ABSTRACT

Title of Document: DNA MISMATCH REPAIR AND RESPONSE TO OXIDATIVE STRESS IN THE EXTREMELY HALOPHILIC ARCHAEON *HALOBACTERIUM* SP. STRAIN NRC-1

Courtney Rae Busch, Doctor of Philosophy, 2008

Directed By: Assistant Professor Dr. Jocelyne DiRuggiero, Department of Cell Biology and Molecular Genetics

Halobacterium is an extremely halophilic archaeon that has homologs of the key proteins, MutS and MutL used in DNA mismatch repair in both Bacteria and Eukarya. To determine whether *Halobacterium* has a functional mismatch repair system, we calculated the spontaneous mutation rate and determined the spectrum of mutation in *Halobacterium* using fluctuation tests targeting genes of the UMP biosynthesis pathway and we performed a sequence analysis of the mutated genes. We found that *Halobacterium* has a low incidence of mutation indicating that some form of DNA repair is taking place, however the mutational spectrum in the Archaea is different from that seen in Bacteria and Eukarya suggesting differences between the archaeal, bacterial, and eukaryal repair systems. To test if the MutS and MutL homologs in *Halobacterium* are essential for the low incidence of mutation, we used in-frame targeted gene deletion and characterized the mutant phenotypes. We found

no phenotypic differences between the mutant strains and the background strain indicating that the MutS and MutL protein homologs found in *Halobacterium* are not essential for maintaining the low incidence of mutation. Since much of the replication and repair processes in *Halobacterium* are similar to that of Eukarya, deciphering how MMR occurs in the Archaea could lead to a new understanding of pathway interactions based on the recruitment of repair enzymes from both bacterial and eukaryal counterparts. In addition, we elucidated the oxidative stress response in *Halobacterium* to hydrogen peroxide and paraquat using a whole genome transcriptional array, in-frame targeted gene deletion, and survival analysis of mutant phenotypes. We showed an overall effort of the cells to scavenge reactive oxygen species and repair damages to the DNA, which has also been seen in response to gamma irradiation. From the mutant analyses, we were able to deduce that Sod1 and PerA proteins played an essential role in removing oxidative stress in *Halobacterium.* Deciphering the stress response to hydrogen peroxide and paraquat in an extreme halophile that lives in an environment subject to long periods of desiccation can further our understanding of the DNA repair and protection systems to oxidative stress in general.

DNA MISMATCH REPAIR AND RESPONSE TO OXIDATIVE STRESS IN THE EXTREMELY HALOPHILIC ARCHAEON *HALOBACTERIUM* SP. STRAIN NRC-1

By

Courtney Rae Busch

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Dr. Jocelyne DiRuggiero, Chair
Dr. Peggy Hsieh
Dr. Steve Hutcheson
Dr. Carol Keefer
Dr. Zvi Kelman
Dr. Douglas Julin

Dedication

This work is dedicated to Rick, Teri, Candace, and Chris Busch. Without your endless support and love, I would not be where I am today.

Acknowledgements

I would like to acknowledge the many people who joined me on this journey through graduate school. Many thanks goes to my advisor, Dr. Jocelyne DiRuggiero, for her guidance and advice the past five years. Special thanks goes to my committee members: Dr. Peggy Hsieh for putting me in contact with other labs willing to send me constructs for experiments, Dr. Zvi Kelman and Dr. Douglas Julin for allowing me to use their lab equipment and time during my experiments, Dr. Steve Hutcheson for helpful suggestions and advice on experiments, and Dr. Carol Keefer for enthusiastically agreeing to be my Dean's representative.

A special thanks goes to my collaborators in Nitin Baliga's lab at the Institute for Systems Biology in Seattle, Washington for collaboration on the transcriptional response of *Halobacterium* to H_2O_2 and paraquat. Many thanks goes to Dr. Tom Kunkel and Alan Clark of Research Park Triangle, North Carolina, for sending me the phage derivatives and *E. coli* strains as well as helpful hints and protocols used in Appendix B.

I would also like to acknowledge the members, past and present, of the DiRuggiero lab, especially Dr. Adrienne Kish, Courtney Robinson, Danish Ahmed, and Kathryn Flanders for helpful discussions about science and other aspects of life.

