Development and Characterization of a Salt-Tolerant Luciferase to Investigate Genetic Regulation in *Haloferax volcanii*

By

Christopher Davis

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Molecular Bioscience

Middle Tennessee State University
May 2018

Dissertation Committee:
Dr. Brian Robertson, Chair
Dr. Jeff Leblond
Dr. Justin Miller
Dr. Matthew Elrod-Erickson
Dr. Elliot Altman

ACKNOWLEDGEMENTS

Foremost, I would like to thank Dr. Brian Robertson for his mentorship, patience, and friendship over the years. He was always accommodating to my peculiarities and strange hours, and I gleaned valuable lessons from him teaching not only myself, but the other graduate and undergraduates in his lab. Brian has taught and inspired me to improve myself as a scientist and educator in too many ways to mention: from presentation and lecturing, to writing and experimental design, almost everything I learned about the scientific process, he had a hand in shaping. To put it simply, I am greatly appreciative and indebted to him for these lessons and gifts.

I appreciate all the time and advice offered by Dr. Jeff Leblond, Dr. Ryan Otter, Dr. Mary Farone, and Dr. Tony Farone regarding my research, teaching development, career options, and general life advice. I also want to thank all my friends (especially my MOBI cohort) for their friendship.

Finally, I would like to extend my thanks and love to my family for their love, support, and general understanding of what ended up being a "slightly" longer than expected process. Without all of you I would not have been able to finish.

ABSTRACT

Haloarchaea, especially *Haloferax volcanii*, are of particular interest for genetics research due to the wide array of tools and procedures available. However, few reporters exist for Archaea that provide a measurable output *in vivo*. Haloarchaeal proteins evolved characteristics that promote proper folding and function at high salt concentrations, but many mesohalic proteins lack these characteristics. However, proteins with salt-stabilizing mutations can lead to proper function in haloarchaea. Using laboratory directed evolution, a salt tolerant luciferase was developed and applied for use in genetic studies involving a previously described inducible promoter and genetic translation inhibitor in *H. volcanii*. A luciferase assay offers several advantages over other reporter systems: capturing data is done in real time due to a high temporal fidelity; data can be captured continuously using solid or liquid media with proper equipment; and the reporter gene itself is innocuous to the organism. Lastly, the individual mutations of the newly developed luciferase were characterized in order to identify changes in the protein leading to improved haline environment luminescence.

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES viii

LIST OF ABBREVIATIONS ix

CHAPTER I 1

1.1 INTRODUCTION 1

1.2 *HALOFERAX VOLCANII* 2

1.3 LUCIFERASE REPORTERS OF GENE EXPRESSION 6

1.4 SIGNIFICANCE OF STUDIES 8

CHAPTER II: DEVELOPMENT AND CHARACTERIZATION OF A SALT-TOLERANT LUCIFERASE TO INVESTIGATE GENETIC REGULATION IN *HALOFERAX VOLCANII* 11

2.1 ABSTRACT 11

2.2 INTRODUCTION 11

2.3 MATERIALS AND METHODS 14

2.3.1 Strains, Plasmids, and Primers 14

2.3.2 Culture and Transformation Procedures 14

2.3.3 Luminescence Image Acquisition and Quantification 15

2.3.4 Molecular Genetic Methods 15

2.3.5 Salinity Range and Plasmid Vector Bioluminescence Assays 16

2.3.6 Luminescence Monitoring of Liquid Cultures 17

2.4 RESULTS AND DISCUSSION 18

2.4.1 Mutant Luciferase Generation and Analysis of Luciferases in *H. volcanii* 18

2.4.2 Salinity Range of Mutant Luciferase 20

2.4.3 Tryptophan Specificity for P_{tnaA} Induction and Half-life Determination 20

2.4.4 Effects of Point Mutations in the Mutant Luciferases 22

2.5 CONCLUSIONS 24

2.6 ACKNOWLEDGEMENTS 26

CHAPTER III: REAL TIME MONITORING OF THE TYPE I-B CRISPR/CAS SYSTEM IN *HALOFERAX VOLCANII* USING A LUCIFERASE REPORTER 44

3.1 INTRODUCTION 44

3.2 MATERIALS AND METHODS 48
3.2.1 Strains, Plasmids, and Primers 48
3.2.2 Culture and Transformation Procedures 48
3.2.3 Molecular Genetic Methods 49
3.2.4 Invader Plasmid Construction 49
3.2.5 P_{tna} Regulated Cas6 Fragment Construction 50
3.2.5.1 Double Homologous Recombination Event Experiments 50
3.2.5.2 Single Homologous Recombination Event Experiments 53
3.3 RESULTS AND DISCUSSION 54
CHAPTER IV: GENETIC TRAPPING AND CREATION OF A PROMOTER LIBRARY UTILIZING A LUCIFERASE REPORTER SYSTEM 66
4.1 INTRODUCTION 66
4.2 MATERIALS AND METHODS 68
4.2.1 Strains, Plasmids, and Primers 68
4.2.2 Culture and Transformation Procedures 68
4.2.3 Molecular Genetic Methods 68
4.2.4 Promoter Library Creation Methods 68
4.2.5 Protein Fusion Library Creation Methods 69
4.3 RESULTS AND DISCUSSION 72
CHAPTER V 84
5.1 RESULTS AND DISCUSSION 84
5.2 FUTURE DIRECTIONS 85
REFERENCES 93

LIST OF FIGURES

CHAPTER II

Figure 2.1. Analysis of luciferases in *H. volcanii* strain H1209 transformed with pTA962 plasmids containing various luciferases. 27

Figure 2.2. Comparison of CBG99 and CBG MutB luciferases in salt-tolerant E. coli.. 28

Figure 2.3. Utilization of CBG MutB luciferase reporter in determining tryptophan specificity of PtnaA promoter in *H. volcanii* strain Hv1209 29

Figure 2.4. Half-life analysis of CBG-MutB in Hv-Minimal liquid media at three salt concentrations. 31

Figure 2.5. Effect of single point mutations on relative luminescence between CBG99 and CBG MutB 32

Figure 2.6. Combined mutations 1, 4, and 10 increase luminescence similar to CBG MutA but not to the amount seen in CBG MutB 33

Figure 2.7. Comparison of CBG99 and CBG MutB coding sequences in pTA962 displaying no significant difference when swapping vectors. 34

Figure 2.8. Luminescence of two intermediately developed luciferases in *H. volcanii*.. 35

Figure 2.9. Expression of luciferase under control of P_{tnA}, a tryptophan inducible promoter, in *H. volcanii* on solid media 36

Figure 2.10. Sequences of luciferases in this study. 39

CHAPTER III

Figure 3.1. Genetic maps of plasmids used for CRISPR/Cas experiments in this paper. 60

Figure 3.2. Type I-B CRISPR/Cas genetic elements and system 62

Figure 3.3. Double and single homologous recombination events integrate a selectable marker and tryptophan inducible promoter to regulate expression of the endogenous *cas* cassette 63

CHAPTER IV

Figure 4.1. Screen for luminescence of *H. volcanii* cells transformed with pTA131- CBG MutB-HvPromoter 77
Figure 4.2. Comparison of luminescence between three luciferase reporters at several volumes 79
Figure 4.3. Measurement of luminescence over time of protein fusion luciferase reporters. 80
Figure 4.4. Patch plate of *H. volcanii* luciferase-protein fusion colonies 81

CHAPTER V

Figure 5.1. Sequences of luciferases to compare CBR and CBG 90
Figure 5.2. Effect of single point mutations in CBR on relative luminescence compared to CBG99 and CBG MutB 91

LIST OF TABLES

CHAPTER II

Table I. Plasmids used. 40

Table II. Oligonucleotides used. 41

Table III. Luciferase mutations and amino acid changes 43

CHAPTER III

Table IV. Plasmids used 64

Table V. Oligonucleotides used 65

CHAPTER IV

Table VI. Plasmids used 82

Table VII. Oligonucleotides used 83

LIST OF ABBREVIATIONS

5-FOA- 5-fluoroorotic acid

5MLDT- 5-methoxy-DL-tryptophan

β-Gal- beta-galactosidase

CBG- click beetle green

CBG MutB- click beetle green mutant B

CBR- click beetle red

CCD- charged-coupled device

CDS- coding sequence

COCBG- codon optimized click beetle green

CRISPR- clustered randomly interspaced palindromic repeat

crRNA- CRISPR RNA

DMSO- dimethyl sulfoxide

FFLuc- firefly luciferase

His- histidine

Hv-Ca- *H. volcanii* casamino acid media

Hv-Min- *H. volcanii* minimal media

Hv-YPC- *H. volcanii* yeast extract, peptone, casamino acid media

IPTG- Isopropyl β-D-1-thiogalactopyranoside

LB- Luria-Bertani media

PAM- protospacer adjacent motif

P_{sdh}- succinate dehydrogenase promoter

P_{tna}- tryptophanase promoter

RLU- relative light unit

ROI- region of interest

Thr- threonine

Thy- thymidine

Trp- tryptophan

CHAPTER I

1.1 INTRODUCTION

The third and most recently described Domain of life, Archaea, is most notably comprised of organisms found surviving in the harshest environments where life is present. Archaeal species are found in geothermal vents, geysers, permafrost, and hypersaline lakes. In order to survive, these organisms produce unique proteins and have pathways found nowhere else on Earth; some of these unique proteins are invaluable tools used in industry and academia every day. Further investigation of Archaea must be conducted to elucidate how these proteins and genetic products function. Among the organisms found in hypersaline environments, the archaeon *Haloferax volcanii* is an ideal model organism for studies regarding the regulation and function of genetic elements in halophilic archaea. This is in part due to its genetic tractability, ease of culture, and ever expanding toolset for conducting such experiments. This dissertation focuses on the development and characterization of one such tool, a salt-tolerant luciferase reporter, and the utilization of this reporter system to answer questions related to the mechanisms of genetic regulation in *H. volcanii*. The methods and work presented underscore a need for further research regarding extremophiles in order to develop a more robust understanding of how these organisms function in their unique environments. In order to contextualize the methods and concepts used through this work, key elements and concepts are covered in the remainder of this introduction, with concepts pertaining to individual chapters found at the beginning of each chapter.

1.2 *HALOFERAX VOLCANII*

While Archaea represents the newest taxonomic Domain addition to the tree of life, many theories exist regarding the role Archaea played on early Earth, specifically regarding hyperthermophiles like those found near hydrothermal vents (Auguet et al., 2010; Stetter 2006). A comparison of 16S ribosomal RNA gene sequence between Eukarya, Bacteria, and Archaea initiated the need for adding the new domain due to the similarity in patterns archaeal species exhibited; Archaea and Eukarya also shared more similarities than Archaea and Bacteria, placing Archaea closer on the tree of life to Eukarya (Woese and Fox 1977; Woese et al. 1990). There are several distinctive features unique to archaea, particularly the presence of ether linked lipids in the cell membrane rather than the predominant ester linked lipids present in both Eukaryota and Bacteria. It is hypothesized the increased stability of ether linked lipids enabled archaea to survive more extreme environments, increasing chances of survival (Jain et al. 2014; Koga 2012; Siliakus 2017). While many scientists support the current diagram of the tree of life, several argue Archaea exist as either a direct linear branch from a shared ancestor of Eukarya, or as a synthesis between Bacteria and a Eukarya ancestor based upon further genetic comparisons (Gupta 1998; Koonin et al. 1997). Interestingly, the ability to survive harsh conditions is also implicated as an evolutionary mechanism which caused ancient gram positive bacteria to evolve into modern Archaea, as many archaeal species are resistant to such antibiotics (Cavalier-Smith 2002; Gupta 2000). Contention about where to place Archaea on the tree of life exists due to the conflicting findings regarding the relationships between archaeal, eukaryotic, and bacterial proteins, membranes, and genetics. Due to these conflicting theories on the origin of Archaea, continued

exploration and studies within the Archaea are crucial to answering taxonomic questions, and understanding how early life evolved on Earth.

The extreme environments some archaea inhabit necessitates unique properties in how these organisms carry out life processes, especially regarding metabolic pathways, gene expression, and protein structure. The mechanisms these archaea utilize to produce these products are still not fully understood, but certainly many of these previously unseen proteins and other useful products can be exploited for biotechnological industries. Of particular note are enzymes produced by Archaea; many of these enzymes retain activity in a polyextreme environment, allowing greater flexibility in industrial use (Littlechild 2015; Rampelotto 2013). The capabilities of surviving polyextreme environments have also indicated some archaeal species as possible candidates for survival in space, especially psychrophilic, xerophilic, or halophilic archaea (DasSarma et al. 2017; Stan-Lotter and Fendrihan 2011). Archaea represent new opportunities to discover naturally occurring products and processes to answer industrial needs; so long as there are ways to study the organisms in a laboratory environment.

Unfortunately, while there are many organisms of commercial and academic interest within Archaea, determining the mechanisms and methods of life processes and obtaining products of interest is difficult for some of the same reasons the organisms are of such interest. Relatively few model organisms within the domain Archaea exist. This is in part due to difficultly in three regards: sustaining environments required to culture these organisms, finding a genetically tractable organism, and developing molecular tools to manipulate organisms on a genetic level. Although Archaea are now known to be

relatively ubiquitous in nature, those found in 'normal' environments are difficult to culture or isolate, providing little help in investigating Archaea (Leigh et al. 2011; Schleper et al. 2005). Similarly, many easily isolated and genetically tractable organisms are unsuitable for study due to the difficulty and expense in culturing these organisms (Khelafia et al. 2016; Lagier et al. 2016). However, an increasing number of methods and protocols for exploiting halophilic organisms, particularly the species *Haloferax volcanii*, along with the relative ease to culture the species in a laboratory setting, grant it status as a model organism for investigating the Archaea (Allers and Ngo 2003; Allers et al. 2004; Hartman 2010).

Haloferax volcanii was first isolated in the Dead Sea (Elazari-Volcani 1943). Preferable growth conditions are at 45°C, pH around 7, and salt concentrations of 1.5M-2.5M; this salt requirement is considerably lower than many other halophilic organisms (Mullakhanbhai and Larsen 1975). *H. volcanii* grows primarily on the sediment of the Dead Sea, and is the predominant species; though time and human activity are slowly increasing the salinity of the sea, causing a drop in microbial survivability and blooms (Oren 2010). *H. volcanii* is a facultative anaerobe and is chemoorganotrophic, capable of growing on a variety of carbon sources and media (Mojica et al. 1995; Zaigler et al. 2003). The doubling time for *H. volcanii* is around 4 hours under ideal conditions, (Oren 2006; Zaigler et al. 2003).

Haloferax volcanii, like many haloarchaea, use a "salt-in" method of osmotic balance, keeping the cytoplasm as concentrated in salts as the outside environment. This method uses potassium as the predominant ion because it attracts less water than sodium,

and is also less costly to the organism than pumping ions from the cell in the "salt-out" method (Oren 1999; Oren 2009). Due to the halophilic conditions within the cell, many archaeal proteins are lower in hydrophobicity, higher in acidic residues, and may feature a 30 amino acid insertion at the N-terminus, a domain not present in homologous mesohalic proteins, to maintain halophilic character (Lanyi 1974; Leigh et al. 2011; Tadeo et. al 2009).

The pleomorphic cell is protected by an S-layer cell wall of glycoprotein and a cell membrane composed of isoprenoid ether lipids (Reuter and Maupin-Furlow 2004). This cell membrane also contains a number of carotenoids, including bacterioruberin, all of which impart a red color to the cells, visible when grown in number (Ronnekleiv 1995; Stoeckenius and Bogomolni 1982). Many of these carotenoids function to protect the microorganism from reactive oxygen species and/or offer ultraviolet light protection, as documented with bacterioruberin (Rodrigo-Banos 2015; Ronnekleiv 1995). Other carotenoids may increase membrane integrity, increasing organism sustainability under a high salt environment (Hochachka and Somero 2002; Ronnekleiv 1995). Interestingly, *H. volcanii* does not produce bacteriorhodopsin, a light driven proton pump common in many haloarchaea (Lozier et al. 1975).

The genome of the DS2 strain of *H. volcanii* is sequenced (Hartman 2010). This genome totals roughly 4 Mb in length, consisting of: one chromosome 2.85Mb in length; three smaller chromosomes ranging from 85-636 kb; and a smaller plasmid around 6.4kb in length (DasSarma 2009; Hartman 2010). Other features of *H. volcanii* include a relatively high G-C content at around 65%; a lack of RNA polyadenylation enzymes or

exosome RNA degradation, unique for *H. volcanii*; and two replication origins on the chromosome (Hartman 2010; Portnoy et al. 2005). *H. volcanii* appears to be totally dependent on RNase R, deleterious mutations to this region appear to render the organism non-viable (Portnoy and Schuster 2006). Similarly to other Haloarchaea, *H. volcanii* is highly recombinogenic, and is capable of forming viable hybrids through recombination with *H. mediterranei*, presumably to increase adaptive phenotypes within each species (Naor et al. 2012; Williams et al. 2012).

A number of genetic tools exist for *H. volcanii*: shuttle vectors with selectable markers (both antibiotic and auxotrophic), reporter genes, and gene knockouts; transformation protocols; and DNA isolation protocols (Allers et al. 2004; Cline et al. 1989; Giroux and MacNeill 2015; Holmes et al. 1991; Holmes et al. 1994). Current understanding of *H. volcanii* genetics and availability of genetic protocols and culture methods make this organism ideal to study for the proposed research.

1.3 LUCIFERASE REPORTERS OF GENE EXPRESSION

Many organisms respond to stimuli in the environment by exhibiting physiological and metabolic shifts over time, but traditionally scientists were limited to phenomena with readily observable changes. Observing these changes on a smaller scale, especially in single-celled organisms presents several challenges, as they could not be easily measured. As molecular biology advanced, the ability to measure changes in an organism over time became crucial in understanding how these organisms function. Elucidating the genetic elements responsible for gene expression requires specialized tools and techniques to report and record changes as they occur. These changes can be

measured via extraction of compounds or genetic material, and although these methods are sufficient, they require the destruction of samples; are relatively cost-prohibitive and time-intensive; are not capable of continuous monitoring unless samples are taken and processed constantly; and are not temporally sensitive to immediate changes in the environment. Fortunately, luciferase reporters can overcome all of these complications for gene regulation studies, as they can be genetically encoded to report activity as it happens *in vivo* (Kai et al. 1995; DasSarma et al. 2013). Luciferase reporters are ideal for several reasons: capturing data is done in real time due to a high temporal fidelity; the data can be captured continuously with proper equipment; and the reporter gene itself is presumably innocuous to the organism.

Beetle luciferase catalyzes the oxidative reaction of a heterocyclic compound termed a luciferin in the presence of magnesium, O_2, and ATP into AMP, carbon dioxide, oxyluciferin, and light (De Wet et al. 1987). This reaction occurs immediately, with a lag time of 25 milliseconds, and 300 milliseconds before maximum rate is achieved (DeLuca and McElroy, 1974). The reaction provides a real time measurement of luciferase activity. Luciferase reporters do have limitations, as ATP and O_2 are required in the cell, and the luciferin substrate must be added exogenously. When luciferin is added to culture media, any regulatory activity of a gene linked to the luciferase can be measured. Temporal fidelity is maintained over time due in part to the short half-life of luciferase at around 2-4 hours (Auld et al. 2009; Ignowski and Schaffer 2004). This half-life can be utilized to show faster changes in transcription rate over time when compared to longer lasting reporters like green fluorescent protein (GFP) or β-Galactosidase (βGal). Utilizing a photomultiplier and other imaging equipment, this luminescence data can be recorded

in real time, providing time course data useful in detecting when genetic elements are upregulated. The luciferase reporter system requires no harmful additives or light sources, \does not damage the organism in any way, or impede on light/dark and dark/dark circadian studies. Luciferase reporter systems are primarily utilized in a way to measure promoter activity, and answer questions of when and how long a gene is expressed. This has made luciferase the gold standard of many chronobiology studies, and has been utilized in mammalian, insect, plant, fungal, and microbial cell lines (Kondo et al. 1993; Ramanathan et al. 2012).

1.4 SIGNIFICANCE OF STUDIES

This dissertation describes work designed to accomplish two goals: the first was to improve the existing molecular toolkit for the proposed model organism *Haloferax volcanii* using methodologies commonly used in other model organisms, altered to accommodate the unique environment of *H. volcanii*; the second was to investigate questions utilizing the new reporter system, especially regarding temporal aspects of genetic regulation and expression in an archaeal organism. Archaea consists of an understudied group of microbes suspected of having many uses in an industrial and academic capacity, providing the impetus for this study.

The second chapter of this dissertation concerns the mutation of commercially available click beetle luciferase reporters over successive rounds in order to increase luminescence in a hypersaline environment. The resulting product of random mutations was then characterized in comparison to both unaltered and earlier round mutants, as well as to individual point mutations. These comparisons were conducted over a range of

salinities using solid media, along with liquid media in chemostats. The final mutant, containing eleven point mutations, was then used to measure specificity of a tryptophan-inducible promoter in *H. volcanii*, providing a new tool for use of measuring the timing of gene expression. After confirming the improved luminescence of the mutant reporter, attempts were made to incorporate the luciferase in measuring the temporality of the CRISPR/Cas system in *H. volcanii* (chapter three). While successful construction of the invader plasmid occurred, modifying the genomic DNA of *H. volcanii* through homologous recombination proved more difficult than previous research suggests, which may indicate some inherent protection mechanism for the CRISPR/Cas region designed to prevent cell destruction or CRISPR/Cas region degradation. Endogenous DNA from *H. volcanii* was also incorporated into a plasmid containing the mutant luciferase using a four base pair enzyme restriction digest, with segments of varying length ligated into the plasmid in an attempt to create a promoter library (chapter four). The construction of this library relied mainly on the ligating genomic DNA to contain a promoter and be in correct reading frame; hypothesized to be a relatively rare event. Once again, results did not match hypothesized or previously published results, particularly regarding transformation rate, or the assumption only one third of the transformants would be in-frame. The results from chapters three and four indicate possible discrepancy between previous literature regarding existing tools and results for *H. volcanii* and actual rates.

The following chapters demonstrate the utility of a luciferase reporter in investigating gene expression. Such a reporter system increases the capabilities of researchers to explore when expression in an organism occurs with high temporal fidelity. Additionally, the characterization of this reporter system provides insight into

amino acid composition and structure changes capable of increasing protein halotolerance. These findings can potentially be utilized in industry to produce more halotolerant products of interest in halophilic organisms, or to improve the halotolerance of genes important in other areas of bioindustry, such as biomining (Coker 2016).

CHAPTER II

DEVELOPMENT AND CHARACTERIZATION OF A SALT-TOLERANT LUCIFERASE TO INVESTIGATE GENETIC REGULATION IN *HALOFERAX VOLCANII*

Chris R. Davis[1], Carl H. Johnson[2], and J. Brian Robertson[1]*

[1] Department of Biology, Middle Tennessee State University, Murfreesboro, TN

[2] Department of Biological Sciences, Vanderbilt University, Nashville, TN

*** james.robertson@mtsu.edu**

2.1 ABSTRACT

Haloarchaea have evolved to thrive in hypersaline environments. *Haloferax volcanii* is of particular interest due to its genetic tractability, however few *in vivo* reporters exist for halophiles. Haloarchaeal proteins evolved characteristics that promote proper folding and function at high salt concentrations, but many mesohalic reporter proteins lack these characteristics. Mesohalic proteins that acquire salt-stabilizing mutations, however, can lead to proper function in haloarchaea. Using laboratory-directed evolution, we developed a salt-tolerant luciferase and demonstrated its use in *H. volcanii.*

2.2 INTRODUCTION

Understanding the physiology and gene regulation of extremophile archaea has proven difficult in the past due to both the environments required to culture these organisms, as well as developing tools for alteration and analysis at the genetic level. However, an increasing number of methods and protocols for exploiting *Haloferax*

volcanii, along with the relative ease to culture the species in a laboratory setting, grant this species status as a model organism for investigating the halophilic archaea (Allers and Ngo, 2003; Allers et al., 2004; Hartman et al., 2010; Reuter and Maupin-Furlow 2004). Similar to many other halophilic archaea, *H. volcanii*, uses a "salt-in" method of osmotic balance, keeping an intracellular concentration of salts isosmotic to the concentration of outside environment (Lanyi, 1974; Oren, 2009). This method uses potassium as the predominant ion, which attracts less water than sodium, and also costs less energy to the organism than an alternative strategy of pumping ions out of the cell by way of a "salt-out" method (Oren, 1999; Oren, 2009). The development of additional tools for genetic regulation studies in *H. volcanii* can provide insight into similar halophiles.

