg White versus Salmonella Enteritidis! A Harsh Medium Meets a Resilient Pathogen. *Food Microbiology* **2016**, *53*, 82–93.

[22]Gantois, I.; Ducatelle, R.; Pasmans, F.; Haesebrouck, F.; Gast, R.; Humphrey, T. J.; Van Immerseel, F. Mechanisms of Egg Contamination by Salmonella Enteritidis. *FEMS Microbiology Reviews* **2009**, *33* (4), 718–738.

[23]Zhao, Y.; Dunker, W.; Yu, Y.-T.; Karijolich, J. The Role of Noncoding RNA Pseudouridylation in Nuclear Gene Expression Events. *Frontiers in Bioengineering and Biotechnology* **2018**, *6*.

[24]Karp, P. D.; Billington, R.; Caspi, R.; Fulcher, C. A.; Latendresse, M.; Kothari, A.; Keseler, I. M.; Krummenacker, M.; Midford, P. E.; Ong, Q.; Ong, W. K.; Paley, S. M.; Subhraveti, P. The BioCyc Collection of Microbial Genomes and Metabolic Pathways. *Briefings in Bioinformatics* **2017**, *20* (4), 1085–1093.

[25]UniProt Consortium. https://www.uniprot.org/ (accessed Mar 5, 2021).

[26]Kuehl, C. J.; Crosa, J. H. The TonB Energy Transduction Systems in Vibrio Species. *Future Microbiology* **2010**, *5* (9), 1403–1412.

[27]Sanchez-Torres, V.; Hu, H.; Wood, T. K. GGDEF Proteins YeaI, YedQ, and YfiN Reduce Early Biofilm Formation and Swimming Motility in Escherichia Coli. *Applied Microbiology and Biotechnology* **2010**, *90* (2), 651–658.

[28]Rida, S.; Caillet, J.; Alix, J. H. Amplification of a Novel Gene, SanA, Abolishes a Vancomycin-Sensitive Defect in Escherichia Coli. *Journal of bacteriology* **1996**, *178* (1), 94–102.

[29]Zhao, Y.; Dunker, W.; Yu, Y.-T.; Karijolich, J. The Role of Noncoding RNA Pseudouridylation in Nuclear Gene Expression Events. *Frontiers in Bioengineering and Biotechnology* **2018**, *6*.

[30]Imlay, J. A. Where in the World Do Bacteria Experience Oxidative Stress? *Environmental Microbiology* **2018**, *21* (2), 521–530.