Lastly, I would like to acknowledge my family and friends who supported me during my time here at Maryland, the Busch clan both near and far for believing in me and

always being excited about my love of science, Mandy Kendrick for letting me move in with her and being a good listener after a long day at the lab, and the Terrapin Masters swim team for giving me a way to work off my stress.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BER – base excision repair
NER – nucleotide excision repair
MMR – mismatch repair
Halobacterium – *Halobacterium* sp. strain NRC-1
PCNA – proliferating cell nuclear antigen
RPA – replication protein A
RFC – replication factor C
8-GO – 8-oxo-guanine
ROS – reactive oxygen species
5-FOA – 5-fluoroorotic acid
ODCase – orotidine 5'-monophosphate decarboxylase
OPRTase – orotate phosphoribosyl transferase
5-FU – 5-fluorouracil
MNNG – N-methyl-N'-nitro-N-nitrosoguanidine
MGMT – O^6-methyl guanine methyltransferase
BSS – basal salts solution
5-FAA – 5-fluoroanthranilic acid
MIC – minimum inhibitory concentration
UDG – uracil-DNA glycosylase
OD – optical density
H_2O_2 – hydrogen peroxide
UDG – uracil-DNA glycosylase
TDG – thymine-DNA glycosylase
RF – replicative form

Chapter 1: Introduction

DNA damage can result in a variety of mutations, many leading to cell death, making it critical for cells to maintain genomic integrity. Examples of DNA damage that result in mutations include the depurination and deamination of DNA, oxidation and methylation of nucleotides, and thymine dimers caused by UV irradiation [1]. If these are left uncorrected when the DNA is replicated, the mutations can cause deletion of a base pair or a base pair substitution leading to a mismatch. All domains of life have systems in place for repairing these damages, including the repair of mismatched bases. These types of damage can be repaired by the DNA repair pathways present in cells. Depurination, deamination, oxidation, and methylation damage is typically repaired through base excision repair (BER) [1, 2]. BER involves a variety of glycosylases, which recognize a specific type of altered base and catalyze its removal. The nucleotide excision repair (NER) pathway is responsible for removing damage caused by UV irradiation, such as pyrimidine dimers [1].

Despite the different ways mismatched bases can arise, they are predominately caused by DNA replication errors. During replication, DNA polymerase can introduce mismatched nucleotides and insertions or deletions, which can result in base pair and frameshift mutations if left uncorrected. Many DNA polymerases contain proofreading activity, which corrects the mistakes during replication by a 3’ to 5’ exonuclease that removes the incorrect base [1]. Following base excision, replication can continue. If the mismatch is not corrected by polymerase proofreading, the DNA mismatch repair (MMR) pathway can correct these errors.

In both Bacteria and Eukarya, repair of mismatched bases is performed by the highly conserved DNA mismatch repair (MMR) pathway [3, 4]. The MMR pathway is critical for maintaining genome integrity. Defects in the MMR system lead to genomic instability which can cause a 50-1000 fold increase in spontaneous mutability, meiotic defects, and resistance to several DNA damaging agents [4-6]. In humans, inactivation of the MMR pathway leads to simple repeat instability resulting in hereditary nonpolyposis colon cancer [4, 5].

The MMR system is highly conserved between Bacteria and Eukarya but little is known about MMR in the Archaea. One of the characteristics of the domain Archaea is that many are able to survive in extreme environments, which can result in extensive DNA damage, including damage caused by desiccation, solar radiation, extreme temperatures and pH. The archaeal proteins involved in DNA repair are more closely related to the Eukarya but they contain fewer proteins allowing a simplified look into complex eukaryotic repair pathways. There are two kingdoms in the Archaea, the *Crenarchaeota* and the *Euryarchaeota,* of which *Halobacterium* sp. strain NRC-1 (*Halobacterium)* is a member (See Figure 1-1).

While the MMR pathway is involved in several biological processes, this review will focus mainly on the repair of mismatched bases after replication in *Halobacterium.* The mismatch repair homologs, MutS1, MutS2, MutL, and UvrD found in *Halobacterium* will be characterized to evaluate their cellular roles.