Monitoring genetic regulation requires specialized tools and techniques to report and record changes as they occur. These changes can be measured via extraction of compounds or genetic material, whereby samples are removed from the culture/population; or a genetic reporter system can be used passively to measure abundance of a reporter protein as it changes *in vivo*, the foremost tools being luciferase reporter systems and green fluorescent protein (Rosson and Nealson, 1981; Reuter and Maupin-Furlow 2004; Whitehead et al., 2009; Yin et al., 1994). Luciferase reporters are ideal for several reasons: capturing data is done in real time due to a high temporal fidelity; the data can be captured continuously; and the reporter gene product is innocuous to the organism (Hastings and Johnson 2003; Robertson et al., 2008). Luciferase catalyzes the oxidative reaction of a heterocyclic compound (luciferin) in the presence of magnesium and ATP into AMP, carbon dioxide, oxyluciferin, and light (De

Wet et al., 1987). The reaction provides a real time measurement of luciferase activity; when luciferin is present in the culture media, regulatory activity of a gene's promoter coupled to the luciferase gene can be quantified using photomultiplier tubes (PMTs) or sensitive cameras. Many luciferase assay systems exist, including assays using click beetle and firefly luciferases (for examples see Kim et al., 2007; Robertson et al., 2008; Krishnamoorthy and Robertson, 2015).

Due to the halophilic conditions within the cell, archaeal proteins are typically lower in hydrophobicity, higher in acidic residues, and adapted for halophilic protein folding at the N-terminus with amino acid insertions (Lanyi, 1974; Paul et al., 2008; Siglioccolo et al., 2011). The reduction in frequency of lysine residues and hydrophobic surface accessibility is also suspected of increasing functionality (Britton et al., 1998; Siglioccolo et al., 2011). While many of the sequences between proteins with halophilic and mesohalic homologs display similarity, there is a noted difference in terms of hydrophobicity reduction in halophiles, suggesting that site-directed changes can be studied to better determine how these differences change protein folding and function (Britton et al., 1998).

Though beetle luciferase (e.g. firefly and click beetle) has been used in a variety of organisms to study gene regulation, no studies were found in extreme halophiles. Given the nature of halophilic proteins and halophilic environments, it was assumed beetle luciferase would suffer some loss of luminescence due to improper folding or increased hydrophobicity on the protein surface. We hypothesized the relative luminescence of this luciferase could be improved for use in *H. volcanii* by altering the genetic sequence to include changes associated with increased halophilic function. We

performed three methods of alteration to study the effect of sequence changes on luminescence: error prone PCR, codon optimization, and site directed mutagenesis. In this paper, we report our development of a halophilic luciferase reporter system in *H. volcanii* which we use to characterize regulation of a tryptophan-inducible promoter, as well as investigate mutants derived from the original beetle luciferase to measure which specific mutations induce the greatest change in luminescence in *H. volcanii*.

2.3 MATERIALS AND METHODS

2.3.1 Strains, Plasmids, and Primers

The plasmids and oligonucleotides used for gene expression and DNA manipulation in both *H. volcanii* and *Escherichia coli* are shown in Tables I and II. *H. volcanii* strain H1209 was used for all luciferase studies in *H. volcanii*, which allowed selection of successful transformants on media without thymidine (Allers et al., 2010). *E. coli* strain ALS1187 is a salt-tolerant isolate of MG1655, capable of growth at 0.98 M Na^{+} concentrations, with mutations affecting five genes: *emrR*, *hfq*, *kil*, *rpsG*, and *sspA* (Wu et al., 2014).

2.3.2 Culture and Transformation Procedures

H. volcanii cultures were grown at 42°C, either on a rich (Hv-YPC) or minimal (Hv-Minimal) agar or broth, according to previously established protocols (Allers et al., 2004). The Hv-YPC medium contained yeast extract, peptone (Oxoid), and casamino acids, as well as salts which varied from 14-25% w/v. When necessary, thymidine was added at a concentration of 40 μg/mL for growth of untransformed H1209. For *H. volcanii*, polyethylene glycol (PEG)-mediated transformation was performed as previously described (Cline et al., 1989) and transformants were selected on 18% salt w/v

Hv-YPC agar. Chemically competent *E. coli* (Top 10) cells were transformed using standard chemical transformation procedure (ThermoFisher), and grown in Luria-Bertani (LB) medium containing 100 μg/mL of ampicillin at 37°C. When necessary, *E. coli* was grown on LB-Miller plates.

2.3.3 Luminescence Image Acquisition and Quantification

Transformed *H. volcanii* colonies were patched onto solid media in diagonal or mixed patterns to decrease the likelihood that luminescence was influenced by location on plate. Plates were incubated at 42°C for 1-2 days then impregnated with a total of 25 μL of 100 μM beetle luciferin (Promega E160E) in 0.5 μL drops placed evenly around the plate 1-3 h before imaging with a Bio-Rad ChemiDoc MP. Data were exported as Tagged Image File Format (TIFF) and quantified using ImageJ software. Regions of interest were selected and measured using the ImageJ software ROI Manager, which rendered the average relative luminescence of a region of interest as numerical data.

2.3.4 Molecular Genetic Methods

All enzymes and reagents for molecular genetic methods were obtained from New England Biolabs unless otherwise stated.

Click beetle green and firefly luciferase reporters were constructed from pCBG99-basic and pGL4.11[*luc2P*] respectively (Promega). The click beetle luciferase coding sequence (CBG99) was amplified using primers 1 and 2, while the firefly luciferase coding sequence was amplified using primers 3 and 4 (Table II), which included the addition of NdeI and BamHI restriction sites to the 5′ and 3′ ends, respectively. These PCR products were then ligated into pTA962, which contained both a

tryptophan-inducible promotor upstream and a synthetic terminator downstream of the insertion site (Allers et al., 2010).

Error-prone PCR of CBG99 was conducted using Roche *Taq* polymerase as described (Cirino et al., 2003). Codon optimization for the luciferase termed codon optimized click beetle green (COCBG) was done by Genewiz utilizing the "one amino acid-one codon" strategy, in which the most prevalent codon for each amino acid replaced the present codon in the click beetle green mutant CBG MutB for the same amino acid (see Fig. 2.10 for sequence of COCBG). Single point mutations of the CBG99 and CBG MutB CBG luciferases were made by site-directed mutagenesis using QuikChange Site-Directed Mutagenesis kit (Stratagene) with primers 5 through 44 (Table II). Mutant luciferases were sequenced using BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) on a 3130xl Genetic Analyzer (Hitachi).

2.3.5 Salinity Range and Plasmid Vector Bioluminescence Assays

Plasmids pTRC99a-CBG99 and pTRC99a-CBG MutB were constructed though PCR of CBG99 and CBG MutB, respectively, with primers 45 and 1. Primer 45 contained a 5′ end EcoRI restriction site. After restriction digest, the pTRC99a plasmid was treated with recombinant shrimp alkaline phosphatase (rSAP) before ligation with the PCR products. This was then transformed into *E. coli* strain ALS1187 (Wu et al., 2014) and plated on 10 cm LB-Miller plates coated with 25 µL of 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 17 µL of 100 mM beetle luciferin (Promega E160E). After incubation overnight at 37°C, the plates were imaged using a Bio-Rad ChemiDoc MP and quantified using ImageJ software.

The luciferase coding sequence (CDS) for plasmids pTA962-CBG99 and pTA962- CBG MutB were excised by NdeI and BamHI restriction digest and transferred to the reciprocal plasmid vector through ligation. The new plasmids were grown on 14.4% salt w/v Hv-YPC media impregnated with 25 µL of 100 mM beetle luciferin per 10 cm plate, imaged using a Bio-Rad ChemiDoc MP, and quantified using ImageJ software (Fig. 2.7).

2.3.6 Luminescence Monitoring of Liquid Cultures

Batch cultures were started using a 5 mL culture of H1209 transformed with pTA962-CBG MutB and used to inoculate 50 mL 14.4% salt w/v Hv-Minimal media. After overnight incubation, two 8 mL cultures were created with 2 mL overnight culture and 6 mL fresh 14.4% salt w/v Hv-Minimal media. These were then added into individual 15 mL conical glass tubes along with 8 µL of 100 mM beetle luciferin (Promega E160E), with moderate bubbling for agitation and aeration. Luminescence was recorded continuously and simultaneously for each pair of cultures using two Hamamatsu photomultiplier tubes (HC135-01) in a black box, according to previously established protocols (Krishnamoorthy and Robertson, 2015). After at least 2 h of simultaneous baseline luminescence recording, one culture received 100 µL of a 50 mM solution of either histidine, threonine, or 5-methyl-DL-tryptophan, while the other (control) culture received 100 µL of 50 mM tryptophan. Luminescence half-life was measured using a similar experimental set-up; 10 mL control and treated cultures received 100 µL of 50 mM tryptophan and incubated for 90 minutes before adding 100 µL dimethyl sulfoxide (DMSO) to the control and 100 µL anisomycin (2 mg/mL in DMSO) to the treated cultures.

2.4 RESULTS AND DISCUSSION

2.4.1 Mutant Luciferase Generation and Analysis of Luciferases in *H. volcanii*

To construct a sufficiently bright halophilic luciferase reporter, we initially tested luminescence of a couple of common beetle derived luciferases in *H. volcanii* strain H1209 by cloning their CDSs into pTNA962 expression vector downstream of P_{tnaA}, a tryptophan-inducible promoter (Allers et al., 2010). Both firefly luciferase (FFLuc) and click beetle green luciferase (CBG99) were very dim when tested at the lowest salt concentration (14.4%) in which *H. volcanii* could efficiently grow and were practically undetectable at the higher salt concentrations (23%) of its salinity range (Fig. 2.1). CBG99 was slightly brighter than FFLuc, so all follow-up development focused on CBG99. We conducted one round of error-prone PCR using CBG99 as the template, ligating the mutant PCR library into pTA962. The resulting transformed H1209 clones were analyzed for bioluminescence for which only ~25% expressed detectable levels (suggesting a stringency of mutagenesis). Five of the brightest colonies were pooled and used as a template for a second round of error-prone PCR. Again, only ~25% of second-round colonies that survived selection also exhibited detectable levels of bioluminescence, however two strains were slightly brighter than the unaltered CBG99 parent (called CBG MutA and CBG MutA2 for this investigation). From one of the brighter strains (CBG MutA), the plasmid was isolated and its mutant luciferase gene sequenced to find it contained seven point mutations. This MutA plasmid was used as a template for a third and final round of error-prone PCR, and the plasmid from the resulting brightest strain was isolated and its mutant luciferase gene (containing eleven

point mutations) was named "CBG MutB" for this study. See Table III and Fig. 2.10 for the identity and location of mutations for CBG MutA and CBG MutB.

We questioned whether CBG MutB's intensity could be further improved by codon optimization. An approximation of *H. volcanii* codon preference was predicted based on codon proportions in 51 ribosomal proteins and then used to synthesize COCBG in which the most prevalent codon for each amino acid was used for each amino acid in the CBG MutB luciferase sequence. Given the GC-content of *H. volcanii* is 65%, a GC rich, codon optimized luciferase was predicted to have increased luminescence over CBG MutB (Hartman et al., 2010). However as Fig. 2.1 shows, COCBG was not significantly brighter than the non-codon-optimized version of CBG MutB, indicating that limits in bioluminescence of CBG MutB were not due to GC content or codon usage in *H. volcanii*.

Finally, we verified that the differences in luminescence between CBG99 and CBG MutB were due to differences in the luciferase coding sequences as opposed to unintentional changes in the expression vector or *H. volcanii* genome. To do this, the CBG99 and CBG MutB CDS were cut out of their respective plasmids, ligated into the reciprocal vector (originally housing the other insert), transformed into naïve H1209, then patch-plated, imaged, and quantified (Fig. 2.8). Results indicated the expression vectors and host genome had no effect on the differences in luminescence, as CBG MutB remained brighter relative to CBG99 regardless of its vector.

2.4.2 Salinity Range of Mutant Luciferase

To test whether CBG MutB was simply a brighter luciferase in general or whether its intensity was specific to salty conditions, we tested luminescence over a range of mesohalic and halophilic salinities. We tested *H. volcanii* at salt concentrations of 14.4%, 18% and 23% salt (Fig. 2.1). However, since 14% is roughly the lowest concentration of salt in which *H. volcanii* can live, we utilized *E. coli* strain ALS1187, which contains mutations to genes known to aid in cellular salt tolerance (Wu et al. 2014) to test the luciferases at salinities less than 14%. This *E. coli* strain was transformed with pTrc99A, containing an IPTG-inducible promoter which regulated either CBG MutB or CBG99, and patched on LB-Miller plates containing 1% NaCl, beetle luciferin, and IPTG (Fig. 2.2A). The luminescence of CBG99 was significantly brighter than CBG MutB at this lower salinity, indicating the mutations present in CBG MutB confer improved luminescence at high salinity ranges, but decreased performance at more mesohalic salinity (Figs. 2.1D, 2.2B). This result was not unexpected, as other proteins used by halophiles have shown decreased function and folding in mesohalic environments, indicating a dependency on halophilic conditions to ensure proper protein folding and function (Siglioccolo et al., 2011; Tadeo et al., 2009).

2.4.3 Tryptophan Specificity for P_{tnaA} Induction and Half-life Determination

The tryptophan-inducible promoter, P_{tnaA}, found in pTA962 is a previously characterized endogenous promoter with swift and vigorous induction (Allers et al., 2010; Large et al., 2007). The addition of 500 µM tryptophan is sufficient for induction, and expression of the luciferase is readily observable after several minutes through imaging and light capture devices (Fig. 2.5). The fidelity of the induction system was

examined to determine the specificity of the promoter to tryptophan in comparison to other amino acids and a tryptophan analogue, both on 14.4% salt w/v Hv-Minimal plates containing beetle luciferin (Fig. 2.9), and in batch cultures of 14.4% salt w/v Hv-Minimal broth containing beetle luciferin (Fig. 2.3). As expected, induction occurred only on plates and in broth cultures with the addition of tryptophan, with little or no change in induction when other amino acids were added to media. Differences in polarity and alkalinity of the amino acids' side chains had no effect on induction, and the tryptophan analogue 5-methyl-DL-tryptophan did not induce a prolonged increase in luminescence. The short increase in luminescence seen during initial treatment with 5-methyl-DL-tryptophan in batch culture was most likely due to the accidental introduction of extra O_2 upon addition of the treatments, as the bioluminescent reaction is O_2 dependent (Hastings et al., 1953; Moriyama et al., 2008). The induction period varied from five to twenty minutes, similar to induction times seen by others (Large et al., 2007). The half-life of CBG MutB was determined (Fig. 2.4) by inducing CBG MutB expression with tryptophan, then repressing expression with anisomycin, an inhibitor of protein translation in haloarchaea (Torreblanca et al., 1986). After luminescence in the non-repressed sample stabilized, the luminescence decay rate was determined for the repressed sample which revealed the half-life of CBG MutB was roughly 3.5-4 hours, similar to that of click beetle luciferase expressed in yeast (Krishnamoorthy and Robertson, 2015).

2.4.4 Effects of Point Mutations in the Mutant Luciferases

The CBG MutB luciferase contained eleven mutations from the original CBG99 luciferase, which are identified in Table III and Fig. 2.10. Three of the mutations were silent and were not expected to influence luminescence, while the other eight mutations resulted in changes to hydrophobicity, polarity, or both. Mutations 7 and 8 were separated by eight nucleotides, and as a consequence we did not measure the individual influence these mutations incurred without the other. However, mutation 7 was silent, and it was assumed any measured difference in signal intensity caused by the two mutations was primarily due to mutation 8. Mutations causing a decrease in hydrophobicity were expected to increase luminescence, as others have found increased hydrophobic interactions decreases protein function in halophiles that use the "salt-in" method of osmotic pressure maintenance (Siglioccolo et al., 2011; Reed et al., 2013). Most halophilic organisms also display an increase in acidic amino acid residues, therefore mutations resulting in substitution of acidic amino acids or a decrease of basic amino acids were expected to increase signal intensity as well, though more recent evidence suggests halophilic protein stability may be independent of protein charge, and instead more reliant on decreased R group size and protein surface area (Britton et al., 1998; Paul et al., 2008; Tadeo et al., 2009).

To see whether any single mutation contributed to (or detracted from) the luminescence differences between CBG MutB and CBG99 in *H. volcanii*, we completed ten separate iterations of site-directed mutagenesis with CBG99 for each mutation present in CBG MutB (except for mutations 7 and 8 which were very close together so were combined into the same mutant luciferase). These ten CBG99 to CBG MutB convertants

were then patched on 18% salt w/v Hv-YPC plates containing beetle luciferin and imaged for analysis. As can be seen from the quantification in Fig. 2.5A, no single mutation resulted in luminescence significantly increasing beyond that of the unaltered CBG99 under salty conditions. Site directed mutagenesis was also conducted on CBG MutB, where we reverted each mutation (other than mutations 7 and 8 separately) from the CBG MutB nucleotide back to the CBG99 nucleotide. This left each revertant luciferase with ten of the mutations (nine for mutations 7 and 8) similar to CBG MutB, and one (two for mutations 7 and 8) as CBG99. We plated the CBG MutB-to-CBG99 revertants on 18% salt w/v Hv-YPC plates containing beetle luciferin and imaged for analysis (Fig. 2.5B). Results indicated an effect wherein all iterations of revertants displayed a luminescence signal brighter than CBG99 and similar to CBG MutB. This is unsurprising for all revertants , as no individual point mutation appeared solely responsible for increasing luminescence, indicating some number of the other point mutations when combined yield a luciferase significantly brighter than CBG99.

With over three hundred combinations of the eight non-silent mutations in CBG99, it was unfeasible to explore all combinations of mutations to determine which were responsible for the increase in luminescence at high salt conditions. However, some insight can be gained from examining the mutations in the intermediately developed luciferase (MutA) which was the immediate predecessor of MutB and had undergone two iterations of laboratory-directed evolution instead of three. The seven mutations present in CBG MutA contribute to luminescence that is brighter than CBG99 but not as bright as CBG MutB (Fig 2.6 and 2.8). As shown in Table III and Fig. 2.10, point mutations 1, 4, 7, and 10 do not occur in CBG MutA. To explore the luminescence contribution of these

mutations working together, we created a mutant CBG99 that contained mutations 1, 4, and 10 of CBG MutB (which we called convertant 1,4,10). We assumed silent mutation 7 was inconsequential and was therefore not included. As Fig. 2.6 shows, convertant 1,4,10 was brighter than CBG99, but not as bright as CBG MutB (similar to CBG MutA). This indicates mutations 1, 4, and 10 are necessary for CBG MutB's intensity but not sufficient to produce its intensity without the other five non-silent mutations in CBG MutA. Similarly, the five non-silent mutations of CBG MutA are necessary for CBG MutB's intensity but are not sufficient to produce its intensity without at least two mutations of 1, 4, and 10.

2.5 CONCLUSIONS

This study has identified that by modifying a preexisting click beetle luciferase unsuitable for studies in high salt concentrations, we improved luminescence in the halophilic archaeon, *Haloferax volcanii*. We have demonstrated an application for the luciferase as a reporter system for studying gene regulation using patches grown on solid medium and liquid batch cultures. This luciferase could be used for other halophilic archaea, such as *Halobacterium salinarum,* that also use a "salt-in" method of maintaining osmotic pressure. This system is ideal for studying growth and regulation behavior when innocuous, real-time monitoring is most suitable for experiments.

We characterized the effects of salinity on the luciferase, which indicated our mutant performed similarly to other halophilic proteins; function (luminescence) was high at halophilic salinity, and low at mesohalic salinity. In addition, we characterized the changes each mutation made to the luciferase and discovered neither a single point

mutation capable of improving luminescence to a level comparable to our mutant, nor a point mutation capable of diminishing luminescence below the preexisting luciferase level, and that some number of mutations display signal intensity similar to both the mutant. Codon optimization was not found to be an effective measure of increasing luminescence beyond that of laboratory-directed evolution, i.e. introducing intentional silent mutations for codon preference yielded luminescence similar to amino acid changes made by chance. This information may be useful in determining which methodology to use for a specific protein, especially if the protein active site responds poorly to mutations. These findings have the potential to further our knowledge regarding how individual mutations affect the shape, surface, and hydrophobicity of halophilic proteins and how modifications to mesohalic proteins can increase their function in a halophilic environment. By utilizing this knowledge, changes in current production procedures of biotechnologically relevant products could be increased, or new procedures for halophilic compounds could be created.

Further study could elucidate the exact mutations responsible for the increase in luminescence of our mutant, though a large number of mutant permutations exist. While structural models and information on protein folding would provide the most information on surface interactions of the protein, given the number of mutations that exist, it is more practical to first compare mutation combinations and identify the brightest mutants before beginning work on structure. The introduction of these mutations to a naïve system, either another luciferase with similar genetic sequence or a different luciferase where similar amino acid substitutions are made, could also further validate the findings of this study.

2.6 ACKNOWLEDGEMENTS

This research was initiated with funding from the NIH/NIGMS (R37 GM067152 and R01 GM107434 to CHJ) and completed with funding from MTSU Center for Molecular Biosciences (to JBR).

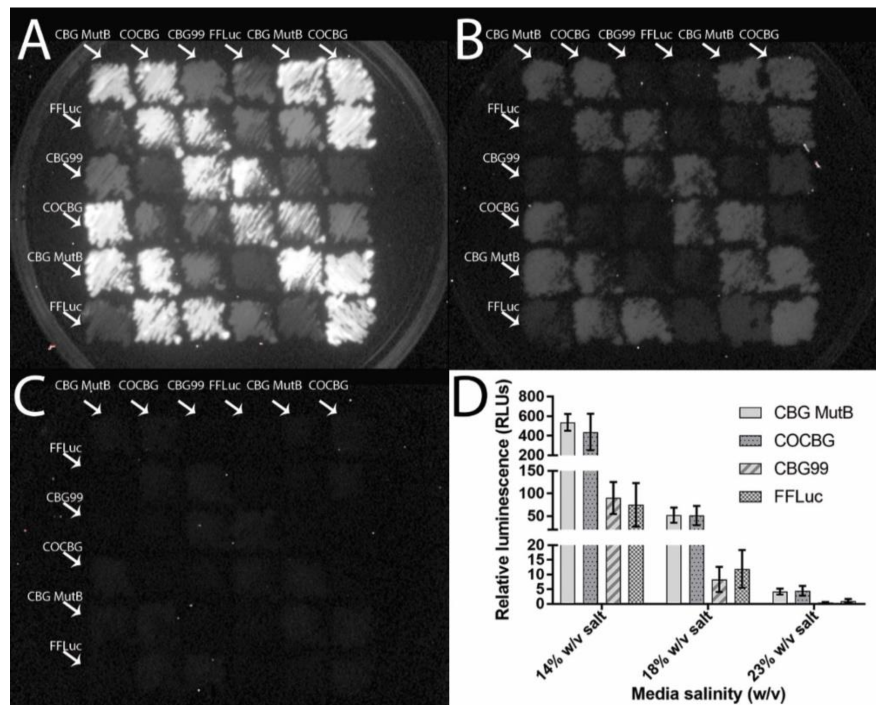


Figure 2.1. Analysis of luciferases in *H. volcanii* strain H1209 transformed with pTA962 plasmids containing various luciferases (CBG99, COCBG, CBG MutB, and FFLuc) driven by P_{tnaA} patched diagonally onto 14.4% salt w/v Hv-YPC solid medium. The plate was replica-plated onto Hv-YPC solid medium containing (A) 14.4% salt, (B) 18% salt, and (C) 23% salt. All plates were imaged (5 min exposure) with a CCD camera after 48 h growth. The brightness and contrast values were uniformly set across all images. The luciferase expressed is indicated by the name and accompanying arrow, which also indicates the direction of the diagonal row of the same luciferase. (D) Luminescence intensity of each luciferase was recorded and quantified using ImageJ software. Signal background was averaged from eight locations on each plate were *H. volcanii* was not patched and subtracted from all measurements from that plate. All bars show means ± standard deviation.