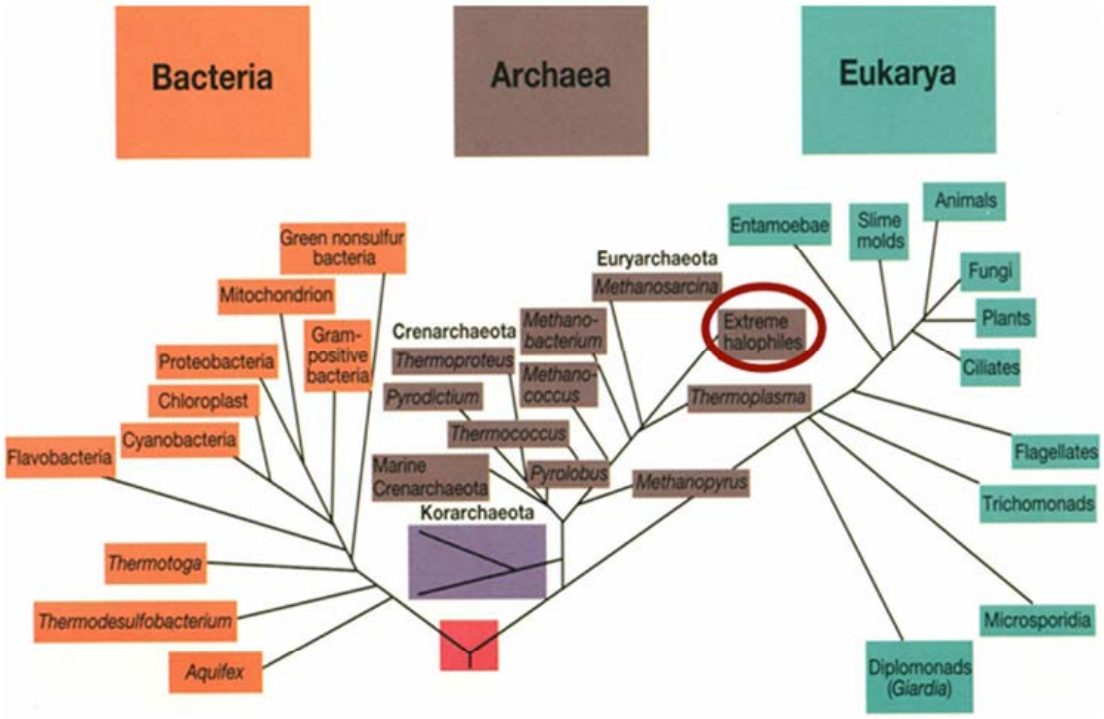


Figure 1-1. Phylogenetic tree of life showing the three domains: Bacteria, Archaea, and Eukarya [7]. The Archaea are divided into three major kingdoms, the *Euryarchaeota* and the *Crenarchaeota* along with the *Korarchaeota* which have only been detected by molecular methods. The branch of the tree where *Halobacterium* is located is circled in red.

1.1 Bacterial MMR

The most characterized MMR pathway in Bacteria is that of *Escherichia coli.* The MMR pathway is responsible for fixing DNA polymerase errors after replication and is critical for maintaining genomic stability. Defects in this pathway can lead to high rates of base substitutions and frameshift mutations and allow recombination between non-homologous sequences [4].

Once DNA polymerase incorporates a mismatched base, the MMR machinery must be directed to the newly synthesized strand in order to remove the base. In *E. coli,* strand discrimination is accomplished by the actions of a GATC sequence specific *dam* methylase, which methylates position 6 on the adenine residue within GATC sequences [8]. MMR occurs on the unmodified strand of the hemimethylated DNA. Cells deficient in *dam* methylase show no strand bias and cells methylated on both strands show no repair [5, 9]. A single hemimethylated GATC sequence is able to direct MMR on either side of the mismatch demonstrating that this pathway is bidirectional. Further evidence supporting GATC methylation directed MMR is the increased rate of spontaneous mutation to streptomycin, rifampin, and valine in cells overproducing *dam* methylase [5, 10]. Further genes involved in MMR were isolated in screens looking for spontaneous mutators. Glickman and Radman isolated mismatch correction deficient mutants by screening for 2-aminopurine, base analog of guanine and adenine, resistant mutants in a *dam* strain of *E. coli* [11]. MutS, MutL, MutH mutants were found to have a 10-1000 fold increase in spontaneous mutation rate [3-5].

Cupples and Miller [12] designed an *in vivo* assay to look at the specificity of MMR for certain mismatches. They constructed six strains of *E. coli* containing a different mutation at the same coding position in the *lacZ* gene. The mutations changed the catalytic amino acid within the *lacZ* gene causing the cells to become Lac^-. Correction of the mismatch led to reversion back to the wildtype codon and was scored by the Lac^+ phenotypes. They used several DNA treatments that create a known mismatch to validate this type of system, which can be useful in detecting new mutator strains.

The development of an *in vitro* assay for MMR led to the underlying mechanisms of methyl-directed MMR [13]. The *in vitro* assay was designed using a heteroduplex from f1 R229 DNA containing a mismatched base within a single restriction site on the duplex. The duplex was incubated with crude cell extracts of *E. coli* with and without the MMR proteins. This experiment elucidated which genes were essential for MMR. Wildtype extracts were able to correct the mismatch, thus restoring the restriction site, whereas *E. coli* extracts of *mutS, mutL, mutH,* and *uvrD* mutants were not. Also implicated in these early experiments were DNA polymerase III, single stranded binding protein, and DNA ligase. Wang and Hays [14] developed an *in vitro* MMR assay to look at correction of mismatches. This assay utilized a double stranded plasmid with two specific endonuclease sites spaced 22 base pairs apart. After cutting by the endonuclease, DNA is removed from the gap and a new piece of DNA can be ligated in. The new piece of DNA has complementary ends to the gapped plasmid and contains a mismatched base within a restriction site. The ability of the restriction enzyme to cut (mismatch corrected) or not (mismatch uncorrected) can be visualized on an agarose gel.

The MMR pathway is highly conserved between Bacteria and Eukarya and one of the key proteins, MutS, has similar structural organization in all three domains of life (See Figure 1-2). The three-dimensional structure of MutS has been resolved in *E. coli* and *Thermus aquaticus* (See Figure 1-3) [15, 16]. It is a 95kDa protein and functions as a dimer *in vivo* [5, 17]. MutS has ATPase activity with Walker A/B sequence motifs and a highly conserved Phe-X-Glu motif responsible for binding DNA [18]. MutS forms a homodimer in bacteria when binding to DNA but the asymmetry of the two subunits bound to the mismatched DNA is similar to that of the MutS heterodimers in the eukaryotes [4]. Crystal structures reveal the two subunits forming a channel in MutS, one of which contains the phenylalanine responsible for binding mismatched DNA with the other subunit contacting the DNA to form a clamp [3]. The C-terminus of MutS contains the helix-turn-helix domain critical for dimerization of the protein [3]. MutL is a 68kDa protein that exists as dimers in solution and is a member of the Bergerat-fold ATPase/kinase family [5, 19]. A precise role for MutL has not been defined but it is known to be essential for MMR. The C-terminus contains the dimerization domain and it has been shown to interact with MutS, MutH, and UvrD [3, 20, 21]. MutH protein plays an essential role in strand discrimination in *E. coli.* It encodes a weak mismatch independent sequence specific endonuclease that cuts 5' to the G in a GATC sequence on the unmethylated strand [4, 5]. This nick can occur either 3' or 5' to the mismatch on the unmethylated strand and the ensuing strain break serves as the signal that directs mismatch repair [5, 22, 23]. MutH is activated by MutS, MutL, ATP, and Mg^{2+}, which increased the endonuclease activity 20-70-fold [23, 24].

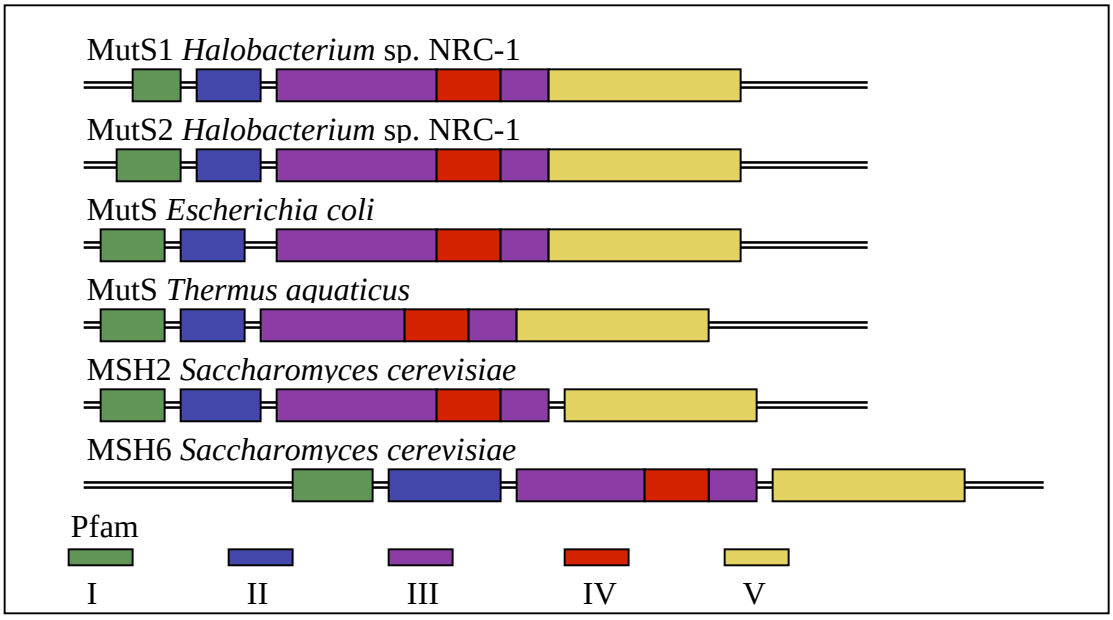


Figure 1-2. Comparison of domain organization for bacterial, eukaryal, and *Halobacterium* MutS proteins [25]. Domains I and IV are involved in DNA binding, domain V is the ATPase and dimerization domain, and domains II and III are connecting domains.