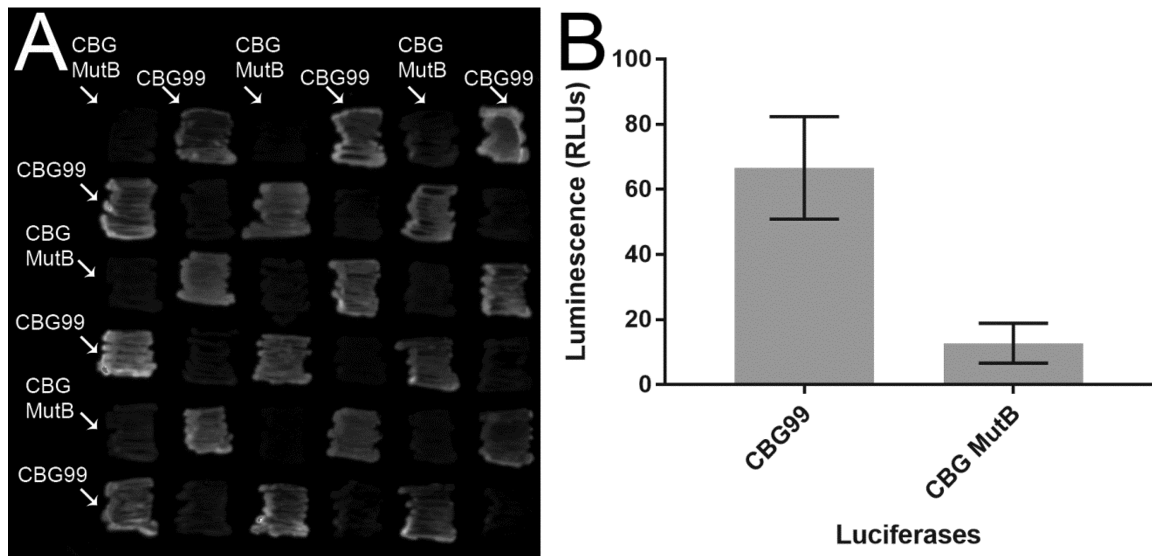


Figure 2.2. Comparison of CBG99 and CBG MutB luciferases in salt-tolerant *E. coli*. (A) *E. coli* strain ALS1187 transformed with pTrc99A plasmid containing either CBR99 or CBG MutB regulated by an IPTG-inducible promoter were patch onto 1% NaCl by volume LB-Miller solid media and imaged (15 sec exposure) after 24 h growth with a CCD camera. (B) Analysis of CBG99 and CBG MutB luminescence at low salinity. Background signal was averaged from 10 locations on the plate containing no patches and was subtracted from all luminescence measurements. Bars show means ± standard deviation.

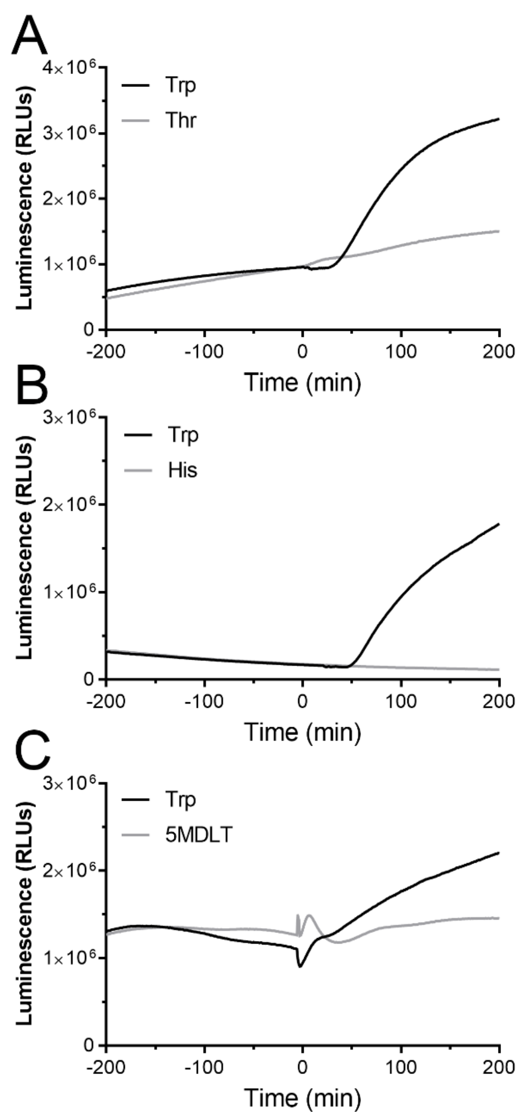
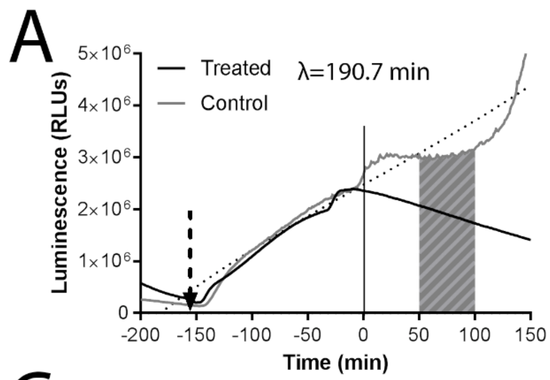
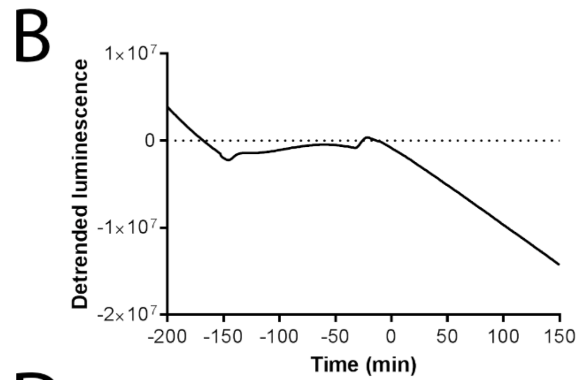
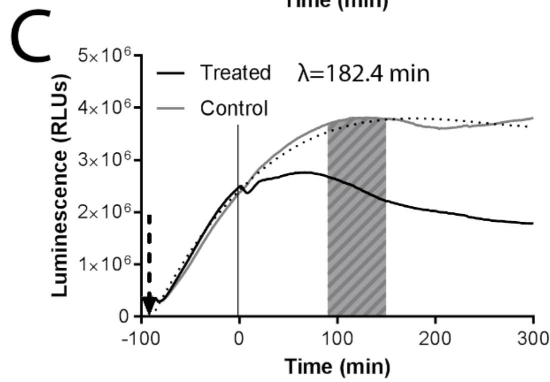
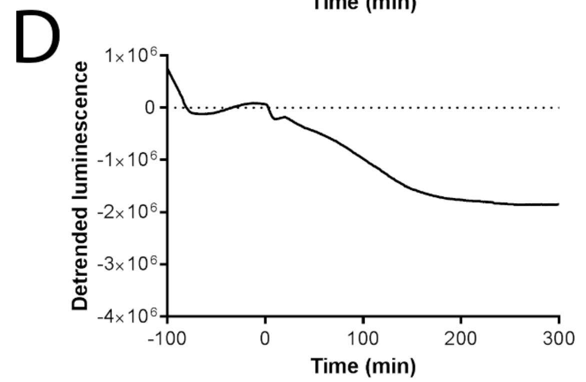


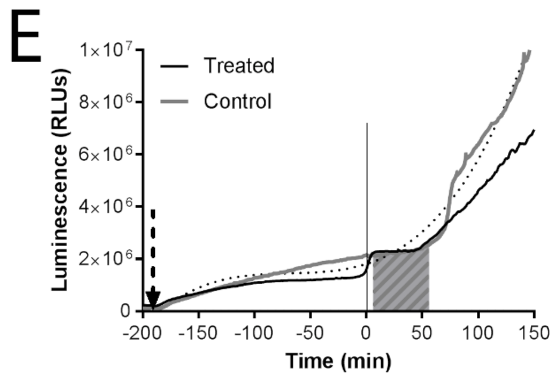
Figure 2.3. Utilization of CBG MutB luciferase reporter in determining tryptophan specificity of PtnaA promoter in *H. volcanii* strain Hv1209. Luminescence was continuously recorded every second for 400 min from two 8 mL 14.4% salt w/v Hv-Minimal media liquid cultures. After 200 min (time point 0), 100 µL of a 50 mM solution of tryptophan (black trace) was added to one tube, and the remaining tube received 100 µL of a 50 mM solution (gray trace) of either (A) threonine, (B) histidine, or (C) 5-methyl-DL-tryptophan, a tryptophan analogue. Tryptophan was the only amino acid which induced immediate and drastic expression of the CBG MutB luciferase.


A
Treated
Control
λ=190.7 min
Luminescence (RLUs)
Time (min)


B
Detrended luminescence
Time (min)


C
Treated
Control
λ=182.4 min
Luminescence (RLUs)
Time (min)


D
Detrended luminescence
Time (min)


E
Treated
Control
Luminescence (RLUs)
Time (min)

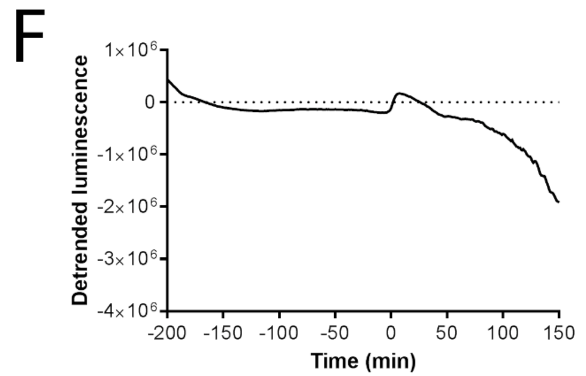

F
Detrended luminescence
Time (min)

Figure 2.4. Half-life analysis of CBG-MutB in Hv-Minimal liquid media at three salt concentrations. *H. volcanii* luminescence half-life was measured after inducing expression of CBG MutB with tryptophan (black dashed line with arrow), then repressed with the introduction of anisomycin (solid vertical line). For the treated culture, luminescence measured in the shaded area (corresponding to the time when the untreated culture had little to no change in luminescence) displayed a luminescence decay with a T½ of 3.5h - 4.5h in (A) 14.4% salt w/v and (C) 18% salt w/v media. (E) At 23.3% salt w/v, anisomycin treatment slowed expression, but no decay was measured. (B, D, and F) Culture growth influence was detrended from the data by subtraction of the polynomial from the luminescence of the culture for each salinity concentration, (B) 14.4%, (D) 18%, and (F) 23.3%.

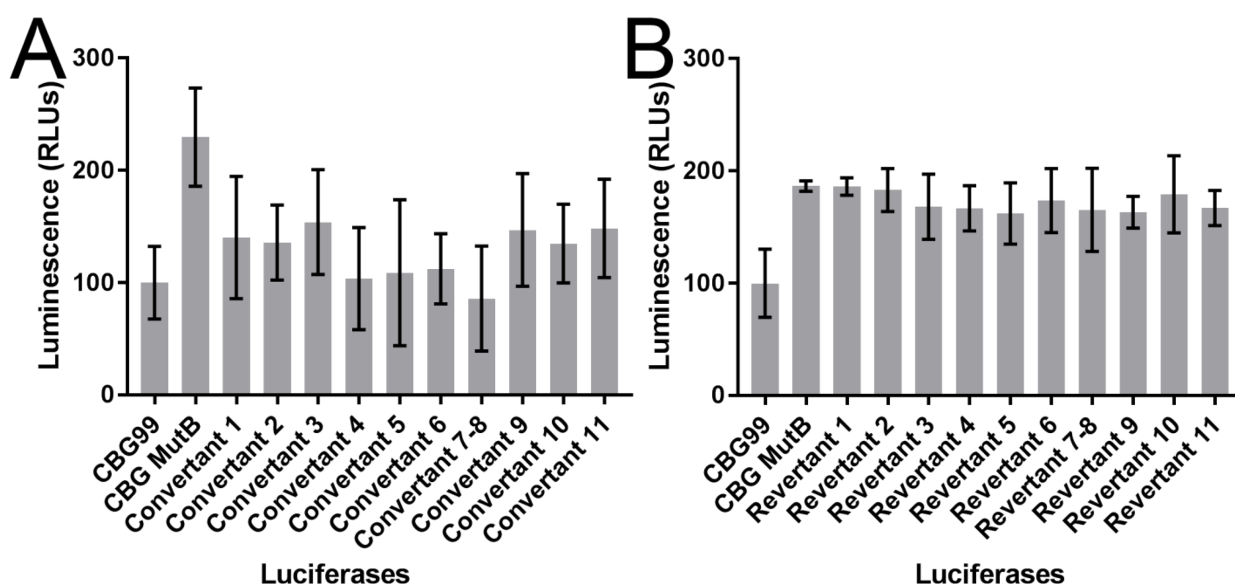


Figure 2.5. Effect of single point mutations on relative luminescence between CBG99 and CBG MutB. (A) Site-directed mutagenesis of CBG99 was conducted in order to induce one of the eleven point mutations (labeled a convertant in this study) seen in CBG MutB. Mutations 7 and 8 sit eight bases apart, and are too close to one another to separate a single mutagenesis event of either point. No single mutation from CBG99 to CBG MutB results in a luciferase significantly brighter than CBG99, though some approach CBG MutB in intensity. (B) Site directed mutagenesis of CBG MutB was also conducted to revert one of the eleven mutations back to the original nucleotide present in CBG99. No single mutation reverting from CBG MutB to CBG99 caused a decrease in luminescence, indicating while mutation point 1 can increase luminescent intensity from CBG99, its absence does not counteract the cumulative effect of all ten other mutations. Similarly, removing mutation point 5 does not increase relative luminescent intensity. Bars show means ± standard deviation.

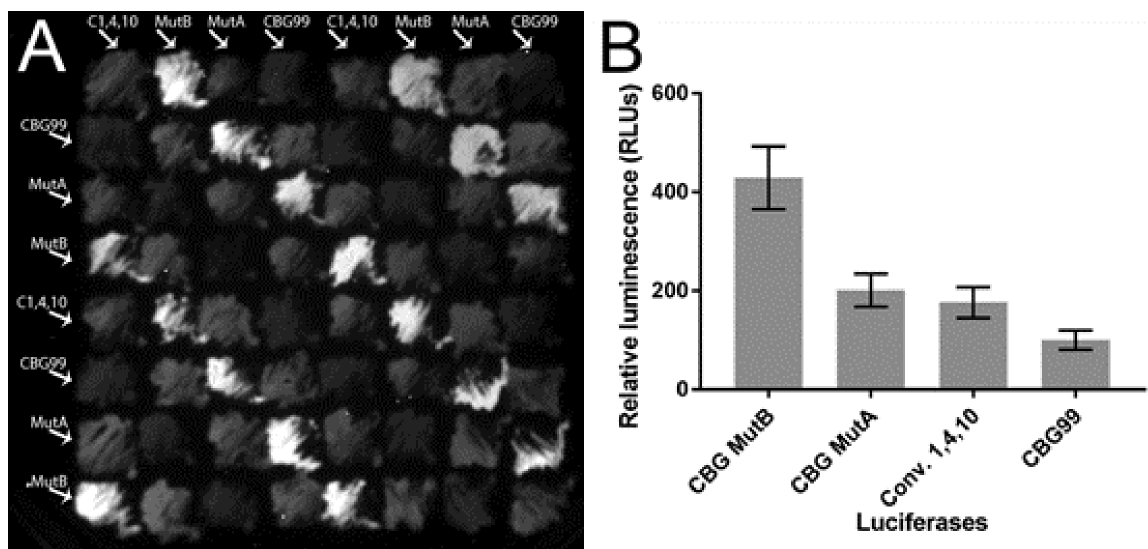


Figure 2.6. Combined mutations 1, 4, and 10 increase luminescence similar to CBG MutA but not to the amount seen in CBG MutB. (A) *H. volcanii* transformed with various luciferases driven by PtnaA were patched diagonally onto 18% salt w/v Hv-YPC solid medium and imaged (5 min exposure) after 48 h growth. The luciferase within each patch is indicated by the name and accompanying arrow, which also indicates the direction of the diagonal row of the same luciferase. CBG MutA is a luciferase similar to CBG99 except that it contains the non-silent mutations 2, 3, 5, 8, 11 of CBG MutB. Conv. 1,4,10 is a luciferase similar to CBG99 except that it contains mutations 1, 4, and 10 of CBG MutB (B) The luminescence intensity of each luciferase was recorded and quantified using ImageJ software. Signal background was averaged from 8 locations on the plate were *H. volcanii* was not patched and subtracted from all luminescence measurements. All bars show means ± standard deviation.

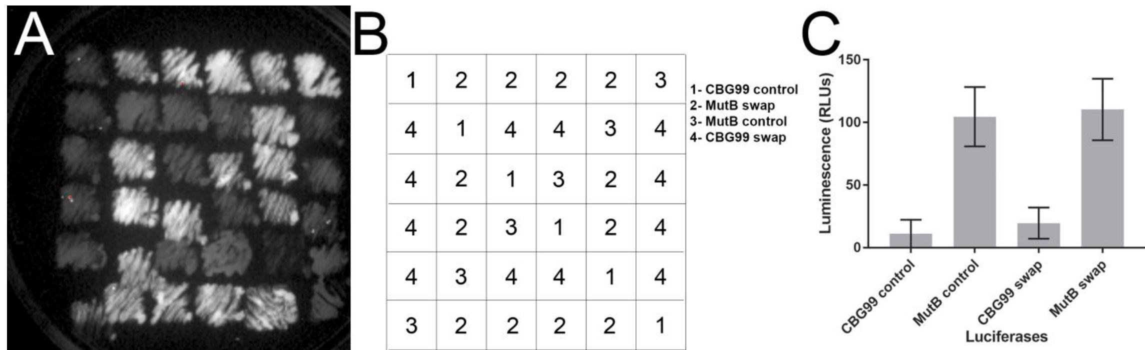


Figure 2.7. Comparison of CBG99 and CBG MutB coding sequences in pTA962 displaying no significant difference when swapping vectors. The two luciferase inserts (labeled here as MutB control and CBG99 control) were digested from original parent vector and swapped into the other reciprocating parent vector (labeled here as MutB swap and CBG99 swap) to demonstrate luminescence intensity is a product of mutations made in the luciferase gene. (A) A 14.4% salt w/v Hv-YPC patch plate used for luminescence quantification. (B) A map denoting the identity of each patch in A. (C) Quantification of insert/vector swap from patch plate displaying no significant difference between either CBG MutB inset/vector combination and CBG99 insert/vector combination. Signal background was averaged from eight locations on the plate were *H. volcanii* was not patched and subtracted from all bioluminescence measurements. All bars show means ± standard deviation.

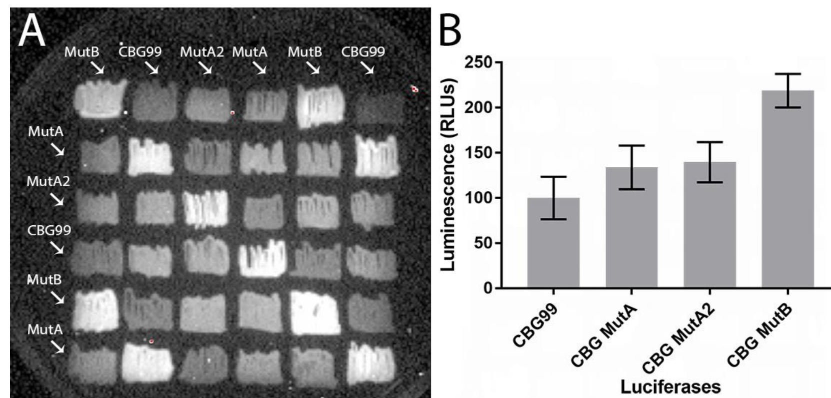


Figure 2.8. Luminescence of two intermediately developed luciferases in *H. volcanii*. (A) *H. volcanii* strain H1209 transformed with pTA962 plasmids containing various luciferase genes at different stages of development (CBG99, intermediates CBG MutA and CBG MutA2, and CBG MutB,) driven by P_{tnaA} were patched diagonally onto 14.4% salt w/v Hv-YPC solid medium and imaged (5 min exposure) after 48 h growth. The luciferase expressed is indicated by the name and accompanying arrow, which also indicates the direction of the diagonal row of the same luciferase. (B) The luminescence intensity of each luciferase recorded and quantified using ImageJ software. Signal background was averaged from eight locations on the plate were *H. volcanii* was not patched and subtracted from luminescence measurements. All bars show means ± standard deviation.