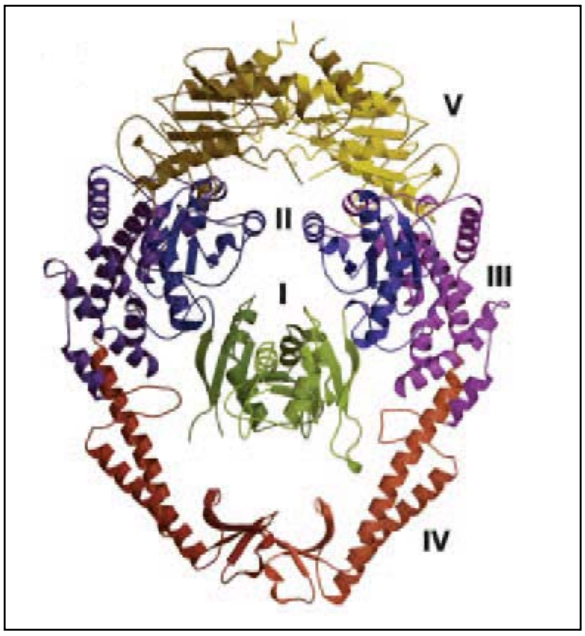


Figure 1-3. *Thermus aquaticus* MutS protein structure modified from [16] showing the ring like structure formed by the two subunits. Ribbon diagram of two subunits of MutS, the five domains are colored green, blue, purple, red, and yellow from N to C terminus. Domains I and IV are involved in DNA binding, domain V is the ATPase and dimerization domain, domains II and III connect the DNA binding domains to the dimerization domain.

A recent analysis of the bacterial MutS homologs showed that they can be grouped into 4 different subfamilies in contrast to two families as previously thought [6]. MutS1 subfamily proteins are the stereotypical MMR MutS1 homologs. They contain four conserved domains including the domains responsible for dimerization, ATPase, and DNA binding activities [15, 16]. This family is widespread among bacterial species. The MutS2 subfamily proteins are also found in many bacterial species but only contain two of the conserved domains including the ones involved in dimerization, ATPase, and DNA binding activities. They have a unique extended C-terminus containing a small MutS related domain that is highly conserved among all MutS2 subfamily proteins. It is hypothesized that this protein may play a role in MMR through the interaction with MutS1 subfamily proteins [26]. MutS3 and MutS4 subfamily proteins are only found in a few distantly related bacterial species. Due to an apparent gene duplication event, most contain two copies of this gene. This family contains the dimerization, ATPase, and DNA binding domains and several species also have one of the other conserved domains. While the biological relevance of MutS2, MutS3, and MutS4 subfamilies are not known, a functional role in repair cannot be ruled out. Studies in *Helicobacter pylori* indicate that they do not function in MMR but play a role in controlling homologous recombination. Genetic studies show that *H. pylori* MutS suppresses homologous and homeologous recombination because inactivated MutS leads to an increased incorporation of exogenous DNA [27, 28]. Pinto *et al* also found that *H. pylori* MutS inhibited DNA strand exchange reactions *in vitro* [27].

The details of the MMR system in *E. coli* have been well characterized due to the availability of purifed MutS, MutL, and MutH proteins and strains deficient in any of the MMR proteins are unable to perform MMR [3, 4]. There are three basic MMR steps: (1) recognition of the mismatch by MutS/MutL; (2) excision of the mismatched base and the DNA surrounding it with MutH endonuclease, UvrD helicase, and 3'-5' or 5'-3' exonucleases; and (3) repair synthesis by DNA polymerase III and a ligase (See Figure 1-4). MutS initiates MMR by recognizing the mismatched base through the highly conserved Phe-X-Glu DNA binding site and recruiting MutL in an ATP-dependent fashion. This MutS/MutL complex activates several downstream activities including MutH, a 25kDa endonuclease [5]. MutH will incise the unmethylated GATC sequence 3' or 5' to the mismatch and create a single strand break [3, 5]. This single strand break is the signal that directs excision repair. The 3' to 5' helicase, UvrD, is loaded on the strand break in an orientation dependent manner by the MutS/MutL complex [4, 5]. The interaction of MutS and UvrD with MutL results in UvrD being loaded onto the appropriate DNA strand in an iterative manner so unwinding can occur towards the mismatch [5, 20, 29]. While the precise role of MutL is not known, it is thought to be responsible for linking mismatch recognition by MutS to repair activities by MutH and UvrD [3, 5]. Studies have also implicated several exonucleases responsible for removing the mismatched DNA. These include 3' to 5' exonucleases ExoI, ExoVII, and ExoX, and 5' to 3' exonucleases ExoVII and RecJ [4, 5, 30, 31]. These exonucleases will degrade the single stranded DNA formed during unwinding by the helicase until they encounter double stranded DNA [32].

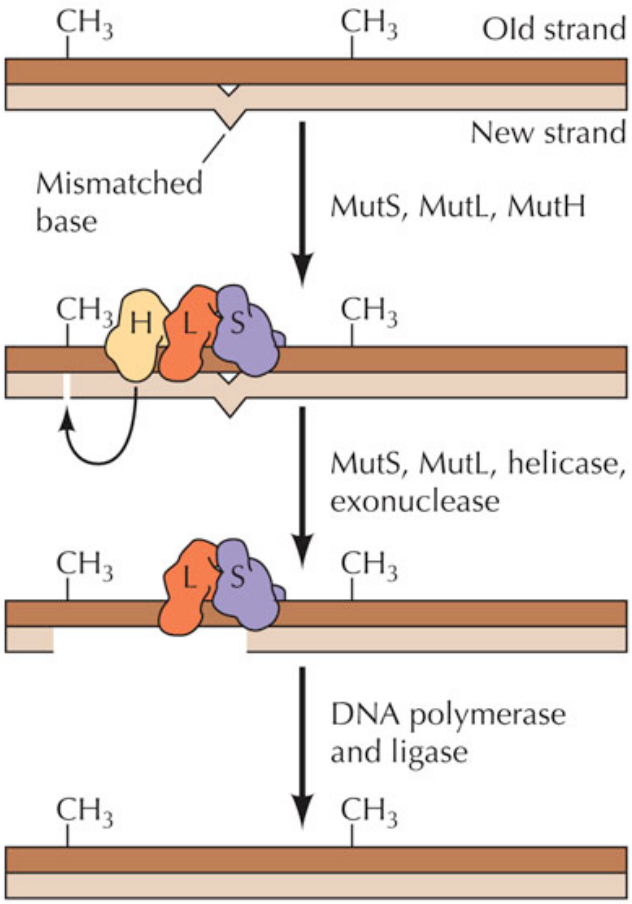


Figure 1-4. Mismatch Repair pathway in *Escherichia coli*. During replication, the DNA is hemimethylated and the polymerase can accidently incorporate a mismatched base into the newly synthesized strand of DNA. The mismatch is recognized by MutS/MutL complex, excised by MutH, RecJ or other exonucleases, and UvrD helicase, and repaired by DNA polymerase III and ligase [33]. Lastly, a *dam* methylase methylates the newly synthesized strand of DNA.

It is important to note that not all bacteria have a methylation-directed MMR system and not all bacteria have a MutH homolog. MutH appears to be an anomaly only found in a few gram negative bacterial species. *Deinococcus radiodurans* only has the MutS/MutL core of the MMR system along with an UvrD homolog yet the pathway is still functional [34]. In the absence of MutH, strand discontinuities have been shown to direct MMR [3, 4]. Strand discontinuities can occur naturally as the 3' terminus on the leading strand or the 3' and 5' ends of Okasaki fragments on the lagging strand [3-5]. One of the major questions in MMR is how strand discrimination is determined in the absence of MutH. Studies are focusing on linking the replication machinery to the MMR machinery.

1.2 Eukaryal MMR

A variety of eukaryotes have homologs of the bacterial MutS and MutL proteins. Many features of MMR are conserved from Bacteria with one major exception. *E. coli* has single MutL and MutS proteins that form homodimers, but eukaryotes have multiple homologs that form heterodimers suggesting a more intricate and complex system with multiple interactions (See Figure 1-5) [3, 4]. Defects in the MMR pathway can lead to a elevated rate of spontaneous mutation, meiotic defects, and resistance to several DNA damaging agents [4-6]. In mammals, inactivation of the MMR pathway can result in microsatellite instability increasing the possibility of hereditary nonpolyposis colon cancer [3-5].