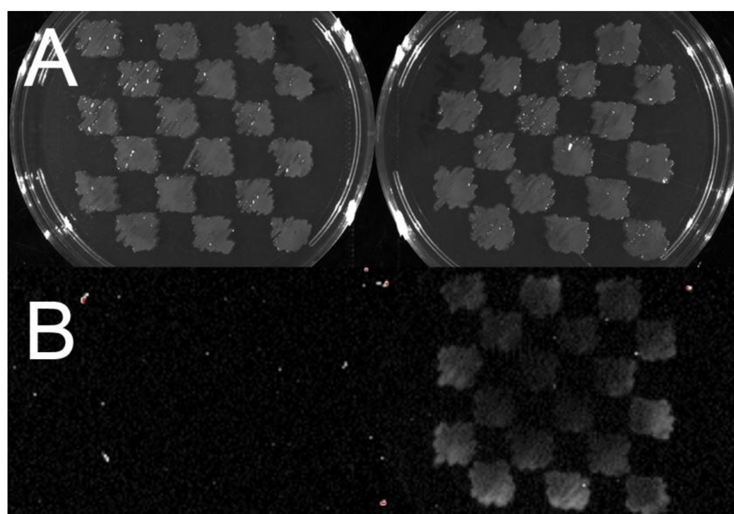


Figure 2.9. Expression of luciferase under control of P_{tnA}, a tryptophan inducible promoter, in *H. volcanii* on solid media. (A) Bright-field image of two 14.4% salt w/v Hv-Minimal solid media plates impregnated with 25 μL of 100 μM beetle luciferin and each patched with *H. volcanii* transformed with plasmid pTA962-CBG MutB. The left plate received no tryptophan, while the right plate received a total treatment of 100μL of 50 mM tryptophan, evenly placed in the center of empty patch areas. (B) Image of same plates (5 min exposure) after addition of tryptophan to the right plate captured with a Bio-Rad ChemiDoc MP.

```
                                  1
prot   M  V  K  R  E  K  N  V  I  Y  G  P  E  P  L  H  P  L  E  D  L  T  A  G  E  M  L  F  R  A  L  R  K
                                  H
CBG99 ATGGTGAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAG
MutA  ATGGTGAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAG
MutB  ATGGTGAAGCGTGAGAAAAATGTCATCCATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAG
COCBG ATGGTCAAGCGCGAGAAGAACGTCATCCACGGCCCCGAGCCCCTCCACCCCCTCGAGGACCTCACCGCCGGCGAGATGCTCTTCCGCGCCCTCCGCAAG

                                  2
prot   H  S  H  L  P  Q  A  L  V  D  V  V  G  D  E  S  L  S  Y  K  E  F  F  E  A  T  V  L  L  A  Q  S  L
                                     E
CBG99 CACTCTCATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATCTTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTC
MutA  CACTCTCATTTGCCTCAAGCCTTGGTCGATGAGGTCGGCGATGAATCTTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTC
MutB  CACTCTCATTTGCCTCAAGCCTTGGTCGATGAGGTCGGCGATGAATCTTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTC
COCBG CACTCCCACCTCCCCCAGGCCCTCGTCGACGAGGTCGGCGACGAATCCCTCAGCTACAAGGAGTTCTTCGAGGCCACCGTCCTCCTGGCCCAGTCCCTC

                     3                                      4
prot   H  N  C  G  Y  K  M  N  D  V  V  S  I  C  A  E  N  N  T  R  F  F  I  P  V  I  A  A  W  Y  I  G  M
                     K                                      T
CBG99 CACAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATG
MutA  CACAATTGTGGCTACAAGAAGAACGACGTCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATG
MutB  CACAATTGTGGCTACAAGAAGAACGACGTCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCACTCCAGTCATCGCCGCATGGTATATCGGTATG
COCBG CACAACTGCGGCTACAAGAAGAACGACGTCGTCAGCATCTGCGCCGAGAACAACACCCGCTTCTTCACCCCGGTCATCGCCGCCTGGTACATCGGCATG

prot   I  V  A  P  V  N  E  S  Y  I  P  D  E  L  C  K  V  M  G  I  S  K  P  Q  I  V  F  T  T  K  N  I  L
CBG99 ATCGTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTG
MutA  ATCGTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTG
MutB  ATCGTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTG
COCBG ATCGTCGCCCCCGTCAACGAGAGCTACATCCCCGACGAGCTCTGCAAGGTCATGGGCATCTCCAAGCCCCAGATCGTCTTCACCACCAAGAACATCCTG

prot   N  K  V  L  E  V  Q  S  R  T  N  F  I  K  R  I  I  I  L  D  T  V  E  N  I  H  G  C  E  S  L  P  N
CBG99 AACAAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAAT
MutA  AACAAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAAT
MutB  AACAAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAAT
COCBG AACAAGGTCCTCGAGGTCCAGAGCCGCACCAACTTCATCAAGCGCATCATCATCCTCGACACCGTCGAGAACATCCACGGCTGCGAGTCCCTCCCCAAC

                                          5
prot   F  I  S  R  Y  S  D  G  N  I  A  N  F  K  P  L  H  F  D  P  V  E  Q  V  A  A  I  L  C  S  S  G  T
                                          R
CBG99 TTCATCTCTCGCTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCACTTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACT
MutA  TTCATCTCTCGCTATTCAGACGGCAACATCGCAAACTTTAGACCACTCCACTTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACT
MutB  TTCATCTCTCGCTATTCAGACGGCAACATCGCAAACTTTAGACCACTCCACTTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACT
COCBG TTCATCTCCCGCTACTCCGACGGCAACATCGCCAACTTCCGCCCCCTCCACTTCGACCCCGTCGAGCAGGTCGCCGCCATCCTGTGCAGCAGCGGCACC

                                                                                     6
prot   T  G  L  P  K  G  V  M  Q  T  H  Q  N  I  C  V  R  L  I  H  A  L  D  P  R  V  G  T  Q  L  I  P  G
CBG99 ACTGGACTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCACGCGTGGGCACTCAGCTGATTCCTGGT
MutA  ACTGGACTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCACGCGTGGGAACTCAGCTGATTCCTGGT
MutB  ACTGGACTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCACGCGTGGGAACTCAGCTGATTCCTGGT
COCBG ACCGGCCTCCCCAAGGGCGTCATGCAGACCCACCAGAACATCTGCGTCCGCCTCATCCACGCCCTCGACCCCCGCGTGGGCACCCAGCTGATCCCCGGC

prot   V  T  V  L  V  Y  L  P  F  F  H  A  F  G  F  S  I  T  L  G  Y  F  M  V  G  L  R  V  I  M  F  R  R
CBG99 GTCACCGTCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTAGCATTACTTTGGGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCCGT
MutA  GTCACCGTCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTAGCATTACTTTGGGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCCGT
MutB  GTCACCGTCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTAGCATTACTTTGGGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCCGT
COCBG GTCACCGTCCTGGTCTACCTCCCCTTCTTCCACGCCTTCGGCTTCAGCATCACCCTCGGCTACTTCATGGTCGGCCTCCGCGTCATCATGTTCCGCCGC

prot   F  D  Q  E  A  F  L  K  A  I  Q  D  Y  E  V  R  S  V  I  N  V  P  S  V  I  L  F  L  S  K  S  P  L
CBG99 TTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACTC
MutA  TTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACTC
MutB  TTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACTC
COCBG TTCGACCAGGAGGCCTTCCTCAAGGCCATCCAGGACTACGAGGTCCGCTCCGTCATCAACGTCCCCAGCGTCATCCTCTTCCTCTCCAAGAGCCCCCTC
```

```
prot   V  D  K  Y  D  L  S  S  L  R  E  L  C  C  G  A  A  P  L  A  K  E  V  A  E  V  A  A  K  R  L  N  L

CBG99 GTGGACAAGTACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATCTT
MutA  GTGGACAAGTACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATCTT
MutB  GTGGACAAGTACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATCTT
COCBG GTCGACAAGTACGACCTCTCCTCGCTCCGTGAGCTGTGTTGCGGCGCCGCGCCGCTCGCCAAAGAAGTCGCCGAGGTCGCGGCCAAGCGCCTCAACCTC

                                                                                            7        8
prot   P  G  I  R  C  G  F  G  L  T  E  S  T  S  A  N  I  H  S  L  G  D  E  F  K  S  G  S  L  G  R  V  T
                                                                                                     H
CBG99 CCAGGGATTCGTTGTGGCTTCGGCCTCACCGAATCTACCAGCGCTAACATTCACTCTCTCGGGGATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCACT
MutA  CCAGGGATTCGTTGTGGCTTCGGCCTCACCGAATCTACCAGCGCTAACATTCACTCTCTCGGGGATGAGTTTAAGAGCGGCTCTTTGGGCCATGTCACT
MutB  CCAGGGATTCGTTGTGGCTTCGGCCTCACCGAATCTACCAGCGCTAACATTCACTCTCTCGGGGATGAGTTTAAGAGCGGCTCCTTGGGCCATGTCACT
COCBG CCGGGCATCCGCTGCGGCTTCGGCCTCACCGAGTCGACCAGCGCCAACATCCACTCGCTCGGCGACGAGTTCAAGAGCGGTTCGCTGGGTCACGTCACG

                                             9
prot   P  L  M  A  A  K  I  A  D  R  E  T  G  K  A  L  G  P  N  Q  V  G  E  L  C  I  K  G  P  M  V  S  K

CBG99 CCACTCATGGCTGCTAAGATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCGAACCAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCAAG
MutA  CCACTCATGGCTGCTAAGATCGCTGATCGCGAAACTGGTAAGGCCTTGGGCCCGAACCAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCAAG
MutB  CCACTCATGGCTGCTAAGATCGCTGATCGCGAAACTGGTAAGGCCTTGGGCCCGAACCAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCAAG
COCBG CCGCTCATGGCCGCGAAGATCGCCGACCGCGAGACCGGCAAGGCCCTCGGCCCGAACCAGGTCGGCGAGCTGTGCATCAAGGGTCCGATGGTCAGCAAG

                                                                                         10
prot   G  Y  V  N  N  V  E  A  T  K  E  A  I  D  D  D  G  W  L  H  S  G  D  F  G  Y  Y  D  E  D  E  H  F
                                                                                          H
CBG99 GGTTATGTCAATAACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGAGCATTTT
MutA  GGTTATGTCAATAACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGAGCATTTT
MutB  GGTTATGTCAATAACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGACATTACGACGAAGATGAGCATTTT
COCBG GGCTACGTCAACAACGTCGAGGCCACCAAGGAGGCCATCGACGACGACGGCTGGCTCCACTCCGGCGACTTCGGCCACTACGACGAGGACGAGCACTTC

                                                                                               11
prot   Y  V  V  D  R  Y  K  E  L  I  K  Y  K  G  S  Q  V  A  P  A  E  L  E  E  I  L  L  K  N  P  C  I  R
                                                                                                N
CBG99 TACGTCGTGGATCGTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGATTCTGTTGAAAAATCCATGCATTCGC
MutA  TACGTCGTGGATCGTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGATTCTGTTGAATAATCCATGCATTCGC
MutB  TACGTCGTGGATCGTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGATTCTGTTGAATAATCCATGCATTCGC
COCBG TACGTCGTGGACCGCTACAAAGAGCTCATCAAGTACAAGGGCAGCCAGGTCGCCCCGGCCGAGCTCGAGGAGATCCTGCTCAACAACCCCTGCATCCGC

prot   D  V  A  V  V  G  I  P  D  L  E  A  G  E  L  P  S  A  F  V  V  K  Q  P  G  K  E  I  T  A  K  E  V

CBG99 GATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTAAAGAAATTACCGCCAAAGAAGTG
MutA  GATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTAAAGAAATTACCGCCAAAGAAGTG
MutB  GATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTAAAGAAATTACCGCCAAAGAAGTG
COCBG GACGTCGCCGTCGTCGGCATCCCGGACCTGGAGGCCGGCGAGCTGCCGTCGGCCTTCGTCGTCAAGCAGCCGGGCAAGGAGATCACCGCCAAGGAGGTG

prot   Y  D  Y  L  A  E  R  V  S  H  T  K  Y  L  R  G  G  V  R  F  V  D  S  I  P  R  N  V  T  G  K  I  T

CBG99 TATGATTACCTGGCTGAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTCGTAACGTAACAGGCAAAATTACC
MutA  TATGATTACCTGGCTGAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTCGTAACGTAACAGGCAAAATTACC
MutB  TATGATTACCTGGCTGAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTCGTAACGTAACAGGCAAAATTACC
COCBG TACGACTACCTGGCCGAGCGCGTCAGCCACACCAAGTACCTGCGCGGCGGCGTCCGTTTCGTCGACTCCATCCCGCGCAACGTCACCGGCAAGATCACC

prot   R  K  E  L  L  K  Q  L  L  E  K  A  G  G  *

CBG99 CGCAAGGAGCTGTTGAAACAATTGTTGGAGAAGGCCGGCGGTTAG
MutA  CGCAAGGAGCTGTTGAAACAATTGTTGGAGAAGGCCGGCGGTTAG
MutB  CGCAAGGAGCTGTTGAAACAATTGTTGGAGAAGGCCGGCGGTTAG
COCBG CGCAAGGAGCTGCTCAAGCAGCTCCTGGAGAAGGCCGGCGGTTAG
```

Figure 2.10. Sequences of luciferases in this study. Seven items are shown. 1) The lines labeled "prot" show the amino acid sequence of the commercially available CBG99 gene. 2) Above this line appear numbers 1-11 indicating the number designations for the mutations to the CBG99 sequence produced by error-prone PCR. 3) Immediately below the "prot" lines are the amino acid substitutions (if any) that result from the mutations. 4) Lines labeled "CBG99" show the nucleotide sequence for CBG99. 5) Lines labeled MutA show the nucleotide sequence for CBG MutA luciferase generated in this study. Differences from the CBG99 parent are highlighted in black. 6) Lines labeled MutB show the nucleotide sequence for CBG MutB from this study. Differences from CBG99 are highlighted in black. 7) Lines labeled COCBG show the sequence for the codon-optimized version of CBG MutB. Differences from CBG99 are highlighted in black.

Table I. Plasmids used.

Plasmid	Relevant Properties	Source or reference
pCBG99-basic	CBG99 CDS	Promega
pGL4.11[luc2P]	luc2P CDS	Promega
pTA962	Overexpression vector with pyrE2 and hdrB markers and pHV2 origin, insertion of t.L11e terminator, PtnaA promoter and t.Syn terminator	Allers et al. 2010
pTA962-CBG99	pTA962 with CBG99 insertion at NdeI and BamHI sites	This study
pTA962-CBG MutA	pTA962 with CBG MutA insertion at NdeI and BamHI sites	This study
pTA962-CBG MutB	pTA962 with CBG MutB insertion at NdeI and BamHI sites	This study
pTA962-FFLuc	pTA962 with luc2P insertion at NdeI and BamHI sites	This study
pTA962-COCBG	pTA962 with COCBG insertion at NdeI and BamHI sites	This study
pTRC99a	Bacterial expression vector with inducible lacI promoter	Amann et al. 1983
pTRC99a-CBG99	pTRC99a with CBG99 insertion at EcoRI and BamHI sites	This study
pTRC99a-CBG MutB	pTRC99a with CBG Mut2B insertion at EcoRI and BamHI sites	This study
pUC19	Ampicillin resistance	Norrander et al. 1983
pUC19-CBG99	CBG99 insertion at EcoRI and BamHI sites	This study
pUC19-CBG MutB	CBG Mut2B insertion at EcoRI and BamHI sites	This study

Table II. Oligonucleotides used.

	Primer	Sequence (5'-3')	Notes
1	CBG srt	TTCGAACATATGGTGAAGCGTGAGAAAAATG	5' NdeI
2	CBG end	AACTATGGATCCTAACCGCCGGCCTTCTC	3'Bam HI
3	FFLuc srt	TTCGAACATATGGAAGATGCCAAAAACATTAAG	5' NdeI
4	FFLuc end	AACTATGGATCCACTCTAGAGTCGCGGCCTTAG	3'Bam HI
5	Convert1 F	GCGTGAGAAAAATGTCATCCATGGCCCTGAGCC	
6	Convert1 R	GGCTCAGGGCCATGGATGACATTTTTCTCACGC	
7	Convert2 F	GCCTTGGTCGATGAGGTCGGCGATGAATC	
8	Convert2 R	GATTCATCGCCGACCTCATCGACCAAGGC	
9	Convert3 F	CCACAATTGTGGCTACAAGAAGAACGACGTCG	
10	Convert3 R	CGACGTCGTTCTTCTTGTAGCCACAATTGTGG	
11	Convert4 F	CCCGTTTCTTCACTCCAGTCATCGCCGCATGG	
12	Convert4 R	CCATGCGGCGATGACTGGAGTGAAGAAACGGG	
13	Convert5 F	CGGCAACATCGCAAACTTTAGACCACTCCACTTCGACCC	
14	Convert5 R	GGGTCGAAGTGGAGTGGTCTAAAGTTTGCGATGTTGCCG	
15	Convert6 F	CGATCCACGCGTGGGAACTCAGCTGATTCC	
16	Convert6 R	GGAATCAGCTGAGTTCCCACGCGTGGATCG	
17	Convert 7-8 F	GATGAGTTTAAGAGCGGCTCGTTGGGCCATGTCACTCCACTCATG	
18	Convert 7-8 R	CATGAGTGGAGTGACATGGCCCAACGAGCCGCTCTTAAACTCTATC	
19	Convert9 F	CGCGAAACTGGTAAGGCCTTGGGCCCGAACC	
20	Convert9 R	GGTTCGGCCCAAGGCCTTACCAGTTTCGCG	
21	Convert10 F	GGTTGCATTCTGGTGATTTTGGACATTACGACGAAGATG	
22	Convert10 R	CATCTTCGTCGTAATGTCCAAAATCACCAGAATGCAACC	
23	Convert11 F	GGAGGAGATTCTGTTGAATAATCCATGCATTCGCG	
24	Convert11 R	CGCGAATGCATGGATTATTCAACAGAATCTCCTCC	
25	Revert1 F	GCGTGAGAAAAATGTCATCTATGGCCCTGAGCC	

Table II (cont.)

	Primer	Sequence (5'-3')	Notes
26	Revert 1 R	GGCTCAGGGCCATAGATGACATTTTTCTCACGC	
27	Revert 2 F	GCCTTGGTCGATGTGGTCGGCGATGAATC	
28	Revert 2 R	GATTCATCGCCGACCACATCGACCAAGGC	
29	Revert 3 F	CCACAATTGTGGCTACAAGATGACGACGTCG	
30	Revert 3 R	CGACGTCGTTCATCTTGTAGCCACAATTGTGG	
31	Revert 4 F	CCCGTTTCTTCATTCCAGTCATCGCCGCATGG	
32	Revert 4 R	CCATGCGGCGATGACTGGAATGAAGAAACGGG	
33	Revert 5 F	CGGCAACATCGCAAACTTTAAACCACTCCACTTC GACCC	
34	Revert 5 R	GGGTCGAAGTGGAGTGGTTTAAAGTTGCGATGTT GCCG	
35	Revert 6 F	CGATCCACGCGTGGGCACTCAGCTGATTCC	
36	Revert 6 R	GGAATCAGCTGAGTGCCCACGCGTGGATCG	
37	Revert 7-8 F	GATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCA CTCCACTCATG	
38	Revert 7-8R	CATGAGTGGAGTGACATGGCCCAACGAGCCGCTC TTAAACTCTATC	
39	Revert 9 F	CGCGAAACTGGTAAGGCTTTGGGCCCGAACC	
40	Revert 9 R	GGTTCGGGCCCAAAGCCTTACCAGTTTCGCG	
41	Revert 10 F	GGTTGCATTCTGGTGATTTTGGATATTACGACGA AGATG	
42	Revert 10 R	CATCTTCGTCGTAATATCCAAAATCACCAGAATG CAACC	
43	Revert 11 F	GGAGGAGATTCTGTTGAAAAATCCATGCATTCGC G	
44	Revert 11 R	CGCGAATGCATGGATTTTTCAACAGAATCTCCTC C	
45	CBG EcoRI	AACTATGAATTCTAACCGCCGGCCTTCTC	5' EcoRI

Table III. Luciferase mutations and amino acid changes

Mutation Number	Nucleotide Position	CBG99	CBG MutB	CBG MutA	Consequence
1	28	T	C		Y→H
2	131	T	A	A	V→E
3	218	T	A	A	M→K
4	266	T	C		I→T
5	536	A	G	G	K→R
6	675	C	A	A	silent
7	1074	T	G		silent
8	1082	G	A	A	R→H
9	1134	T	C	C	silent
10	1264	T	C		Y→H
11	1371	A	T	T	K→N

CHAPTER III

REAL TIME MONITORING OF THE TYPE I-B CRISPR/CAS SYSTEM IN *HALOFERAX VOLCANII* USING A LUCIFERASE REPORTER

3.1 INTRODUCTION

The clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR-associated (Cas) mechanisms are endogenous genetic elements which act as a prokaryotic immune system based on previous exposure to foreign genetic elements. Almost all archaeal species and half of bacteria contain a CRISPR/Cas system (Maier et. al 2012). Effectively, the CRISPR loci act as a genetic record of previous exposure to pathogenic genetic elements, as they contain sequences of DNA originally copied from foreign sources, or "spacer DNA", separated by short repeated DNA segments. This immune response is a three step process: invader DNA/RNA must be acquired (adaptation), CRISPR RNAs (crRNAs) must be expressed (biogenesis), and degradation of invader DNA/RNA must occur (interference). Three types of CRISPR/Cas systems exist (type I-III), due to differences in invader DNA integration and Cas proteins; these three categories can be further divided into a total of ten subtypes (Makarova 2012). Since *Haloferax volcanii* utilizies type I-B, this paper will focus primarily on the I-B CRISPR/Cas system and related elements in *H. volcanii*, described previously for use with molecular genetic techniques (Gophna et. al 2017).

Within *H. volcanii* there exist eight Cas (Cas1-Cas8b) proteins and three DNA loci, with one contained on the main chromosome, and the remaining two flanking a polycistronic operon of the Cas genes on the pHV4 chromosomal plasmid (Fig. 3.1)

(Fischer et al. 2012). The repeat sequences between spacers at all three loci are 30 nucleotides long, and differ only by one base near the end of the spacer sequence (Maier et al. 2013).

Adaptation to invader DNA is completed by acquisition and integration of a foreign DNA sequence known as a protospacer into a CRISPR locus, where the protospacer is then termed a spacer. New spacers are added to a CRISPR locus preferentially, typically adjacent to the leading 5' spacer sequence. Type I-B systems require the presence of invader DNA containing a specific short sequence motif, known as a protospacer adjacent motif (PAM), directly upstream of the protospacer (Maier et al. 2015). This PAM sequence is used to identify and select protospacer integration sites, as well as direct the interference step of the Cas system. The endogenous CRISPR/Cas system contains no PAMs, which is required in conjunction with recognition of the protospacer sequence in order for DNA degradation to occur; this prevents autoimmunity within the organism. There are two adaptation pathways in type I systems, the naïve and priming adaptation pathways (Fineran and Charpentier 2012). The naïve pathway utilizes only Cas1 and Cas2, and is the less common pathway. This pathway acts as the primary method of spacer acquisition to completely foreign DNA, as Cas1 and Cas2 proteins can identify the PAMs required to initiate adaptation of foreign DNA (Wang et al. 2015). The priming pathway, of which *H. volcanii* preferentially uses, requires Cas1 and Cas2, but also involves Cas3, the complex known as the Cascade (CRISPR-associated complex for antiviral defense) complex, and a pre-existing spacer that fully or partially matches the invader DNA (Li et al. 2017; Sternberg et al. 2016). This pathway is based upon previous exposure to foreign DNA, and provides the organism with additional spacers for more

reliable recognition of invader DNA in the future. Whereas protospacer sequences are incorporated into genomic DNA as spacers in both methods, the Cas1 and Cas2 proteins are not as strict in PAM conservation, whereas the crRNAs in priming adaptation are much stricter in conservation of these motifs in invader DNA recognition (Savitskaya 2013).

The second stage of CRISPR activity covers the biogenesis of the CRISPR-derived RNAs (crRNAs); the CRISPR loci are transcribed into long precursor sequences starting from a promoter located within the leader repeat sequence, and Cas6 cleaves these into mature crRNAs, each with a spacer sequence unique for a particular invader (Charpentier et al. 2015). Once the loci RNA cleaving is completed by Cas6, the protein remains tightly bound to the 3' end of the crRNA structure, acting as an anchor for other Cas proteins to bind to once the crRNA recognizes a foreign protospacer; the 5' end is bound by Cas5 in type I systems (Weidenheft et al. 2011; Li et al. 2014). Apart from Haloarchaea, the Cas6 protein of *H. volcanii* shows very little sequence similarity to other archaea or bacteria (Brendel et al. 2014). Deletion of the Cas6 protein in *H. volcanii* halts the formation of crRNAs, but crRNA biogenesis also hinges upon protection using Cas5 and Cas7 (Brendel et al. 2014). These three proteins form a stable Cascade complex capable of protecting the spacer before and during the interference step; Cas8b is also associated with this complex as a necessary component of interference, though unnecessary for biogenesis and stability (Brendel et al. 2014; Maier et al. 2015). A key difference between Haloarchaeal type I-B systems and other species is the addition of an extra Cas7 protein to the complex, speculated to compensate for the extra length in spacer sequence of Haloarchaea in comparison to other type I-B systems (Maier et al. 2015).

The final stage of the defense reaction is interference, which can only occur when invader DNA carrying a protospacer sequence enters the cell. During interference, the Cascade-crRNAs complex recognizes a specific seeding sequence of the protospacer located within the invader DNA, forming an R-loop, and then the Cascade-crRNA complex initiates degradation of foreign DNA with Cas3 (Cass et al. 2015). Initially, the DNA seeding event occurs in the portion of the protospacer most proximal to the PAM sequence; here the invader DNA must exactly match the crRNA sequence in order to continue the nucleolytic process (Semenova et al. 2011). After seeding, an R-loop is formed: crRNA forms base pairs with complementary invader DNA and displaces noncomplementary DNA into a loop (Jore et al. 2011). This R-loop causes a conformational change in the Cascade complex, allowing Cas3 to associate with the complex and begin cleavage of invader DNA (Xiao et al. 2017).

The purpose of this experiment was to study the temporal characteristics of the CRISPR/Cas system in detecting foreign genetic elements using a salt-tolerant luciferase. The experiment required the creation of both a *H. volcanii* strain in which genetic elements responsible for interference in the CRISPR/Cas mechanism were regulated by the tryptophan-inducible promoter found in pTA962 and selectable on uracil-deficient plates, and an “invader” plasmid housing a constitutive promoter; thymidine and uracil selectable markers; and a spacer and protospacer adjacent motif (PAM). Once this invader plasmid was transformed into the tryptophan-controlled Cas cassette *H. volcanii* strain, a decay in luminescence was expected after the addition of tryptophan, as the invader plasmid containing the spacer and PAM elements was detected and destroyed by the CRISPR/Cas system. The same invader plasmid without the PAM or protospacer

elements would have acted as a control, as without the invader elements, luminescence should remain stable and detectable over time. While CRISPR/Cas systems exist to integrate luciferase reporters within cell lines, no studies were found utilizing a luciferase reporter to study temporal aspects of the CRISPR/Cas system.

3.2 MATERIALS AND METHODS

3.2.1 Strains, Plasmids, and Primers

The plasmids, primers and strains used for CRISPR/Cas studies or DNA manipulation in *H. volcanii* or *E. coli* are shown in Table IV and Table V, respectively. Initially, *H. volcanii* strain H1209 (Δ*pyrE2*, Δ*hdrB*) was used for all studies, moving to a mutant strain of H26 (Δ*pyrE2* along with a mutation in Cas8b) obtained from Anita Marchfelder (Ulm University, Ulm, Germany) which contained a deficiency to the Cas system to address and attempt to overcome issues in cloning and manipulation of the CRISPR system later (Fischer et al. 2012). This strain required passage of transformation DNA through a dam-3 host; *E. coli* strain ALS 513 was used to prepare unmethylated plasmid DNA for efficient transformation (obtained from Nancy Kleckner , Harvard University).

3.2.2 Culture and Transformation Procedures

H. volcanii cultures were grown at 42°C, either on a rich (Hv-YPC), minimal (Hv-Min), or casamino acid (Hv-Ca) agar or broth, according to previously established protocols (Allers et al., 2004). The Hv-YPC medium contained yeast extract, peptone (Oxoid), and casamino acids, as well as salts which vary from 14-25% w/v. When necessary, thymidine was added at a concentration of 40 μg/mL for growth of non-transformed H1209, and uracil at a concentration of 50 μg/mL for H26. For *H. volcanii*,

PEG-mediated transformation was performed as previously described (Cline et al., 1989); transformants were selected on either 18% salt w/v Hv-YPC agar, or 18% salt w/v Hv-Ca containing 50 µg/mL5-fluoroorotic acid (5-FOA), a toxic analogue of uracil. Chemically competent Top 10 cells were transformed using standard chemical transformation procedure (ThermoFisher), and grown in Luria-Bertani (LB) medium containing 100 µg/mL of ampicillin at 37°C. When necessary, *E. coli* was grown on LB plates containing 100 µg/mL of ampicillin at 37°C. After all ligations, plasmids were transformed into *E. coli* as described to increase plasmid amount.

3.2.3 Molecular Genetic Methods

All enzymes and reagents for molecular genetic methods were obtained from New England Biolabs unless otherwise stated. For PCR, genomic DNA prepared from exponentially growing cells was amplified for 30 cycles (Promega).

3.2.4 Invader Plasmid Construction

Plasmid pTA409-PAM3-P1.1, which contained protospacer adjacent motif 3 (PAM3) and spacer 1 from locus P1 (P1.1), was obtained from Anita Marchfelder (Fischer et al. 2012). This plasmid was modified by the addition of a luciferase reporter regulated by the constitutive P_{sdh} promoter and synthetic terminator, using the plasmid pTA962-P_{sdh}-CBG MutB created previously by amplifying the succinate dehydrogenase promoter from *H. volcanii* strain Hv1209 with primers Psdh F and Psdh R, then digesting the amplicon and pTA962- CBG MutB with restriction enzymes ApaI and NdeI, and ligating the amplicon to the vector. The reporter element was amplified using primers RepElem-5 and RepElem-3, which added an EcoRI and BglII site to the 5’ and 3’ ends, respectively. The PCR product containing P_{sdh}, and pTA409-PAM3-P1.1 plasmid were

digested with EcoRI and BamHI, then ligated together to create pTA409- PAM3-P1.1-P_{sdh}- CBG MutB (Fig. 3.2A). The BglII and BamHI sites have compatible ends, and the reporter element sat downstream from the PAM3-P1.1 element. A control plasmid was created by restriction digestion of the new plasmid with EcoRI and XhoI, recircularization with Klenow, and blunt end healing to remove the PAM3-P1.1 element, leaving a reporter which should not diminish over time due to a lack of the CRISPR recognition site (plasmid pTA409-P_{sdh}- CBG MutB). Prior to transformation into *H. volcanii* strain H26, both plasmids were passed through *E. coli* strain ALS 513 to prepare unmethylated plasmid DNA for efficient transformation (obtained from Nancy Kleckner, Harvard University).

3.2.5 P_{tna} Regulated Cas6 Fragment Construction

Generating a controllable *cas* gene element was attempted via five routes, all of which involved utilizing the P_{tna} promoter to drive expression of Cas6 and/or other CRISPR associated genes (Fig. 3.2). Three attempts relied upon a double homologous recombination event, while the other two used a single homologous recombination event (Fig. 3.3).

3.2.5.1 *Double Homologous Recombination Event Experiments*

A linearized fragment containing the Cas6 gene and upstream intergenic region which bypassed the Cas6 promoter was needed in order to induce homologous recombination and create a strain in which the Cas6 coding sequence promoter was under control of the tryptophan inducible promoter. This linearized fragment was constructed utilizing pTA962, obtained from Thorsten Allers (University of Nottingham, Nottingham,

UK) (Allers et al. 2010). The plasmid pTA962-P_{tna}-Cas6 was generated by amplifying the 800 base pair Cas6 gene from genomic Hv1209 DNA using the forward primer Cas6-5(NdeI), which added a NdeI restriction site to the 5′ end, and the reverse primer Cas6-3(NotI), which included the addition of a NotI site to the 3′ end. The NotI site was utilized in order to remove an extra XbaI site present on the plasmid, this enabled use of the remaining XbaI site for the addition of upstream Cas6 intergenic region. The Cas6 PCR reaction fragment and pTA962 plasmid were digested with NotI and NdeI then ligated, producing a plasmid wherein Cas6 was under regulation by P_{tna}. Several upstream Cas6 intergenic regions were PCR amplified in order to bypass the endogenous Cas6 promoter and include native terminators to end transcription of the selectable marker present in the plasmid. This upstream intergenic region PCR product (termed US for use in plasmid naming) required both a 5' AflII and 3' end XbaI restriction sites. Several 5' primers were utilized to increase the likelihood of including all desired features, these primers were: CasIngen-5(AflII); alt1; alt2; alt3; alt4; and alt5. The 3' end primer used was CasIngen-3(XbaI). This second cloning step removed the thymidine selectable marker (*hdrB*) from the plasmid, leaving only uracil (*pyrE2*) as the method of selection. A sequential digest of the plasmid, named pTA962-US-P_{tna}-Cas6, using AflII then NotI linearized the fragment of interest, US-P_{tna}-Cas6, which contained the uracil selectable marker, the Cas6 intergenic region upstream of the native Cas6 promoter, the P_{tna} promoter upstream of the Cas6 gene, and the Cas6 gene (Fig. 3.2C).

Once the linearized fragment was created, 5 μg of total digested plasmid was used for transformation, as the fragment of interest was roughly 30 percent of the entire plasmid mass. Transformation followed the protocol described previously (Cline et al.

1989). After this initial transformation, the following modified version of a gene knock-out protocol by (Bitan-Banin et al. 2003) was used. Transformed cells were suspended in transformant plating solution, and 100 µL at a dilution of 1:100 was plated on Hv-Ca +thy plates. Single colonies from this plate were streaked onto another 14.4% salt w/v Hv-Ca +thymidine plate, and individual colonies were cultured in 5 mL 18% salt w/v Hv-YPC broth. After 24 h growth, a 1:500 dilution of this broth was cultured in 5 mL fresh 18% salt w/v Hv-YPC broth; this process was repeated three times. Plates of 18% salt w/v Hv-Ca +5-FOA were cultured using 100 µL of the broth from the last day at a 1:100 dilution, along with a control plate of 14.4% salt w/v Hv-Ca +ura. Any resistant colonies from this plate were then transferred to an 18% salt w/v Hv-YPC +thy plate, then cultured for 24 h. A 5 mL 18% salt w/v Hv-YPC broth sample was then obtained for genomic DNA testing. This DNA was screened using the following primers: LSS-5 and LSS-3, which bind in the genomic CRISPR region and P_{fdx}, respectively; as well as RSS-5 and RSS-3, which bind in P_{tna} and after the genomic Cas6 CDS, respectively. Positive identification of these primer products indicate successful homologous recombination at the point of interest.

The second attempt at a double homologous recombination strategy used oligonucleotides Cas6-5(NdeI) and Cas6fl-3(NotI) to generate the P_{tna}-Cas6 fragment in pTA962, and Alt3-5(KpnI) and CasIngen-3(ApaI) to generate the Cas6 intergenic fragment. Digestion of pTA962 was done with NdeI and NotI to create pTA962-P_{tna}-Cas6, this was followed by the addition of the upstream intergenic region via a sequential digestion of ApaI then KpnI, creating pTA962-US-P_{tna}-Cas6 (Fig. 3.2C). This plasmid was then sequentially digested with KpnI and NotI, and ligated into pTA131 to create

pTA131- US-P_{tna}-Cas6. Two transformations in strain Hv1209 were performed using linearized or circular plasmid DNA, and plated on media containing thymidine, but lacking uracil.

3.2.5.2 *Single Homologous Recombination Event Experiments*

The construction of the first plasmid for use in a single homologous recombination site began with a digest of the plasmid pTA962-P_{tna}-Cas6 using MfeI and ligating the 5.2 kilobase segment together. This plasmid was then digested with XbaI and AflII in order to remove the *hdrB* gene. After heat inactivation of the restriction enzymes, the plasmid was treated with Klenow to form blunt ends, with were then ligated together This new pTA962- P_{tna}-Cas6-Δ*hdrB* plasmid was then transformed into strain Hv1209.

The second experiment utilizing a single recombination event required a new plasmid, and the use of strain H26 containing mutations to the Cas8b gene and pTA409 plasmid obtained from Anita Marchfelder (Fischer et al. 2012). The oligonucleotides used for construction were Cas6-5(NdeI), Cas5-3(BamHI), Cas7-3(BamHI), and Cas8-3(BamHI). Each 3' oligonucleotide was run with Cas6-5(NdeI) to produce a product of 4.9 kilobases (Cas5-3), 4.0 kilobases (Cas7-3), or 2.2 kilobases (Cas8-3). These PCR reactions were completed using genomic DNA from strain Hv1209 and PfuUltra polymerase (Agilent). The PCR product was digested with NdeI and BamHI, then ligated into pTA962. This plasmid was then digested with KpnI and BamHI into pTA131 for later transformation into a modified H26 strain. This modified H26 (Δ*pyrE2*) strain was constructed by homologous recombination using plasmid pTA155, which contained *pyrE2* selectable marker and sequences of genomic DNA flanking *hdrB* for knock-out (Allers et al. 2004). Prior to this knockout protocol, the strain required pTA409 plasmid

curing, which was conducted by growing the H26 strain in 18% salt w/v Hv-YPC +uracil +thymidine liquid media overnight at 42°C, then inoculating more media at a 1:500 dilution and repeating the process over three days. This was then plated at a 1:100 dilution on 14.4% salt w/v Hv-Ca +5-FOA agar and grown overnight at 42°C. Resistant +5-FOA colonies were streaked onto both 14.4% salt w/v Hv-Ca +5-FOA and DNA samples were tested for plasmid via PCR using oligonucleotides AMP and *E. coli* ori.

3.3 RESULTS AND DISCUSSION

Integrating a P_{tna} regulated Cas6/Cas cassette into genomic *H. volcanii* DNA was unsuccessful despite several strategies and attempts (Fig. 3.2 and Fig. 3.3). Due to this, the latter steps of invader plasmid transformation and luminescence measurement were not completed. Had successful confirmation of homologous recombination of the P_{tna}-Cas6 fragment occurred, the invader plasmid pTA-409-PAM3-P1.1-P_{sdh}- CBG MutB would have then been transformed into the *H. volcanii* strain containing the genomic Cas6 gene under control of the inducible P_{tna} promoter, selected under *hdrB* (thymidine) marker. Transformation would have been selected on thymidine 18% w/v salt Hv-Min plates, and colonies would have been patch plated onto a similar plate impregnated with 25 µL of 100 mM beetle luciferin per plate and imaged for luminescence using a Bio-Rad ChemiDoc MP. After initial imaging, 50 µL of 50 mM tryptophan would have been added to each plate to induce *cas* cassette expression. This would have driven the innate CRISPR response from the cell, forcing recognition of the spacer and protospacer adjacent motif in the modified pTA409 plasmid, which would have then been degraded by the cell. This degradation would have presumably lead to a decrease in luminescence observable in real time, and measurable using both ImageJ software for patch plates, and

data from a photomultiplier tube in a test tube set up as described previously in chapter two. A control test using the pTA409 plasmid with the CRISPR elements removed would have demonstrated no change in luminescence after Cas6 expression through tryptophan induction.

Initially, a linear fragment was obtained after digesting pTA962-US-P_{tna}-Cas6 for a double homologous recombination event (Fig 3.2B). Construction of the recombination fragment began with the 3' side, which contained the endogenous tryptophan inducible promoter, P_{tna}, driving transcription of Cas6. This portion of the fragment was created by cloning the Cas6 region from genomic DNA with primers Cas6-5(NdeI) and Cas6-3(NotI), then ligating the created fragment into pTA962-P_{tna}. Restriction enzyme NotI was chosen as it removes a secondary XbaI site from pTA962, which enabled use of XbaI for later ligation. Once the total fragment was linearized and transformed, integration was confirmed via screening transformants using the right side screening oligonucleotides RSS-5, RSS-wt-5 (to test for wild type *H. volcanii* genomic DNA), and RSS-3 (Table IV). While positive results, indicating successful right side integration, were obtained in right side screening, these were likely false positives, as later results indicated only the wild-type oligonucleotide would screen positive with RSS-3.

As successful right side screening was only of use upon left side integration, it was suspected the double homologous recombination event needed to regulate Cas6 with P_{tna} did not occur. The upstream half of the fragment needed for homologous recombination, which contained intergenic regions between Cas6 and the P1 locus of spacers proved problematic to generate, and a variety of oligonucleotides were used in

order to increase the chance of successful PCR; these were labeled alt1-5 (Table V). Digestions of PCR reactions intended to produce this fragment were smaller than the roughly 1 kilobase fragment, and screening using our left side screening oligonucleotides (LSS-5 and LSS-3) never tested positive simultaneously with the RSS oligonucleotides after transformation of the suspected correct linearized fragment into Hv1209. The melting temperature for the left and right side screens varied by several degrees, an all confirmation PCRs utilized multiple samples tested over a 6ºC gradient (55-61ºC) of annealing temperature. None of these temperatures yielded positive results for left side screening. In addition to using different annealing temperatures for confirmation of left side screening, some of the 3' end primers (alt3 and alt5) used to create the Cas6 intergenic fragment were used in conjunction with the 5' left side screening primer to test suspect transformants for any presence of expected left side PCR products; these tests also came up negative.

Testing of the genomic DNA used to obtain the intergenic and Cas6 region fragments was done utilizing oligonucleotides for HVO_0133, the gene for chaperonin thermosome subunit 1, which produced successful PCR products in the past. This test verified the genomic DNA was suitable for PCR reactions, as the expected product, and 300 base pair band, was obtained. The genomic sample also yielded the expected 250 base pair band from a CasIngen-5 and CasIngen-3 oligonucleotide PCR reaction. The pTA962 plasmid was tested via restriction mapping; results confirmed correct sequence using the enzymes AflII, NotI, BglII, NdeI, and XmaI. After repeated attempts to integrate linearized fragments, pTA131 was used to assess whether any issues resulted from use of pTA962. Both linearized and circularized plasmids were obtained from

pTA131 and transformed into Hv1209; neither produced positive PCR screens other than wild type.

As attempts to transform using linearized plasmids proved unsuccessful, a circularized plasmid was used to attempt to integrate the P_{tna}-Cas6 fragment for single recombination events. The first strategy relied upon integration at Cas6; the endogenous Cas6 controlled a truncated Cas6 gene generated during plasmid construction, while P_{tna} regulated the endogenous and complete Cas6 gene (Fig. 3.2 and Fig. 3.3). Digestion of the plasmid with MfeI removed nearly half of the 3' end of the Cas6 gene, which was expected to further reduce any function of the truncated gene controlled by the endogenous *cas* promoter upon successful integration, though the effects of the truncated gene would not be known until transformation. The second attempt utilized a more complete *cas* operon consisting of portions from Cas5, Cas6, Cas7, and Cas8b under P_{tna} regulation. This plasmid would have been integrated into a strain of H26 wherein *hdrB* and *pyrE2* were knocked out, and a mutant in the endogenous Cas8b was present. The integration would have to occur downstream of this mutation in order for the inducible promoter to control a functioning Cas8b gene. The fragment using primers Cas6-5(NdeI) and Cas8-3(BamHI) was successfully created and moved to pTA962, which was then moved to pTA131 with P_{tna} upstream. The required pop-out of the *hdrB* gene in the H26 strain with plasmid pTA155 was never conducted as the strain was never cured of the pTA409 plasmid. Had the strain been cured of pTA409, the curing protocol would have been run once again to remove plasmid pTA155, utilizing media supplemented with thymidine and uracil.

Throughout the study, homologous recombination events were never proven to have occurred. The foundation of this work was based upon this phenomena, which is well documented and utilized heavily in other *H. volcanii* studies (Allers and Ngo 2003; Bitan-Banin et al. 2003; Kapatai et al. 2006). A possible explanation for a lack of homologous recombination include a difference in H1209 and H26 strain sequence and the previously published sequence of the DS2 *H. volcanii* strain; Fischer et al. (2012) reported a 1.5 kilobase discrepancy between the published genome and H119 located at the P1 locus, near where the upstream intergenic region was obtained for this experiment. If this sequence segment does not exist in this strain, double recombination events cannot occur. It is also possible homologous recombination sites under replication stress induces replication errors by the replisome; causing the PCR screens to show up negative (Carr and Lambert 2013). The region consisting of the L11e rRNA terminator, P_{tna}, and perhaps some portion of Cas6 in the introduced DNA may have also provided a third point for homologous recombination between where one of the screening oligonucleotides for either the left or right side would amplify. If this occurred, then successful homologous recombination incorporating the selectable marker could occur while simultaneously providing negative amplification results for screening. The CRISPR/Cas region may be protected from homologous recombination in order to preserve defenses against invader DNA. No studies were found where the promoter for the *cas* cassette was replaced within genomic DNA, only a study involving promoter replacement within a plasmid (Stachler and Marchfelder 2016).

Though the intended experiment never came to fruition, this study emphasizes the need for further understanding of the CRISPR/Cas system. The mechanisms providing

protection from genetic alteration, should they exist, may potentially offer new avenues of research regarding CRISPR/Cas function and Haloarchaea genetics. Studies involving temporal aspects of CRISPR/Cas typically focus on changes over long periods of time in viral or microbial environments, and predominantly utilize microarrays (Emerson et al. 2013; Snyder et al. 2010). As mentioned previously, no studies were found regarding the CRISPR/Cas defense system response time. The luciferase reporter system is uniquely suited for this investigation, and provided the genetic elements can be constructed and integrated properly, the many subtypes of CRISPR systems could quickly be measured and compared.

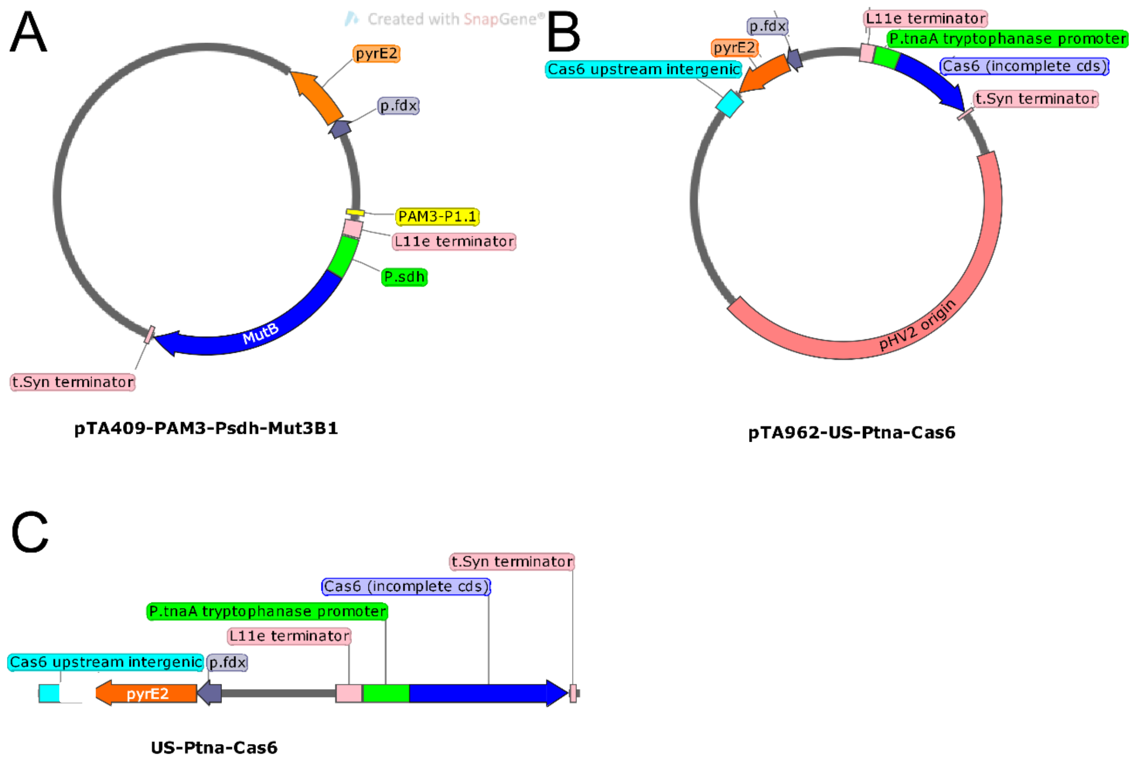
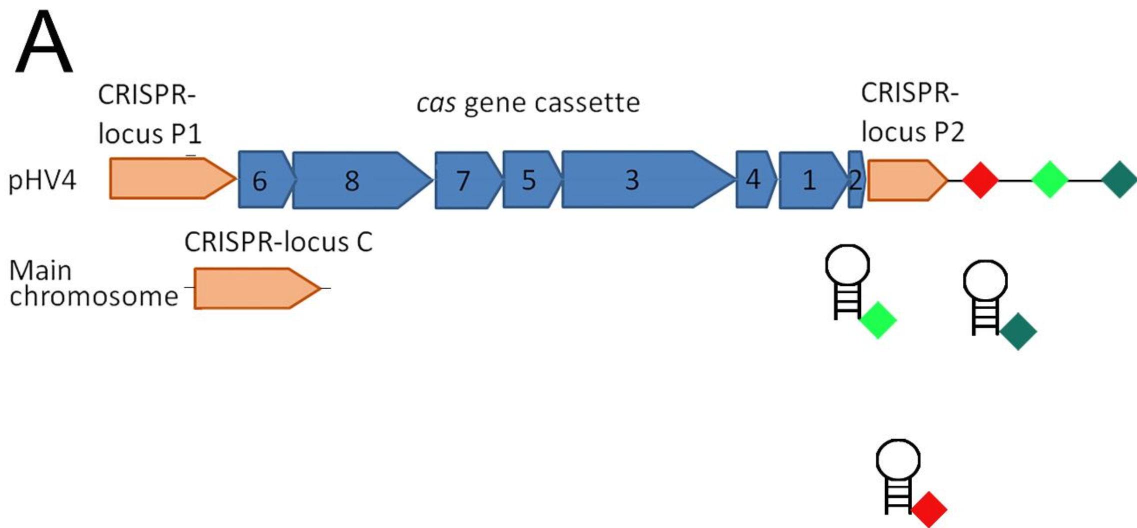
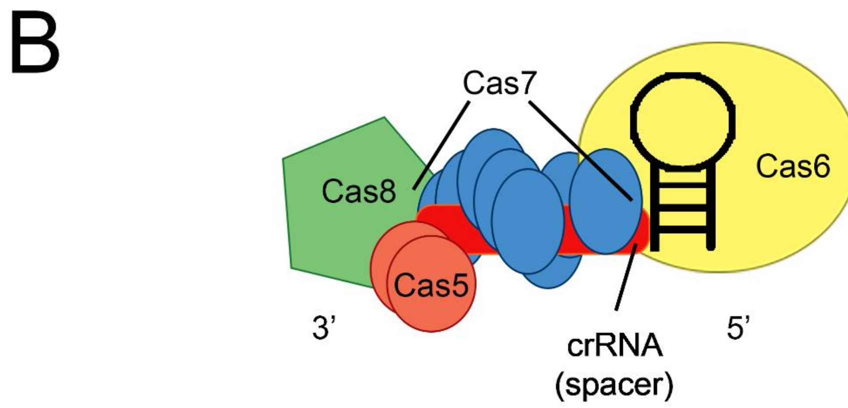


Figure 3.1. Genetic maps of plasmids used for CRISPR/Cas experiments in this paper. (A) Invader plasmid pTA409-PAM3-P_{sdh}- CBG MutB prior to swap of *pyrE2* gene with *hdrB* gene with NcoI and AflII digestion. The constitutive promoter P_{sdh} drives expression of CBG MutB luciferase. Recognition of the protospacer adjacent motif (PAM3) in the host organism drives degradation of the plasmid in order to measure luminescence decay over time. (B) Linearized fragment US-P_{tna}-Cas6, which was transformed in *H. volcanii* as a linear fragment and within (C) pTA962-US-P_{tna}-Cas6 to induce a double homologous recombination event causing P_{tna} to regulate *cas* cassette transcription. The crossover events were intended to occur in the sequence from Cas6 upstream intergenic to Cas6, which would include the *pyrE2* selectable marker as a method for verifying proper recombination. The US-P_{tna}-Cas6 fragment was incorporated into plasmid pTA131 for double crossover homologous recombination.