The best characterized MMR system in eukaryotes is that of the yeast *Saccharomyces cerevisiae*. *S. cerevisiae* contains six MutS (MSH1-6) and four MutL (MLH1-3, PMS1) homologs of which only MSH2/MSH6, MSH2/MSH3, and MLH1/PMS1 are involved in

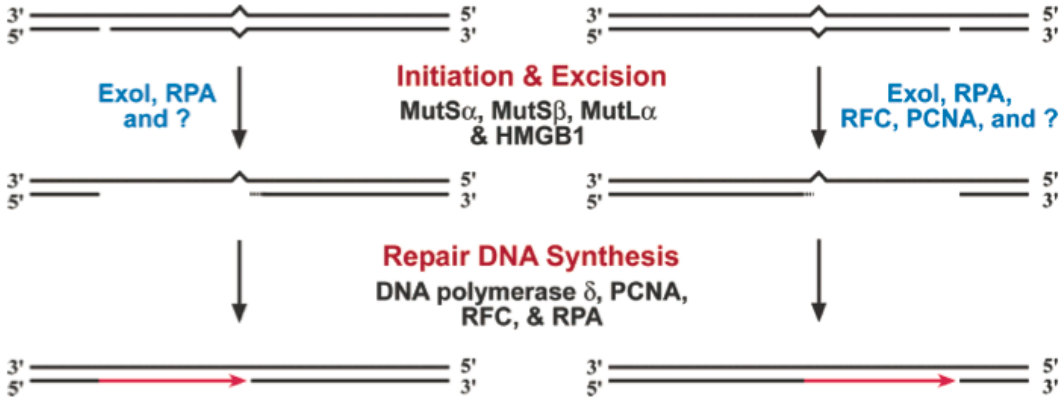


Figure 1-5. Human MMR pathway [5]. MMR pathway is bidirectional *in vitro* and the proteins involved are listed. MutSα is comprised of MutS homologs MSH2/6, MutSβ is comprised of MutS homologs MSH2/3, and MutLα is comprised of MutL homologs MLH1/PMS1. Along with HMGB1, ExoI, and RPA, these proteins are thought to be involved in the initiation and excision of MMR. DNA repair synthesis is completed by DNA polymerase, PCNA, RFC, and RPA. Question marks indicate unidentified proteins.

MMR (See Figure 1-6) [3, 4, 35-38]. Yeast strains deficient in MSH2 display a mutator phenotype and have a 40-fold increase in spontaneous mutation as measured in forward mutation rate assays [39]. MSH1 is required for mitochondrial DNA stability [5]. MSH4 and MSH5 are involved in meiosis for the formation of crossovers and are important in both yeast and mammals [5]. MLH2 and MLH3 also appear to play roles in meiosis [5]. Similarly to *E. coli* MutS, the C-terminal domain of eukaryotic MutS homologs contain the ATP binding and hydrolysis domain responsible for interactions with MSH6 as shown in deletion studies [3, 4, 40]. MSH2/6 binds duplex DNA with base pair mismatches or insertion/deletions loops whereas MSH2/MSH3 only binds to DNA containing insertion deletion loops.

Eukaryotes contain members of the MutS1 and MutS2 subfamilies [6]. The MutS1 subfamily includes the MMR MutS proteins MSH1-6. Originally, MSH4 and MSH5 were classified as part of the MutS2 subfamily based on their divergent sequences and functions but new evidence has shown that they are indeed members of the MutS1 subfamily. Only chloroplast containing species encode members of the MutS2 subfamily and many have multiple copies of these genes.

There are four MutL homologs in *S. cerevisiae* of which PMS1 was the first to be identified based on its mutator phenotype. Along with PMS1, MLH1 plays the most important role in MMR. Yeast strains deficient in MLH1 and PMS1 display a 30-50-fold increase in spontaneous mutation as described in forward mutation rate assays [38, 41].