A
CRISPR-
locus P1
cas gene cassette
CRISPR-
locus P2
pHV4
6
8
7
5
3
4
1
2
CRISPR-locus C
Main
chromosome


B
Cas7
Cas8
Cas6
Cas5
3'
5'
crRNA
(spacer)

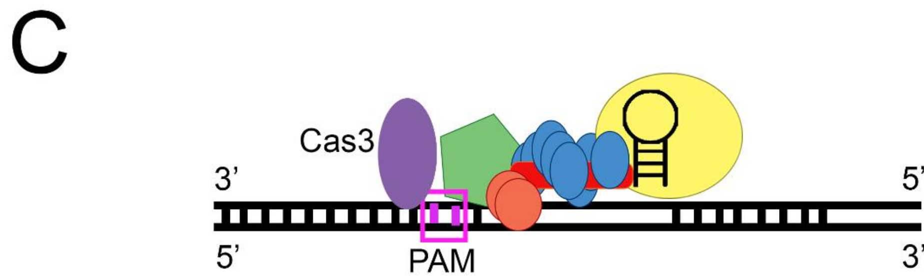

C
Cas3
3'
5'
5'
3'
PAM

Figure 3.2. Type I-B CRISPR/Cas genetic elements and system. (A) *H. volcanii* contains three CRISPR loci; CRISPR locus C, located on the main chromosome. and loci P1 and P2, located on the minichromosome pHV4. The protospacer used for the invader plasmid was obtained from locus P1. Unique spacers are represented by the three colored diamonds, which are then processed with the predecessor repeat forming a hairpin stem at the 5' end. (B) After initial biogenesis, the spacer and repeat are protected by the Cascade complex, consisting of two Cas5, one Cas6, eight Cas7, and one Cas8 protein. (C) This complex will then bind to a specific protospacer adjacent motif (PAM) sequence occurring in invader DNA, recruiting Cas3 protein to cleave this DNA.

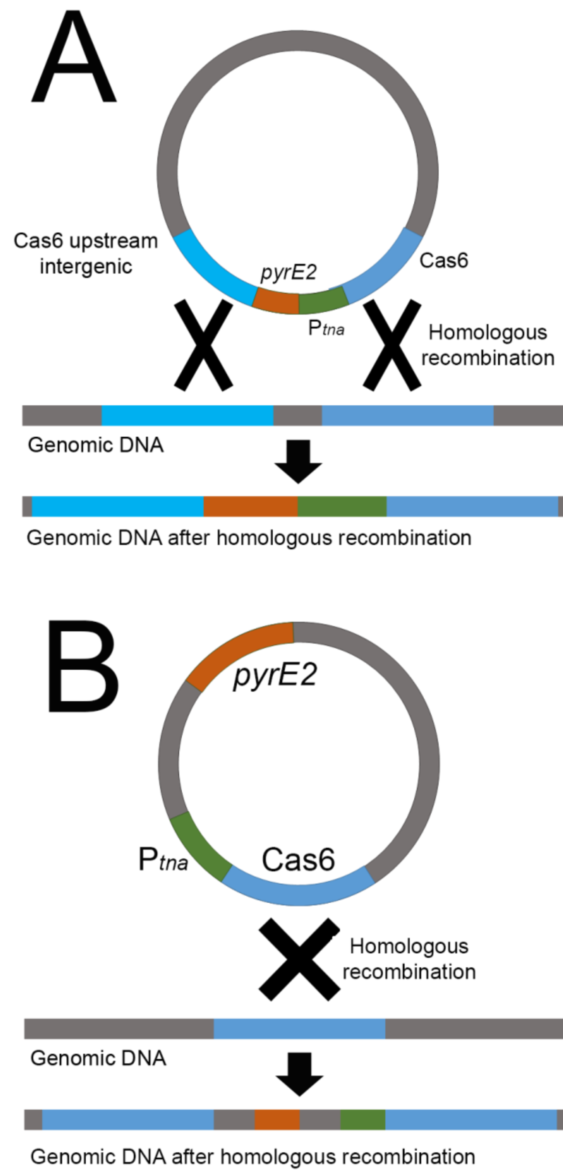


Figure 3.3. Double and single homologous recombination events integrate a selectable marker and tryptophan inducible promoter to regulate expression of the endogenous *cas* cassette. (A) The double recombination event used to integrate pTA962-US-P_{tna}-Cas6 and pTA131-US-P_{tna}-Cas6 into *H. volcanii*. T (B) Single recombination event.

Table IV. Plasmids used

Plasmid	Relevant Properties	Source or reference
pTA409	Shuttle vector based on pBluescript II, with *pyrE2* and *hdrB* markers and *ori-pHV1/4* replication origin	Delmas et al. 2009
pTA409-PAM3-P1.1	pTA409 including spacer 1 of the CRISPR locus P1 (P1.1) and the PAM sequence TTC (PAM3)	Maier et al., 2013
pTA409-PAM3-P1.1-P_{sdh}- CBG MutB	pTA409-PAM3-P1.1 including CBG MutB luciferase regulated by *H. volcanii* succinate dehydrogenase promoter	This study
pTA962	Overexpression vector with *pyrE2* and *hdrB* markers and pHV2 origin, insertion of t.L11e terminator, P_{tna}A promoter and t.Syn terminator	Allers et al. 2010
pTA962-CBG MutB	pTA962 with CBG MutB insertion at NdeI and BamHI sites	This study
pTA962- P_{sdh} -CBG MutB	pTA962 with CBG MutB luciferase regulated by constitutive *H. volcanii* succinate dehydrogenase promoter	This study
pTA962-P_{tna}-Cas6	pTA962 with engodenous cas6 regulated by by tryptophan inducible promoter	This study
pTA962-US-P_{tna}-Cas6	pTA962-P_{tna}-Cas6 with cas6 intergenic region upstream of tryptophan inducible promoter	This study
pTA962-P_{tna}-Cas6-Δ*hdrB*	pTA962-P_{tna}-Cas6 with *hdrB* coding sequence excised	This study
pTA131	Integrative vector based on pBluescript, with pyrE2 marker	Allers et al. 2004
pTA131- US-P_{tna}-Cas6	pTA131 with cas6 intergenic region upstream of tryptophan inducible promoter	This study
pTA155	pTA131 with HindIII-XbaI PCR fragment containing flanking regions of hdrB	Allers et al. 2004

Table V. Oligonucleotides used

	Primer	Sequence (5'-3')	Notes
1	Psdh F	ATCAATGGGCCCGACCGCATCTGCACGATTAC	5'ApaI
2	Psdh R	TCAGTACATATGCTCTTGTGAGTGTCGGGAAG	3'NdeI
3	RepElem5	ATGCCTGAATTCCGTAATACGACTCACTATAGGG	
4	RepElem3	ATGCCTAGATCTAACAAAAGCTGGAGCTCCAC	
5	HVO0133F	TCAAATGGGCCCAACGGAAGCACAACGAACTG	5'ApaI
6	HVO0133R	ACGAGCCATATGAGTCATCGCCTGATTGTTTG	3'NdeI
7	Cas6-5	ACTGTTCATATGCGTATAGAATTAGCGCTCG	5'NdeI
8	Cas6-3	ACTGTTGCGGCCGCAGGAAGCCAAACCCGTGTTC	3'NotI
9	casIngen5	ACTGTTGGGCCCTCTCGCGTACGACTAACTTCTG	5'AflII
10	casIngen3 (XbaI)	ACTGTTTCTAGATCTCGCGTACGACTAACTTCTG	3'XbaI
11	casIngen3 (ApaI)	ACTGTTGGGCCCTCTCGCGTACGACTAACTTCTG	3'ApaI
12	alt1	AGTTCTTAAGACCACGTTGCTAATAGCTCC	5'AflII
13	alt2	ATGTTCTTAAGTGCAGCTCGTTTGATGAGAC	5'AflII
14	alt3	ATGTTCTTAAGTGTATGCGATATCCCTTAAAAC	5'AflII
15	alt4	ACTGTTCTTAAGCGACTACGACCTGATGCG	5'AflII
16	alt5	ACTGTTCTTAAGAAGCCGTCGATGAAGTCCT	5'AflII
17	LSS5	GATTGAAGTCTAGACGGCAC	
18	LSS3	CCCGAACAGCAACTACTATG	
19	RSS5	CTGCCGATTACTTCACATTC	
20	RSS3	CGATATCTGGACCTGTCATC	
21	Cas6fl-3 (NotI)	CATGTTGCGGCCGCAGTCGTCGATATCTGGACCTG TC	3'NotI
22	Alt3-5 (KpnI)	ACTGTTGGTACCGTGTATGCGATATCCCTTAAAAC	5' KpnI
23	Cas5-3	CATCGATGGATCCAAAATACGGAGGGTGTTCGG	3'Bam HI
24	Cas7-3	ATCCGATGGATCCGACATTGTCGCCTCGTTAGG	3'Bam HI
25	Cas8-3	CATCGATGGATCCTTAGTTCGTGGTGCTCTCAG	3'Bam HI
26	Amp	CTGGATGGAGGCGGATAAAG	
27	*E. coli* ori	GTTCGGTGTAGGTCGTTCGC	

CHAPTER IV

GENETIC TRAPPING AND CREATION OF A PROMOTER LIBRARY UTILIZING A LUCIFERASE REPORTER SYSTEM

4.1 INTRODUCTION

Genetic traps are useful tools in identifying elements of genetic regulation in an organism. A variety of methods exist which enable researchers to design vectors to identify enhancers, promoters, or genes depending on the reporter system (Springer 2000). Two methods were used in this study to create a library of random fusion proteins with luciferase: one was integrating reporter gene downstream of host DNA possibly containing a promoter into the host chromosome, and the other was integrating portions of host DNA upstream of a promoterless reporter in a plasmid containing an origin of replication of the host organism. In both cases, expression of the reporter can only occur if the inserted DNA is within frame and correct orientation. These studies were intended to reveal promoters and other genetic elements regulating protein expression in *H. volcanii*, which would then be later used to investigate the effects of environmental changes on protein expression.

The first library was constructed using a plasmid containing an insertion of fragments from a partial digestion of genomic *Haloferax volcanii* DNA upstream of a luciferase reporter, lacking a promoter and a start codon, with an origin of replication for *H. volcanii*. After transformation into the haloarchaeon, if the fragment of genomic DNA contained a promoter and a start codon in-frame and with the promoter, then transcription of the luciferase would be detectable by the appearance of a luminescent colony, if the

colony was grown in an environmental condition whereby the promoter was active. A similar method which relied on homologous recombination used a plasmid containing no origin of replication or homology other than the integrated fragment for *H. volcanii*. Single recombination of potential fragments would result in stable integration of the reporter plasmid. The majority of luciferase insertions into the genome are expected to predominately be out of frame or include DNA fragments lacking promoters and start codons. While other studies in this paper detail integration of selected promoter regions upstream of the luciferase, this avenue of research potentially provides a much cheaper and efficient method of surveying activity of multiple promoters simultaneously. While a complete promoter library is not very realistic regarding the limitations in the methodology, given the sheer number of outcomes, a large working library should be assembled, with potentially hundreds of promoters available for screening.

After generation of the protein fusion and promoter libraries, screening can be done for responses to changes in the environment; including light/dark transitions, UV irradiation, heat/cold shocks, hyper- and hypoosmotic transitions, pH shifts, and nutrient changes. Upon completion of the genetic trap, the luciferase reporter is expected to integrate near a variety of promoters. If the promoters are either directly or indirectly regulated by the presence or absence of any of the specific environmental changes investigated, a measurable change in luminescence over time should occur. Because the genome of *H. volcanii* is published and sequenced, determining what gene the endogenous promoter regulates should prove relatively easy by using primers that bind and sequence upstream or downstream of the luciferase.

4.2 MATERIALS AND METHODS

4.2.1 Strains, Plasmids, and Primers

The plasmids and primers used for library construction or DNA manipulation in *H. volcanii* or *E. coli* are shown in Table VI and Table VII, respectively. *H. volcanii* strain H1209 was used for all luciferase studies in *H. volcanii*, which allowed selection of successful transformants on media without thymidine (Allers et al., 2010).

4.2.2 Culture and Transformation Procedures

H. volcanii cultures were grown at 42°C, either on a rich (Hv-YPC), minimal (Hv-Min), or casamino acid (Hv-Ca) agar or broth, according to previously established protocols (Allers et al., 2004). The Hv-YPC medium contains yeast extract, peptone (Oxoid), and casamino acids, as well as salts at 18% w/v. When necessary, thymidine was added at a concentration of 40 μg/mL for growth of non-transformed H1209. For *H. volcanii*, PEG-mediated transformation was performed as previously described (Cline et al., 1989), transformants were selected on 18% salt w/v Hv-YPC agar. Chemically competent Top 10 *E. coli* cells were transformed using standard chemical transformation procedure (ThermoFisher), and grown on Luria-Bertani (LB) medium containing 100 μg/mL of ampicillin at 37°C.

4.2.3 Molecular Genetic Methods

All enzymes and reagents for molecular genetic methods were obtained from New England Biolabs unless otherwise stated.

4.2.4 Promoter Library Creation Methods

The first library attempted for construction was a promoter library wherein endogenous promoters would drive expression of CBG MutB. Genomic *H. volcanii* DNA

was isolated from strain Hv1209 using a previously published protocol with the following modifications: after cells were lysed, extraction was completed with 400 μL phenol/chloroform/isoamyl alcohol for two rounds, followed by one round of 400 μL chloroform extraction, and ethanol precipitation (Allers et al. 2004). Samples of genomic DNA were restriction digested with AciI for either 20, 30, or 60 minutes, while plasmid pTA131- CBG MutB was restriction digested with ClaI for 1 hour at 37°C and was then treated with TSAP. Both the plasmid and digested genomic DNA were run on an agarose gel. Genomic DNA was isolated and purified according to base pair length and time spent during restriction digest, with samples separated from 200-500, 500-1000, and 1000-1500 base pairs. After isolation and purification, these nine samples were independently ligated with the digested pTA131- CBG MutB plasmid, creating a library of plasmids labeled pTA131- CBG MutB-HvPromoter. The ligations were then transformed into *E. coli,* plated, and cultured overnight. The colonies and DNA were then collected in the same manner as in the protein fusion protocol. After library plasmid DNA was obtained, strain Hv1209 transformations were conducted using 1000 ng, 500 ng, 200 ng, and 50 ng of plasmid DNA and plated. Colonies were patched by transformation DNA amount onto 18% salt w/v Hv-Min agar and imaged as previously described in chapter two (Fig. 4.1). Plasmids from the pTA131- CBG MutB-HvPromoter library were digested with AciI and run on an agar gel for confirmation of *H. volcanii* genomic DNA.

4.2.5 Protein Fusion Library Creation Methods

To create the protein fusion library, CBG MutB was amplified, with an EcoRI site before the stop codon and a three frame stop cassette added. The oligonucleotide CBG99endFUZ contained these genetic elements, and was used in conjunction with

CBG99srt for luciferase amplification. Thermo Phusion polymerase was used to amplify the luciferase from pTA962-MutB plasmid. The amplicon and pTA962 were digested with NdeI and BamHI and ligated together, creating pTA962- CBG MutB-Fuz. The new pTA962- CBG MutB-Fuz plasmid, as well as pTA962- CBG MutB and pTA962-CBG99, were transformed into Hv1209, then cultured in 3 mL 18% salt w/v Hv-YPC broth overnight. Optical density readings were recorded and each vial was impregnated with 3 µL of 100 µM beetle luciferin (Promega E160E). Culture media was removed from each vial and placed within a single well of a 96 well plate in triplicate at the following volumes: 200 µL, 100 µL, and 50 µL. These were then measured for luminescence intensity using a Bio-Rad ChemiDoc MP to observe if changes occurred due to sequence differences between CBG MutB-Fuz, CBG MutB, and CBG99 (Fig. 4.2). The plate was incubated at 42°C overnight and measured again. In addition to these measurements, 1 mL of the remaining culture from the first day of measurement was placed in an Eppendorf tube and measured once a day for five days via a Berthold Detection Systems FB 12 Luminometer after opening and closing the Eppendorf tube lid, then shaking the Eppendorf tube 4 times.

Genomic *H. volcanii* DNA was isolated from strain Hv1209 using a previously published protocol with the following modifications: after cells were lysed, extraction was completed with 400 µL phenol/chloroform/isoamyl alcohol for two rounds, followed by one round of 400 µL chloroform extraction, and ethanol precipitation (Allers et al. 2004). Genomic *H. volcanii* DNA was digested with MluCI, a four base pair restriction enzyme, for 2 hours at 37°C. The pTA962- CBG MutB-Fuz plasmid was digested according to the following protocol: EcoRI for one hour at 37°C followed by heat

inactivation for 20 minutes at 65°C, then TSAP treatment for 30 minutes at 37°C followed by heat inactivation at 74°C for 20 minutes. These products were run on an agarose gel for verification and isolation. The genomic DNA from the 500-1500 base pair length and the pTA962- CBG MutB-Fuz plasmid were purified individually, and ligated. The ligations were then transformed into *E. coli,* plated, and cultured overnight. Luria-Bertani broth containing 100 μg/mL of ampicillin at a volume of 10 mL was poured over the plates and left for 20 minutes; the colonies were then lifted off the plates into the LB broth, and cultured for 6 hours. A miniprep was conducted on the culture, and the new plasmid pTA962- CBG MutB-Fuz-HvGenomic was transformed into Hv1209 and plated.

Twenty colonies from this transformation were then grown in 3 mL of 18% salt w/v Hv-YPC broth, along with pTA962-CBG99, pTA962- CBG MutB, and pTA962-CBG Mut-Fuz controls. After overnight growth, 1 mL was taken from these cultures and diluted with 18% salt w/v Hv-YPC broth to an optical density of OD_{600}= 2; each culture then received 1 μL of 100 μM beetle luciferin. Luminescence measurements were conducted using a luminometer every 20 minutes for 3 hours. Samples were kept at shaking at 150 rpm at 42°C when not being measured. After this experiment, four colonies were selected based on luminescence (one bright, one dim, one intermediate, and one displaying a broad range intensities) and grown overnight in 3 mL 18% salt w/v Hv-YPC broth, along with three pTA962- CBG MutB and one pTA962-Mut-Fuz controls. These cultures were similarly diluted to OD_{600}= 2, received 1 μL of 100 μM beetle luciferin, and luminescence was monitored every 20 minutes. After the first hour of measurements, 3 μL of 10 mg/mL anisomycin in DMSO was added to each culture. Monitoring continued for a total of ten time points over 200 minutes (Fig. 4.3).

Eighty colonies from original Hv1209 pTA962-CBG MutB-Fuz-HvGenomic transformation plate, along with ten colonies each of both pTA962- CBG MutB-Fuz and pTA962- CBG MutB were patch plated as a 10 x 10 grid onto 14.4% salt w/v Hv-YPC agar and cultured. This plate was cultured overnight at 42°C and then received 48.4 µL of 100 µM beetle luciferin evenly distributed across the plate at the intersections between patches (0.4 µL luciferin per intersection). This plate sat in darkness at 42°C for 30 minutes and was then imaged every 30 minutes for 1 hour as previously described (Fig. 4.4).

4.3 RESULTS AND DISCUSSION

Despite the innate differences of how each was designed and implemented, the two libraries were created in order to further study gene expression in *H. volcanii*. Integration of genomic *H. volcanii* DNA in both libraries relied upon digestion using a 4-base pair cutter to form many fragments of varying size for ligation into the host vector. Due to the design of these experiments, plasmids and transformations were divided by fragment size rather than sequence. Because of this, there was an inherent risk of out-of-frame fragment ligations, individual ligations containing multiple fragments, and cells being transformed with multiple plasmids. Out-of-frame ligations would result in either unsuccessful protein-protein fusions or luciferase without a working promoter, depending on which library was used. The digested genomic fragments could also ligate to one another before integration into the plasmid in either library. Though genetic sequencing would verify if this event occurred, the cost to conduct these screens would have offset the advantages of using the techniques to make the libraries. Similarly, it was never determined if only one plasmid was transformed into each successful transformant. If

multiple plasmids with in-frame genetic elements transformed into a cell, it could not be determined what effect each had on luminescence expression or protein-protein fusion. As every colony in the promoter library was luminescent, it is suspected multiple plasmid copies containing different promoter regions remained in the cell post transformation. Additionally, *H. volcanii* is known to regularly undergo gene conversion; and therefore, the number of chromosomes with any integrated luciferase caused by homologous recombination would increase until the number of chromosomes with the change equaled the number of chromosomes without (Lange et al. 2011). This would further ensure observing luminescence in colonies transformed with multiple plasmids.

Using restriction digestion and PCR, integration of CBG MutB-Fuz into pTA962 was confirmed, and the modified 3' end of the fusion library luciferase amplicon allowed protein fusion at the C-terminal end of the luciferase protein. Positioning the endogenous sequence of the hypothetical fusion protein at the end of the luciferase sequence increases the likelihood of creating a functional fusion protein, as most proteins fold with either the N- or C-terminus exposed on the protein surface (Hovmoller and Zhou 2004). Functional protein-protein fusion events would result in a library of proteins linked with a luciferase which could then be used to measure native protein expression levels *in vivo*. The individual protein-protein fusions created in the performed experiments were regulated by the promoter P_{tna}, though this promoter could be excised and the region containing the native promoter could be placed upstream after sequencing to determine what protein was produced. Homologous recombination using one of these plasmids would allow incorporation of the reporter system into the genome, which could then be used for further study in gene expression and regulation under a variety of environmental changes.

A potential concern regarding this approach is the luciferase fusion causing possible disruption of a gene's function, downstream signaling, or the natural response to environmental change. *Haloferax volcanii* is a polyploidy organism, with 10 to 25 genomic copies in each cell, so these interruptions were considered to be inconsequential to cell survival, so integrating into one or several is expected to have only minor (if any) consequence (Breuert et al. 2006). To ensure the integrated reporter does not disrupt normal cellular behavior, confirmation in wild type cells by q-PCR should report any changes of gene expression seen by the promoter-reporter strategy. However, in some cases the stimuli used to induce a shift in regulation may affect the luciferase reaction itself rather than be a true change in regulation/expression of the luciferase. These stimuli should be universal among all luminescent colonies subjected to the stimuli, and therefore distinguishable from single colonies actually responding due to alterations in underlying genetic control. If this phenomenon proved to be a problem, utilizing environmental pulses (e.g. a 30 min heat shock) rather than long-lasting environmental shifts would have been one method used in an attempt to minimalize the decrease in luminescence.

Results regarding the first genetic trap method indicate the strategy does indeed work, but as every colony screened positive, there are still modifications needed before this library is of use (Fig. 4.1). Such a seemingly high rate of transformation may have occurred due to a high number of plasmids transforming, as previously suggested, or be a result of higher transformation competency than previously thought. As only a single restriction enzyme was used, ligations including multiple genomic insets were possible; depending on the size and orientation of these inserts, a greater chance of obtaining an in

frame promoter may have been possible than expected. Further experiments where done wherein the amount of genetic material being transformed was reduced (down to 1 μg DNA per transformation), though this did not prevent every subsequent colony from expressing luminescence. If a region of unknown homology exists within the plasmid, homologous recombination may have occurred at this site, also explaining why every colony tested positive. This is unlikely however, as the pTA962 and pTA131 plasmids were extensively used for other experiments over several years, and never presented similar issues.

Luminescence comparison between CBG99, CBG MutB, and Mut-Fuz suggest that the Mut-Fuz luciferase is brighter than CBG99, and is brighter than CBG MutB (Fig. 4.2). Genetic differences between Mut-Fuz and CBG MutB appear only as a total of 15 base pair insertion at the end of the coding sequence, and may contribute to the difference in luminescence. This experiment was conducted with stationary and shaken cultures, and each group exhibited similar results; therefore it is unlikely available O_2 levels were the cause of the difference in luminescence. Results from anisomycin treatment to reduce protein expression in Hv1209 in the experiment involving four colonies and controls helps to confirm the fusion does not significantly alter persistence of the luciferase over time, which may cause a perceived increase in luminescence (Fig. 4.3). The pTA962-CBG MutB-Fuz-HvGenomic transformations into Hv1209 resulted in every screened colony exhibiting luminescence. Differences in observed luminescence intensities between colonies screened indicated some alteration of the luciferase occurred, and suggest portions, if not the entire protein, were translated; though it was never determined if the other portion of the suspected protein fusions were functional (Fig. 4.3 and Fig.

4.4). In addition, if the protein fusions are observed to cause little change in how the luciferase operates, it is possible the endogenous protein may retain some function within the cell, which would allow investigations regarding natural protein degradation within *H. volcanii* to be measured over time.

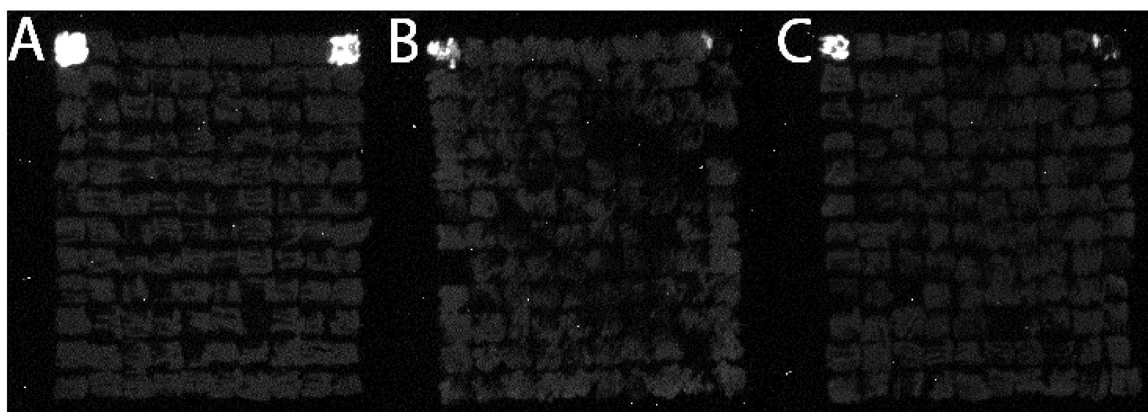
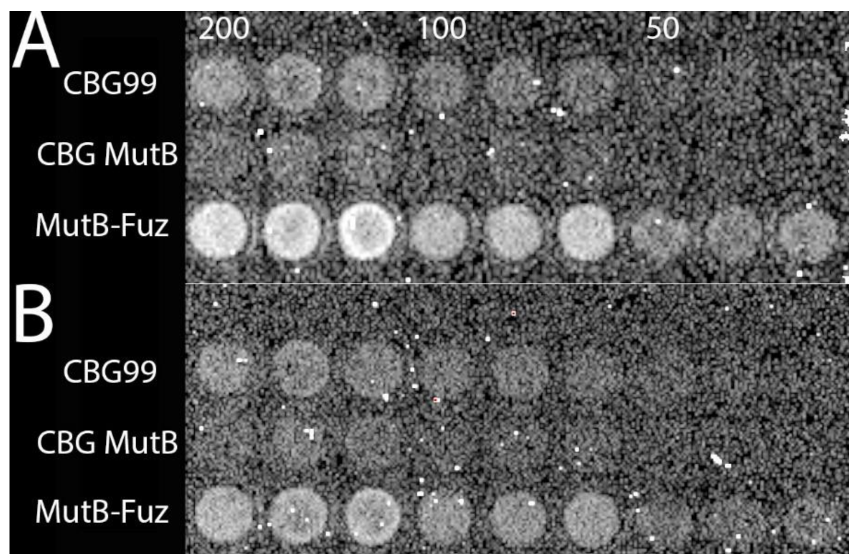
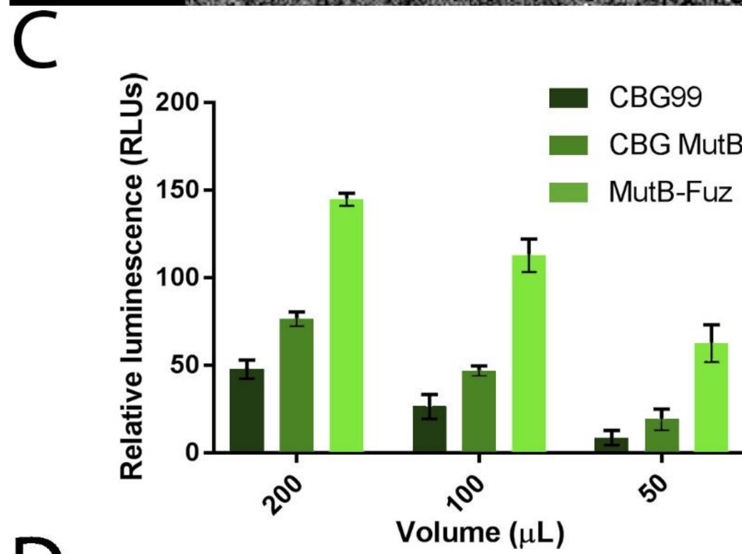


Figure 4.1. Screen for luminescence of *H. volcanii* cells transformed with pTA131- CBG MutB-HvPromoter. The HvPromoter region was a sequence of genomic *H. volcanii* DNA obtained after restriction digestion with AciI. Each plate represents HvPromoter sequences obtained from this digestion of varying lengths: (A) 1000 base pairs; (B) 500 base pairs; (C) 200 base pairs. All colonies of any HvPromoter length successfully lumineced, regardless of the amount of plasmid DNA transformed, indicating an unexpected issue regarding, previously known transformation data rate about *H. volcanii*, or plasmid contruction.


A
B
200
100
50
CBG99
CBG MutB
MutB-Fuz
CBG99
CBG MutB
MutB-Fuz


C
Relative luminescence (RLUs)
200
150
100
50
0
CBG99
CBG MutB
MutB-Fuz
200
100
50
Volume (μL)

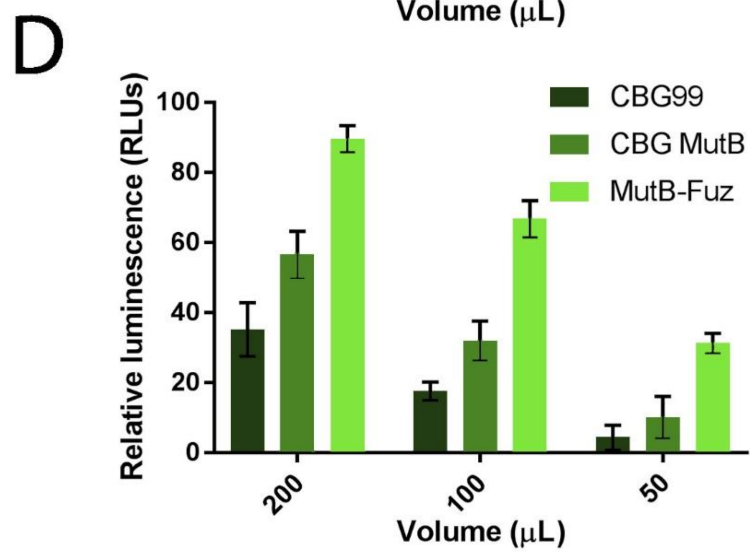

D
Relative luminescence (RLUs)
100
80
60
40
20
0
CBG99
CBG MutB
MutB-Fuz
200
100
50
Volume (μL)

Figure 4.2. Comparison of luminescence between three luciferase reporters at several volumes. Under control of a tryptophan inducible promoter, luciferase reporters CBG99, CBG MutB, and CBG MutB-Fuz were compared in *H. volcanii*. (A) Three luciferases were separated in a 96 well plate, with CBG MutB, CBG99, and CBG MutB-Fuz placed in the top, middle, and bottom rows, respectively. Aliquots of 200, 100, and 50 μL were done in triplicate for each sample, and the plate was imaged for 10 min. (B) The plate incubated at 42 ºC for 24 h and was reimaged for 10 min. Comparison of luminescence between luciferases at day 1 (C) and day 2 (D). All bars show mean ± standard deviation.

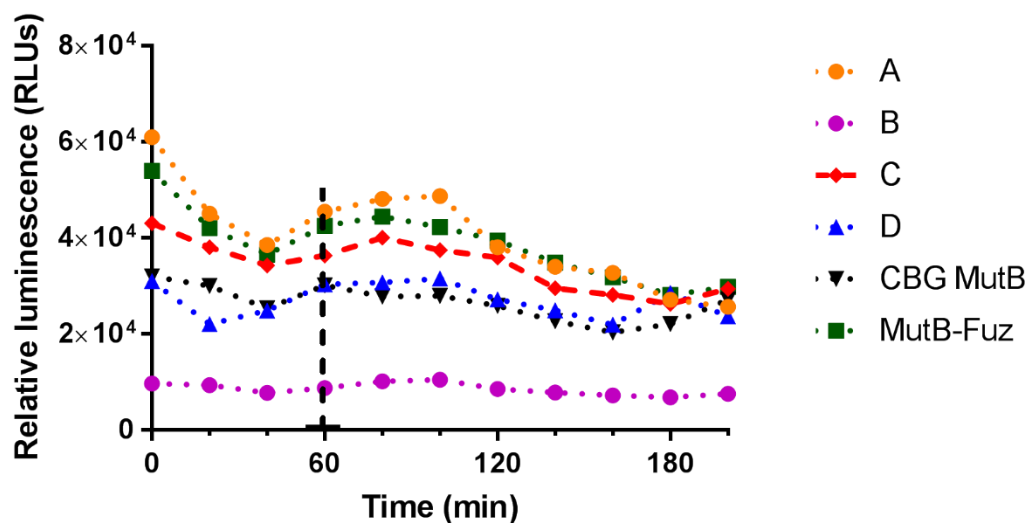


Figure 4.3. Measurement of luminescence over time of protein fusion luciferase reporters. Four colonies from an initial twenty were selected based on differences luminescence when compared to the CBG MutB-Fuz control: one bright colony; one dim colony; and two intermediate colonies; labeled A, B, C, and D, respectively. Each colony was cultured overnight and diluted to an optical density of OD_{600}= 2. Luciferin was then added to each sample, and luminescence was measured every 20 min for 200 min. After 60 min, 3 μL of 10 mg/mL anisomycin in DMSO was added to each culture to measure changes in expression. Luminescence peaked for each colony between 60-100 min, then fell until all samples but colony B were measured around $3x10^4$-$4x10^4$ RLUs. All samples displayed similar behavior regardless of luminescence intensity.

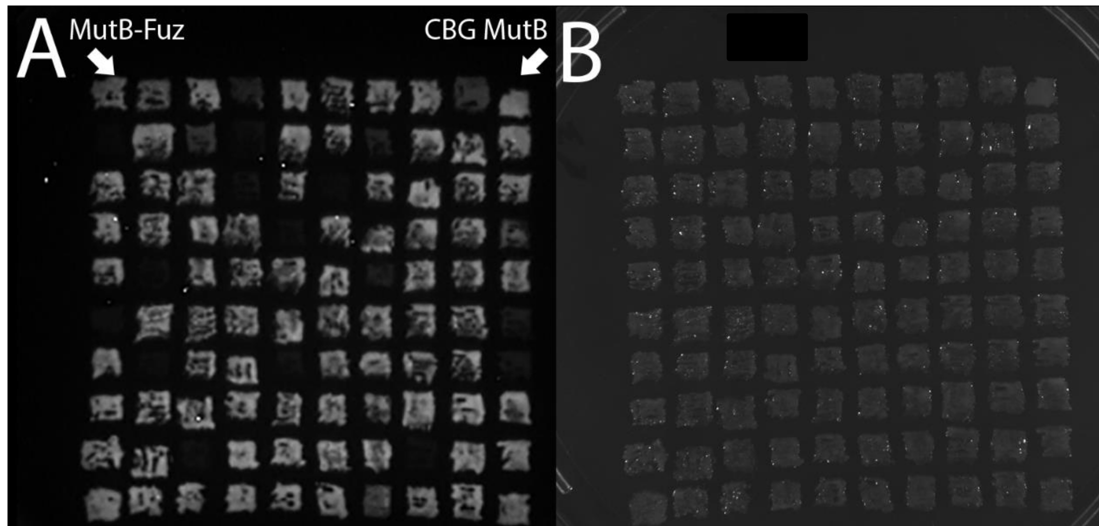


Figure 4.4. Patch plate of *H. volcanii* luciferase-protein fusion colonies. (A) Eighty tranformants were selected on plates transformed with pTA962- CBG MutB-Fuz-HvGenomic and were compared with ten colonies each of either pTA962- CBG MutB-Fuz or pTA962- CBG MutB (arranged diagonally as indicated starting in either the upper left or right corner). After a five minute exposure, several colonies exhibit decreased luminescence, presumably caused by the additional protein fusion to the luciferase resulting from ligation of genomic *H. volcanii* DNA to luciferase DNA. (B) Brightfield image of (A).

Table VI. Plasmids used

Plasmid	Relevant Properties	Source or reference
pTA131	Integrative vector based on pBluescript, with pyrE2 marker	Allers et al. 2004
pTA131-CBG MutB	pTA131 with CBG MutB insertion at NdeI and BamHI sites	This study
pTA131-CBG MutB-HvPromoter	pTA131-CBG MutB with *H. volcanii* genome sequence inserted at ClaI site	This study
pTA962	Overexpression vector with *pyrE2* and *hdrB* markers and pHV2 origin, insertion of t.L11e terminator, P_{tna}A promoter and t.Syn terminator	Allers et al. 2010
pTA962-CBG MutB	pTA962 with CBG MutB insertion at NdeI and BamHI sites	This study
pTA962-CBG MutB-Fuz	pTA962-CBG MutB with EcoRI site intergrated upstream of a stop codon and a three frame stop cassette added	This study
pTA962-CBG MutB-Fuz-HvGenomic	pTA962-CBG MutB-Fuz with *H. volcanii* genome sequence inserted at EcoRI site	This study
pTA962-CBG99	pTA962 with CBG99 insertion at NdeI and BamHI sites	This study

Table VII. Oligonucleotides used

	Primer	Sequence (5'-3')	Notes
1	CBG srt	TTCGAAATCGATATGGTGAAGCGTGAGAAAAATG	5' ClaI
2	CBG99endFUZ	AATCTTGGATCCCTCTCCAACAATTG	3' EcoRI

CHAPTER V

5.1 RESULTS AND DISCUSSION

This work demonstrates the utility of the salt-tolerant luciferase reporter in investigations regarding gene regulation and protein expression in *H. volcanii*. Traditionally, research involving the temporal aspects of genetic elements in *H. volcanii* relied heavily on reporter systems with high equipment cost or time investiture; this reporter system provides researchers with an inexpensive, high throughput tool with real time results. Organisms can be maintained *in vivo* with no need for collection or sacrifice, allowing time point studies to observe changes in an entire population without removal of any organisms from the population. The luciferase and controlling elements are easily shuttled using common genetic methods with no adverse effect, and the high salt environment required to culture *H. volcanii* appears to have little effect upon luciferin. Several mutations to the luciferase CBG99 were found to greatly increase luminescence and suitability for genetic investigations in a high salt environment. Currently, these mutations are characterized by effect on luminescence intensity and amino acid changes to the protein. The reporter system is capable of function in liquid and solid media, though O_2 availability in liquid media may influence the luciferase-luciferin reaction, causing a decrease in luminescence not necessarily reflective of actual expression; this can be mitigated by aeration of the culture. The newly developed luciferase appears to function much like other halophilic proteins, with decreased mesohalic function. Codon optimization was not found to have effect upon luminescence, which may contradict findings in other *H. volcanii* studies (Allers et al. 2010; Hartman et al. 2010). The challenges regarding correct homologous recombination and transformation rates

observed also serve to further demonstrate current knowledge regarding genetic transformation and integration may not be as well established as previously thought, or perhaps additional regulators to such processes still undiscovered are present. While the luciferase is currently suitable for use in studies, further characterization and manipulation is expected to improve the luciferase in regards to its use in measuring the timing of genetic elements as well as current knowledge on said genetic elements and protein engineering for halophilic organisms.

5.2 FUTURE DIRECTIONS

The information presented in this work regarding characterization of the CBG MutB luciferase would greatly benefit from protein structure studies to determine where each amino acid change occurs in the folded protein. As mentioned previously, mutations resulting in an increase of acidic amino acid residues, or a decrease of hydrophobicity, R group size, or protein surface area are expected to increase function/luminescence. Interestingly, convertant 5 at first appeared to reduce luminescence (though later analysis proved the reduction was not significant) compared to CBG99. This would be unusual as the hydrated amino acid arginine is greatly selected over lysine in halophilic proteins, and arginine has also been indicated as improving thermostability in firefly luciferase when present at the protein surface (Jaenicke, 1991; Solgi et al., 2016). Because of this, it is speculated the amino acid is either not present on the protein surface, or is directly inhibiting the protein active site. Convertant 11, a lysine to asparagine mutation, was expected to increase signal intensity, as lysine is generally not favored for halophilic proteins; though other studies report lysine to asparagine mutations in green fluorescent protein increases protein stability while decreasing signal strength (Siglioccolo et al.,

2011; Sokalingram et al., 2012; Solgi et al., 2016). As the experiments in the previous chapters primarily concerned *when* gene expression occurred, the ideal luciferase for such studies would present a strong yet brief signal. A procedure similar to the error prone PCR luciferase development could be conducted on CBG MutB, specifically searching for mutations with no reduction in luminescence intensity and reduction in signal half-life. Knowing the actual structure of the protein would help answer all of these questions, and provide further information to be applied in the engineering of other proteins.

Improving the luciferase based upon information from chapter two does not wholly rely upon determination of protein structure for future protein engineering efforts. Experiments involving the mutation of click beetle red luciferase (CBR) were conducted to demonstrate the mutations improving luminescence intensity in CBG99 could be applied to a similar, but naive system. The DNA sequence of click beetle red is very similar to CBG99; sixteen bases are different in the entire sequence (Fig. 5.1). Interestingly, in *H. volcanii* CBR and CBG display similar luminescence intensity (Fig. 5.2). Site-directed mutagenesis was performed as described in chapter two for CBR using all CBG MutB mutation points. Transformants were then plated on 18% YPC agar and imaged; and the luminescence intensity was measured for comparison (Fig. 5.2). Further research could be conducted reverting the nucleotides differing from CBR to CBG99 to investigate if changes could result in a luminescence to the same intensity of CBG MutB in CBR. Selected mutations could then be applied to CBG MutB to determine if luminescence could be increased further. The two luciferases could then be used in tandem to investigate a variety of experiments; Figure 5.3 displays the results of imaging of the luciferases with a green and a red filter. Promoters regulating different

metabolic pathways could be used to measure if *H. volcanii* exhibits substrate preferences. An experiment on plasmid retention rates could be conducted on a population of Hv1209 transformed with two plasmids conferring either *pyrE2* or *hdrB* selectable markers and measured for changes in red and green luminescence when grown in minimal, rich, +thy/-ura, or +ura/-thy media. Competition studies, using two populations of *H. volcanii*, or *H. volcanii* and another organisms transformed with either green/red luciferase could be investigated. Additionally, a dual luciferase reporter could be used to research the effects of plasmid origin and copy number in the CRISPR/Cas system, or for developing different genetic traps in a single population.

Although the CRISPR/Cas and genetic trapping experiments presented challenges preventing the experiments from working as intended, other experiments involving the underlying principles are still possible. The CRISPR/Cas response has been shown to respond to invader plasmids differently depending on the presence of a *H. volcanii* origin of replication, with different origins of replication eliciting different responses, though it is unknown whether this is due to steric hindrance occurring on the plasmid (Maier et al. 2013). This publication also suggested overcoming the CRISPR/Cas response by a high copy number of plasmids is not important. As mentioned earlier, using a luciferase reporter, with the results measured real time, an important factor given the CRISPR/Cas response is thought to initiate destruction of invader plasmid immediately after the invader is detected (this hypothesis could also potentially be investigated with a luciferase reporter). While the method for trapping genetic elements did not result in a library, certain promoters of interest from *H. volcanii* were incorporated into plasmids to regulate CBG MutB expression, particularly those thought to likely influence any

rhythmic behavior. The previously mentioned constitutive succinate dehydrogenase promoter from *H. volcanii* was initially developed for use as a control for experiments involving other promoters; these targeted promoters were similarly isolated from genomic DNA using oligonucleotides with restriction sites at either the 5′ or 3′ end for later incorporation into plasmids. These promoters were initially selected due to their potential role in affecting time keeping elements, if such elements are present in *H. volcanii*. Depending on the promoter regulating CBG MutB, transformed *H. volcanii* would be subjected to various environmental conditions changing over time, such as reactive oxygen species, increased UV radiation, light, and heat. These colonies would be regularly imaged and measured for changes in luminescence; luciferase intensity fluctuated in time with environmental changes, the promoter would likely be responding to the environmental changes. The dual luciferase reporter could be used to determine if the environmental change was directly affecting the luciferin-luciferase reaction by placing the other luciferase under the control of a constitutive promoter and measuring if any changes in luminescence occurred when the environmental change took place.

The CBG MutB reporter has also been transformed into another halophilic archaeon, *Halobacterium salinarum*. Results indicate the luciferase is suitable for experiments in *H. salinarum*; only slight modifications allotting for the differences between *H. volcanii* and *H. salinarum* to the previously described protocols are required, and the luminescent signal intensity is strong. This organism is also used as a model organism for Archaea, and has a variety of genetic tools available. Using a dual luciferase reporter system, *H. salinarum* and *H. volcanii* could be co-cultured to measure growth overlap at different salinities and in different media, and in other competition studies.

```
prot  M V K R E K N V I Y G P E P L H P L E D L T A G E M L F R A L R K

CBG99 ATGGTGAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAG
CBR   ATGGTAAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAG

prot  H S H L P Q A L V D V V G D E S L S Y K E F F E A T V L L A Q S L

CBG99 CACTCTCATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATCTTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTC
CBR   CACTCTCATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATCTTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTC

prot  H N C G Y K M N D V V S I C A E N N T R F F I P V I A A W Y I G M

CBG99 CACAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATG
CBR   CACAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATG

prot  I V A P V N E S Y I P D E L C K V M G I S K P Q I V F T T K N I L

CBG99 ATCGTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTG
CBR   ATCGTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTG

prot  N K V L E V Q S R T N F I K R I I I L D T V E N I H G C E S L P N

CBG99 AACAAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAAT
CBR   AACAAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAAT

prot  F I S R Y S D G N I A N F K P L H F D P V E Q V A A I L C S S G T

CBG99 TTCATCTCTCGCTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCACTTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACT
CBR   TTCATCTCTCGCTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCACTTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACT

prot  T G L P K G V M Q T H Q N I C V R L I H A L D P R V G T Q L I P G
                                                     Y
CBG99 ACTGGACTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCACGCGT-GGGCACTCAGCTGATTCCTGG
CBR   ACTGGACTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCACGC-TACGGCACTCAGCTGATTCCTGG

prot  V T V L V Y L P F F H A F G F S I T L G Y F M V G L R V I M F R R
                                H
CBG99 TGTCACCGTCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTT-AGCATTACTTTGGGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCC
CBR   TGTCACCGTCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTCA-TATTACTTTGGGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCC

prot  F D Q E A F L K A I Q D Y E V R S V I N V P S V I L F L S K S P L

CBG99 TTTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACT
CBR   TTTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACT

prot  V D K Y D L S S L R E L C C G A A P L A K E V A E V A A K R L N L

CBG99 GCGTGGACAAGTACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATC
CBR   GCGTGGACAAGTACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATC

prot  P G I R C G F G L T E S T S A N I H S L G D E F K S G S L G R V T
                                I   Q
CBG99 TTCCAGGGATTCGTTGTGGCTTCGGCCTCACCGAATCTACCAGCGCTAACATTCACTCTCTCGGGGATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCA
CBR   TTCCAGGGATTCGTTGTGGCTTCGGCCTCACCGAATCTACCAGTGCGATTATCCAGACTCTCGGGGATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCA

prot  P L M A A K I A D R E T G K A L G P N Q V G E L C I K G P M V S K

CBG99 CTCCACTCATGGCTGCTAAGATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCGAACCAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCA
CBR   CTCCACTCATGGCTGCTAAGATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCGAACCAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCA

prot  G Y V N N V E A T K E A I D D D G W L H S G D F G Y Y D E D E H F

CBG99 AGGGTTATGTCAATAACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGAGCATT
CBR   AGGGTTATGTCAATAACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGAGCATT
```

```
prot   Y  V  V  D  R  Y  K  E  L  I  K  Y  K  G  S  Q  V  A  P  A  E  L  E  E  I  L  L  K  N  P  C  I  R

CBG99  TTTACGTCGTGGATCGTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGATTCTGTTGAAAAATCCATGCATTC
CBR    TTTACGTCGTGGATCGTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGATTCTGTTGAAAAATCCATGCATTC

prot   D  V  A  V  V  G  I  P  D  L  E  A  G  E  L  P  S  A  F  V  V  K  Q  P  G  K  E  I  T  A  K  E  V
                                                                                T
CBG99  GCGATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTAAAGAAATTACCGCCAAAGAAG
CBR    GCGATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTACAGAAATTACCGCCAAAGAAG

prot   Y  D  Y  L  A  E  R  V  S  H  T  K  Y  L  R  G  G  V  R  F  V  D  S  I  P  R  N  V  T  G  K  I  T

CBG99  TGTATGATTACCTGGCTGAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTCGTAACGTAACAGGCAAAATTA
CBR    TGTATGATTACCTGGCTGAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTCGTAACGTAACAGGCAAAATTA

prot   R  K  E  L  L  K  Q  L  L  E  K  A  G  G  *

CBG99  CCCGCAAGGAGCTGTTGAAACAATTGTTGGAGAAGGCCGGCGGTTAG
CBR    CCCGCAAGGAGCTGTTGAAACAATTGTTGGTGAAGGCCGGCGGTTAG
```

Figure 5.1. Sequences of luciferases to compare CBR and CBG. Three items are shown. 1) The lines labeled "prot" show the amino acid sequence of the commercially available CBG99 gene. 2) Immediately below the "prot" lines are the amino acid substitutions (if any) that result from the mutations. 3) Lines labeled "CBG99" show the nucleotide sequence for CBG99. 4) Lines labeled "CBR" show the nucleotide sequence for CBR luciferase.. Differences from the CBG99 parent are highlighted in black.

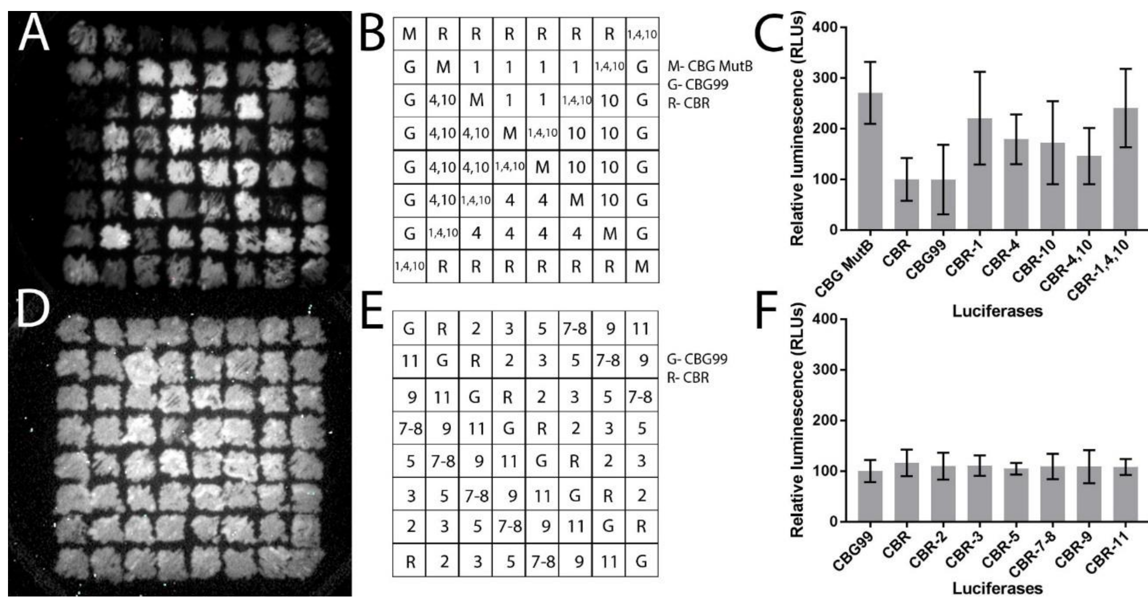


Figure 5.2. Effect of single point mutations in CBR on relative luminescence compared to CBG99 and CBG MutB. (A) Site directed mutagenesis of CBR was conducted in order to induce one of the eleven point mutations (labeled CBR-X, where X is the mutation point) seen in CBG MutB. Mutations 7 and 8 sit eight bases apart, and are too close to one another to separate a single mutagenesis event of either point. No single mutation is significantly brighter than CBG99 or CBR, though a combination of mutation points 1, 4, and 10 appears brighter. (B) A map denoting the identity of each patch in A. (C) The luminescence intensity of each luciferase was recorded and quantified using ImageJ software. (D) *H. volcanii* transformed with various luciferases driven by P_{tnaA} were patched diagonally onto 18% salt w/v Hv-YPC solid medium and imaged (5 min exposure) after 48 h growth. (E) A map denoting the identity of each patch in D. Conv. 1,4,10 is a luciferase similar to CBG99 except that it contains mutations 1, 4, and 10 of CBG MutB (F) The luminescence intensity of each luciferase was recorded and quantified using ImageJ software. Signal background for each plate was averaged from

10 locations on the plate were *H. volcanii* was not patched and subtracted from all luminescence measurements. All bars show means ± standard deviation.

\

REFERENCES

Allers, T., and Ngo, H.-P. (2003). Genetic analysis of homologous recombination in archaea: *Haloferax volcanii* as a model organism. Biochem. Soc. Trans. *31*, 706–710.

Allers, T., Ngo, H.-P., Mevarech, M., and Lloyd, R.G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. Appl Environ Microbiol *70*, 943–953.

Allers, T., Barak, S., Liddell, S., Wardell, K., and Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of his-tagged proteins in *Haloferax volcanii*. Appl Environ Microbiol *76*, 1759–1769.

Amann, E., Brosius, J., and Ptashne, M. (1983). Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene *25*, 167–178.

Auguet, J.-C., Barberan, A., and Casamayor, E.O. (2010). Global ecological patterns in uncultured archaea. ISME J *4*, 182–190.

Auld, D.S., Thorne, N., Maguire, W.F., and Inglese, J. (2009). Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. Proc. Natl. Acad. Sci. U.S.A. *106*, 3585–3590.

Bitan-Banin, G., Ortenberg, R., and Mevarech, M. (2003). Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the pyrE Gene. J Bacteriol *185*, 772–778.

Brendel, J., Stoll, B., Lange, S.J., Sharma, K., Lenz, C., Stachler, A.-E., Maier, L.-K., Richter, H., Nickel, L., Schmitz, R.A., et al. (2014). A complex of Cas proteins 5, 6, and 7 is required for the biogenesis and stability of clustered regularly interspaced short palindromic repeats (CRISPR)-derived RNAs (crRNAs) in Haloferax volcanii. J Biol Chem *289*, 7164–7177.

Breuert, S., Allers, T., Spohn, G., and Soppa, J. (2006). Regulated polyploidy in halophilic archaea. PLOS ONE *1*, e92.

Britton, K.L., Stillman, T.J., Yip, K.S., Forterre, P., Engel, P.C., and Rice, D.W. (1998). Insights into the molecular basis of salt tolerance from the study of glutamate dehydrogenase from *Halobacterium salinarum*. J. Biol. Chem. *273*, 9023–9030.

Carr, A.M., and Lambert, S. (2013). Replication Stress-Induced Genome Instability: The dark side of replication maintenance by homologous recombination. Journal of Molecular Biology *425*, 4733–4744.

Cass, S.D.B., Haas, K.A., Stoll, B., Alkhnbashi, O.S., Sharma, K., Urlaub, H., Backofen, R., Marchfelder, A., and Bolt, E.L. (2015). The role of Cas8 in type I CRISPR interference. Biosci. Rep. *35(3),* e00197.

Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. Int. J. Syst. Evol. Microbiol. *52*, 297–354.

Charpentier, E., Richter, H., van der Oost, J., and White, M.F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol Rev *39*, 428–441.

Cirino, P.C., Mayer, K.M., and Umeno, D. (2003). Generating mutant libraries using error-prone PCR. Methods Mol. Biol. *231*, 3–9.

Cline, S.W., Schalkwyk, L.C., and Doolittle, W.F. (1989). Transformation of the archaebacterium *Halobacterium volcanii* with genomic DNA. J. Bacteriol. *171*, 4987–4991.

Coker, J.A. (2016). Extremophiles and biotechnology: current uses and prospects. F1000Res *5*.

DasSarma, P., Laye, V.J., Harvey, J., Reid, C., Shultz, J., Yarborough, A., Lamb, A., Koske-Phillips, A., Herbst, A., Molina, F., et al. (2017). Survival of halophilic archaea in Earth's cold stratosphere. International Journal of Astrobiology *16*, 321–327.

DasSarma, S., Capes, M., and DasSarma, P. (2009). Haloarchaeal megaplasmids. In Microbial Megaplasmids, (Springer, Berlin, Heidelberg), pp. 3–30.

DasSarma, S., Karan, R., DasSarma, P., Barnes, S., Ekulona, F., and Smith, B. (2013). An improved genetic system for bioengineering buoyant gas vesicle nanoparticles from Haloarchaea. BMC Biotechnol *13*, 112.

Delmas, S., Shunburne, L., Ngo, H.-P., and Allers, T. (2009). Mre11-Rad50 Promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. PLOS Genetics *5*, e1000552.

Emerson, J.B., Andrade, K., Thomas, B.C., Norman, A., Allen, E.E., Heidelberg, K.B., and Banfield, J.F. (2013). Virus-host and CRISPR dynamics in archaea-dominated hypersaline Lake Tyrrell, Victoria, Australia. Archaea *2013*, 370871.

Fineran, P.C., and Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. Virology *434*, 202–209.

Fischer, S., Maier, L.-K., Stoll, B., Brendel, J., Fischer, E., Pfeiffer, F., Dyall-Smith, M., and Marchfelder, A. (2012). An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to Target Invader DNA. J. Biol. Chem. *287*, 33351–33363.

Giroux, X., and MacNeill, S.A. (2015). Molecular Genetic methods to study DNA replication protein function in *Haloferax volcanii*, a model archaeal organism. Methods Mol. Biol. *1300*, 187–218.

Gophna, U., Allers, T., and Marchfelder, A. (2017). Finally, Archaea get their CRISPR-Cas toolbox. Trends in Microbiology *25*, 430–432.

Gupta, R.S. (1998). Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. Microbiol. Mol. Biol. Rev. *62*, 1435–1491.

Gupta, R.S. (2000). The natural evolutionary relationships among prokaryotes. Critical Reviews in Microbiology *26*, 111–131.

Hartman, A.L., Norais, C., Badger, J.H., Delmas, S., Haldenby, S., Madupu, R., Robinson, J., Khouri, H., Ren, Q., Lowe, T.M., et al. (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. PLoS One *5(3),* e9605.

Hastings, J.W., and Johnson, C.H. (2003). Bioluminescence and chemiluminescence. Meth. Enzymol. *360*, 75–104.

Hastings, J.W., McElroy, W.D., and Coulombre, J. (1953). The effect of oxygen upon the immobilization reaction in firefly luminescence. J. Cell. Comp. Physiol. *42*, 137–150.

Hochachka, P.W., and Somero, G.N. (2002). Biochemical adaptation: mechanism and process in physiological evolution (Oxford University Press).

Holmes, M., Pfeifer, F., and Dyall-Smith, M. (1994). Improved shuttle vectors for *Haloferax volcanii* including a dual-resistance plasmid. Gene *146*, 117–121.

Holmes, M.L., Nuttall, S.D., and Dyall-Smith, M.L. (1991). Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase. J. Bacteriol. *173*, 3807–3813.

Hovmöller, S., and Zhou, T. (2004). Why are both ends of the polypeptide chain on the outside of proteins? Proteins *55*, 219–222.

Ignowski, J.M., and Schaffer, D.V. (2004). Kinetic analysis and modeling of firefly luciferase as a quantitative reporter gene in live mammalian cells. Biotechnol. Bioeng. *86*, 827–834.

Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. Eur. J. Biochem. *202*, 715–728.

Jain, S., Caforio, A., and Driessen, A.J.M. (2014). Biosynthesis of archaeal membrane ether lipids. Front Microbiol *5*, 641.

Jore, M.M., Lundgren, M., Duijn, E. van, Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, Ü., Wurm, R., Wagner, R., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by cascade. Nature Structural and Molecular Biology *18*, nsmb.2019.

Kapatai, G., Large, A., Benesch, J.L.P., Robinson, C.V., Carrascosa, J.L., Valpuesta, J.M., Gowrinathan, P., and Lund, P.A. (2006). All three chaperonin genes in the archaeon *Haloferax volcanii* are individually dispensable. Molecular Microbiology *61*, 1583–1597.

Khelaifia, S., Lagier, J.-C., Nkamga, V.D., Guilhot, E., Drancourt, M., and Raoult, D. (2016). Aerobic culture of methanogenic archaea without an external source of hydrogen. Eur. J. Clin. Microbiol. Infect. Dis. *35*, 985–991.

Kiljunen, S., Pajunen, M.I., Dilks, K., Storf, S., Pohlschroder, M., and Savilahti, H. (2014). Generation of comprehensive transposon insertion mutant library for the model archaeon, *Haloferax volcanii*, and its use for gene discovery. BMC Biology *12*, 103.

Koga, Y. (2012). Thermal adaptation of the archaeal and bacterial lipid membranes. Archaea *2012*, 789652.

Kondo, T., Strayer, C.A., Kulkarni, R.D., Taylor, W., Ishiura, M., Golden, S.S., and Johnson, C.H. (1993). Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. Proc. Natl. Acad. Sci. U.S.A. *90*, 5672–5676.

Koonin, E.V., Mushegian, A.R., Galperin, M.Y., and Walker, D.R. (1997). Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. Molecular Microbiology *25*, 619–637.

Krishnamoorthy, A., and Robertson, J.B. (2015). Dual-color monitoring overcomes the limitations of single bioluminescent reporters in fast-growing microbes and reveals Phase-dependent protein productivity during the metabolic rhythms of *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. *81*, 6484–6495.

Lagier, J.-C., Khelaifia, S., Alou, M.T., Ndongo, S., Dione, N., Hugon, P., Caputo, A., Cadoret, F., Traore, S.I., Seck, E.H., et al. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. Nature Microbiology *1*, 16203.

Lange, C., Zerulla, K., Breuert, S., and Soppa, J. (2011). Gene conversion results in the equalization of genome copies in the polyploid haloarchaeon *Haloferax volcanii*. Molecular Microbiology *80*, 666–677.

Lanyi, J.K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriol Rev *38*, 272–290.

Large, A., Stamme, C., Lange, C., Duan, Z., Allers, T., Soppa, J., and Lund, P.A. (2007). Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene. Molecular Microbiology *66*, 1092–1106.

Leigh, J.A., Albers, S.-V., Atomi, H., and Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. FEMS Microbiol Rev *35*, 577–608.

Li, M., Wang, R., Zhao, D., and Xiang, H. (2014). Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. Nucleic Acids Res *42*, 2483–2492.

Li, M., Gong, L., Zhao, D., Zhou, J., and Xiang, H. (2017). The spacer size of I-B CRISPR is modulated by the terminal sequence of the protospacer. Nucleic Acids Res *45*, 4642–4654.

Littlechild, J.A. (2015). Archaeal Enzymes and Applications in Industrial Biocatalysts. Archaea *2015*.

Lozier, R.H., Bogomolni, R.A., and Stoeckenius, W. (1975). Bacteriorhodopsin: a light-driven proton pump in *Halobacterium halobium*. Biophys J *15*, 955–962.

Mackwan, R.R., Carver, G.T., Drake, J.W., and Grogan, D.W. (2007). An Unusual Pattern of Spontaneous Mutations Recovered in the Halophilic Archaeon *Haloferax volcanii*. Genetics *176*, 697–702.

Maier, L.-K., Fischer, S.M., Stoll, B., Brendel, J., Pfeiffer, F., Dyall-Smith, M., and Marchfelder, A. (2012). The immune system of halophilic archaea. Mobile Genetic Elements *2*, 228–232.

Maier, L.-K., Lange, S.J., Stoll, B., Haas, K.A., Fischer, S., Fischer, E., Duchardt-Ferner, E., Wöhnert, J., Backofen, R., and Marchfelder, A. (2013). Essential requirements for the detection and degradation of invaders by the *Haloferax volcanii* CRISPR/Cas system I-B. RNA Biol *10*, 865–874.

Maier, L.-K., Dyall-Smith, M., and Marchfelder, A. (2015). The Adaptive Immune System of *Haloferax volcanii*. Life *5*, 521–537.

Mojica, F. j. m., Ferrer, C., Juez, G., and Rodríguez-Valera, F. (1995). Long stretches of short tandem repeats are present in the largest replicons of the archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. Molecular Microbiology *17*, 85–93.

Moriyama, E.H., Niedre, M.J., Jarvi, M.T., Mocanu, J.D., Moriyama, Y., Subarsky, P., Li, B., Lilge, L.D., and Wilson, B.C. (2008). The influence of hypoxia on bioluminescence in luciferase-transfected gliosarcoma tumor cells in vitro. Photochem. Photobiol. Sci. *7*, 675–680.

Mullakhanbhai, M.F., and Larsen, H. (1975). *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. Arch. Microbiol. *104*, 207–214.

Naor, A., Lapierre, P., Mevarech, M., Papke, R.T., and Gophna, U. (2012). Low species barriers in halophilic archaea and the formation of recombinant hybrids. Curr. Biol. *22*, 1444–1448.

Norais, C., Hawkins, M., Hartman, A.L., Eisen, J.A., Myllykallio, H., and Allers, T. (2007). Genetic and Physical Mapping of DNA Replication Origins in Haloferax volcanii. PLOS Genetics *3*, e77.

Norrander, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene *26*, 101–106.

Oren, A. (1994). The ecology of the extremely halophilic archaea. FEMS Microbiology Reviews *13*, 415–439.

Oren, A. (1999). Microbiological studies in the Dead Sea: future challenges toward the understanding of life at the limit of salt concentrations. Hydrobiologia *405*, 1–9.

Oren, A. (2009). Microbial diversity and microbial abundance in salt-saturated brines: Why are the waters of hypersaline lakes red? Natural Resources and Environmental Issues *15,* Article 49.

Oren, A. (2010). Industrial and environmental applications of halophilic microorganisms. Environ Technol *31*, 825–834.

Paul, S., Bag, S.K., Das, S., Harvill, E.T., and Dutta, C. (2008). Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. Genome Biol. *9*, R70.

Portnoy, V., and Schuster, G. (2006). RNA polyadenylation and degradation in different Archaea; roles of the exosome and RNase R. Nucleic Acids Research *34*, 5923–5931.

Portnoy, V., Evguenieva-Hackenberg, E., Klein, F., Walter, P., Lorentzen, E., Klug, G., and Schuster, G. (2005). RNA polyadenylation in Archaea: not observed in *Haloferax* while the exosome polynucleotidylates RNA in *Sulfolobus*. EMBO Rep *6*, 1188–1193.

Ramanathan, C., Khan, S.K., Kathale, N.D., Xu, H., and Liu, A.C. (2012). Monitoring cell-autonomous circadian clock rhythms of gene expression using luciferase bioluminescence reporters. J Vis Exp *67*, 4234.

Rampelotto, P.H. (2013). Extremophiles and extreme environments. Life *3*, 482–485.

Reed, C.J., Lewis, H., Trejo, E., Winston, V., and Evilia, C. (2013). Protein adaptations in archaeal extremophiles. Archaea *2013*, 373275.

Robertson, J.B., Stowers, C.C., Boczko, E., and Johnson, C.H. (2008). Real-time luminescence monitoring of cell-cycle and respiratory oscillations in yeast. PNAS *105*, 17988–17993.

Rodrigo-Baños, M., Garbayo, I., Vílchez, C., Bonete, M.J., and Martínez-Espinosa, R.M. (2015). Carotenoids from haloarchaea and their potential in biotechnology. Mar Drugs *13*, 5508–5532.

Ronnekleiv, M. (1995). Bacterial carotenoids 53∗ C50-carotenoids 23; carotenoids of *Haloferax volcanii* versus other halophilic bacteria. Biochemical Systematics and Ecology *23*, 627–634.

Rosson, R.A., and Nealson, K.H. (1981). Autoinduction of bacterial bioluminescence in a carbon limited chemostat. Arch. Microbiol. *129*, 299–304.

S. Makarova, K., H. Haft, D., Barrangou, R., J. J. Brouns, S., Charpentier, E., Horvath, P., Moineau, S., J. M. Mojica, F., I. Wolf, Y., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol *9*, 467–477.

Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A., and Severinov, K. (2013). High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in E. coli. RNA Biol *10*, 716–725.

Schiraldi, C., Giuliano, M., and De Rosa, M. (2002). Perspectives on biotechnological applications of archaea. Archaea *1*, 75–86.

Schleper, C., Jurgens, G., and Jonuscheit, M. (2005). Genomic studies of uncultivated archaea. Nature Reviews Microbiology *3*, 479–488.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., Oost, J. van der, Brouns, S.J.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. PNAS *108*, 10098–10103.

Siglioccolo, A., Paiardini, A., Piscitelli, M., and Pascarella, S. (2011). Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. BMC Structural Biology *11*, 50.

Siliakus, M.F., van der Oost, J., and Kengen, S.W.M. (2017). Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. Extremophiles *21*, 651–670.

Snapp, E. (2005). Design and use of fluorescent fusion proteins in cell biology. Curr Protoc Cell Biol *CHAPTER*, Unit-21.4.

Snyder, J.C., Bateson, M.M., Lavin, M., and Young, M.J. (2010). Use of cellular CRISPR (clusters of regularly interspaced short palindromic repeats) spacer-based microarrays for detection of viruses in environmental samples. Appl. Environ. Microbiol. *76*, 7251–7258.

Sokalingam, S., Raghunathan, G., Soundrarajan, N., and Lee, S.-G. (2012). A study on the effect of surface lysine to arginine mutagenesis on protein stability and structure using green fluorescent protein. PLOS ONE *7*, e40410.

Solgi, Z., Khalifeh, K., Hosseinkhani, S., and Ranjbar, B. (2016). Surface arginine saturation effect on unfolding reaction of firefly luciferase: a thermodynamic and kinetic perspective. Photochem Photobiol *92*, 688–693.

Soppa, J. (2013). Evolutionary advantages of polyploidy in halophilic archaea. Biochemical Society Transactions *41*, 339–343.

Springer, P.S. (2000). Gene Traps. Plant Cell *12*, 1007–1020.

Stachler, A.-E., and Marchfelder, A. (2016). Gene Repression in Haloarchaea Using the CRISPR (clustered regularly interspaced short palindromic repeats)-Cas I-B system. J. Biol. Chem. *291*, 15226–15242.

Stan-Lotter, H., and Fendrihan, S. (2011). Adaption of microbial life to environmental extremes: novel research results and application (Wien ; New York: Springer).

Sternberg, S.H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas systems. Molecular Cell *61*, 797–808.

Stetter, K.O. (2006). Hyperthermophiles in the history of life. Philos Trans R Soc Lond B Biol Sci *361*, 1837–1843.

Stoll, B., Maier, L.-K., Lange, S.J., Brendel, J., Fischer, S., Backofen, R., and Marchfelder, A. (2013). Requirements for a successful defence reaction by the CRISPR–Cas subtype I-B system. Biochemical Society Transactions *41*, 1444–1448.

Tadeo, X., López-Méndez, B., Trigueros, T., Laín, A., Castaño, D., and Millet, O. (2009a). Structural basis for the aminoacid composition of proteins from halophilic archea. PLoS Biol. *7*, e1000257.

Tadeo, X., López-Méndez, B., Trigueros, T., Laín, A., Castaño, D., and Millet, O. (2009b). Structural basis for the aminoacid composition of proteins from halophilic archaea. PLOS Biology *7*, e1000257.

Torreblanca, M., Rodriguez-Valera, F., Juez, G., Ventosa, A., Kamekura, M., and Kates, M. (1986). Classification of non-alkaliphilic halobacteria based on numerical taxonomy and polar lipid Composition, and description of *Haloarcula* gen. nov. and *Haloferax* gen. nov. Systematic and Applied Microbiology *8*, 89–99.

W Stoeckenius, and Bogomolni, and R.A. (1982). Bacteriorhodopsin and related pigments of halobacteria. Annual Review of Biochemistry *51*, 587–616.

Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., and Wang, Y. (2015). Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. Cell *163*, 840–853.

de Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R., and Subramani, S. (1987). Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. *7*, 725–737.

Whitehead, K., Pan, M., Masumura, K., Bonneau, R., and Baliga, N.S. (2009). Diurnally entrained anticipatory behavior in Archaea. PLOS ONE *4*, e5485.

Wiedenheft, B., van Duijn, E., Bultema, J.B., Bultema, J., Waghmare, S.P., Waghmare, S., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J.R., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. U.S.A. *108*, 10092–10097.

Williams, D., Gogarten, J.P., and Papke, R.T. (2012). Quantifying homologous replacement of loci between haloarchaeal species. Genome Biol Evol *4*, 1223–1244.

Woese, C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. PNAS *74*, 5088–5090.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eukarya. Proc. Natl. Acad. Sci. U.S.A. *87*, 4576–4579.

Wu, X., Altman, R., Eiteman, M.A., and Altman, E. (2014). Adaptation of *Escherichia coli* to elevated sodium concentrations increases cation tolerance and enables greater lactic acid production. Appl. Environ. Microbiol. *80*, 2880–2888.

Xiao, Y., Luo, M., Hayes, R.P., Kim, J., Ng, S., Ding, F., Liao, M., and Ke, A. (2017). Structure basis for directional R-loop formation and substrate handover mechanisms in type I CRISPR-Cas system. Cell *170*, 48–60.e11.

Yin, Y., Williamson, B.D., and Rutherford, C.L. (1994). An autonomously propagating luciferase-encoding vector for *Dictyostelium discoideum*. Gene *150*, 293–298.

Zaigler, A., Schuster, S.C., and Soppa, J. (2003). Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon *Haloferax volcanii*. Molecular Microbiology *48*, 1089–1105.