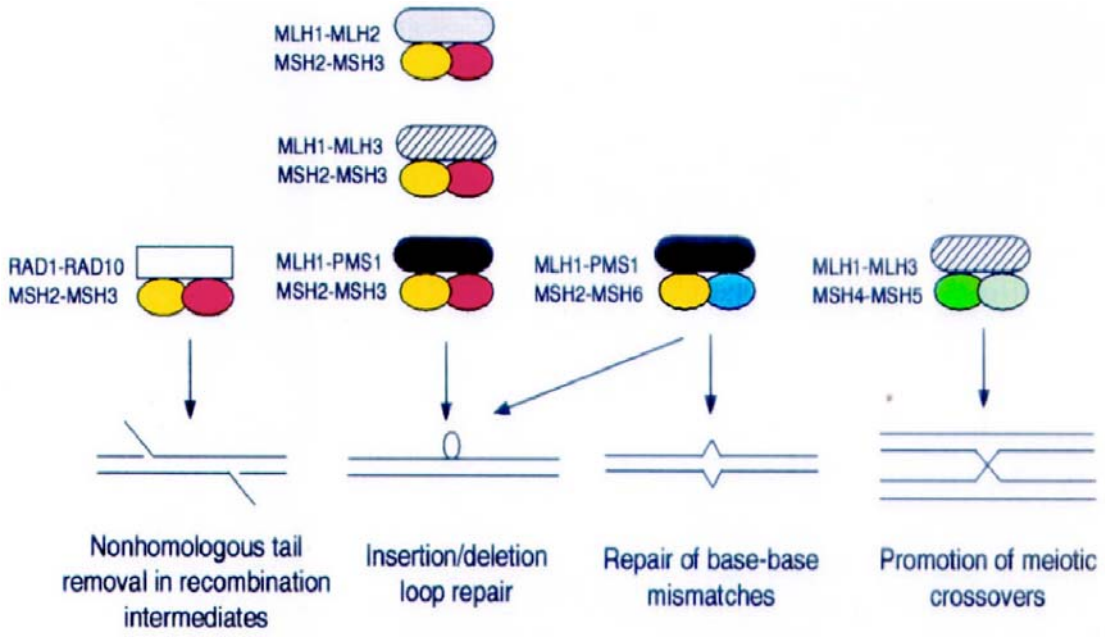


Figure 1-6. The diverse functions of the MMR homologs in *Saccharomyces cerevisiae* [3]. There are 6 MutS homologs, MSH1 is involved in mutation avoidance in the mitchondria, MSH4 and MSH5 are involved in meiosis, and MSH2/MSH3 and MSH2/MSH6 are involved in the repair of base mismatches and frameshift mutations. There are 4 MutL homologs of which MLH1/PMS1 are the major players in MMR.

MLH2 and MLH3 play lesser roles in MMR. The MutL homologs in yeast include a highly conserved N-terminal domain responsible for ATPase activities [4].

Biochemical analyses have shown that the mechanism of Eukaryal MMR functions similarly to that of bacterial MMR with the major exception being strand discrimination mechanisms. Like Bacteria, Eukaryal MMR is thought to be bidirectional (see Figure 1-5) [5]. Human cell extracts along with circular heteroduplexes containing a mismatch and a strand break were used to examine excision tracts [42-44]. With the addition of polymerase inhibitors or the absence of dNTPs, gaps were shown to extend from the strand break to the mismatch regardless of orientation demonstrating that MMR is bidirectional. Further studies in human cell extracts have implicated proliferating cell nuclear antigen (PCNA), a DNA clamp that increases the processivity of the polymerase, as playing a role in MMR. MSH3 and MSH6 both contain a PCNA interaction motif (QXX(LI)XXFF), called a PIP box, in their N-terminus and mutations in this motif can confer a partial mutator phenotype [4]. Four exonucleases are suggested to be involved in MMR: EXO1 and RAD27, which are 5'-3' exonucleases, and the exonuclease subunits of DNA polymerases δ and ε, which are 3'-5' [3, 5, 43, 45, 46]. The most convincing evidence has been shown for ExoI which, in addition to its 5'-3' exonuclease activity, can also function as a 5' flap endonuclease [4, 5]. Using purified human proteins, MSH2, MSH6, MLH1, PMS1, ExoI, and a single stranded DNA binding protein, replication protein A (RPA), MMR occurs exclusively in a 5'-3' direction [47]. With the addition of PCNA and the clamp loader, replication factor C (RFC), MMR becomes bidirectional. It

is possible that PCNA and RFC regulate the directionality of excision by suppressing the 5'-3' capability of ExoI activating 3'-5' excision. Genetic studies in yeast have shown that polymerase δ is the likely polymerase in MMR but it does not rule out a role for polymerases α and ε [5]. In contrast, a helicase has yet to be characterized for this pathway although this could be because of a redundancy in helicases.

Similarly to many bacteria, eukaryotes do not have a MutH homolog and methods of strand discrimination are not clear. *In vitro* and *in vivo* studies have shown that strand discontinuities and nicks/gaps can direct MMR but the natural signal is still a mystery [48]. The general theories of strand discrimination are that it may be nick-directed using Okasaki fragments created during replication of the lagging strand or directed by PCNA thus coupling replication and mismatch repair [3-5, 49-52]. Studies in yeast on mutation rate differences between leading and lagging strands demonstrated that the leading strand has a higher rate of mutations suggesting that the 5' ends of Okasaki fragments along with PCNA can provide the strand discrimination signal [51, 52]. *In vitro* studies of interactions between the MutS homologs and PCNA suggest that MMR and replication may be coupled [5, 48, 53]. MutS homologs contain an interaction motif, Qxx(LI)xxFF, in the N-terminus that is essential for these interactions, and it is hypothesized that PCNA delivers the MutS homologs to the mismatch in the newly synthesized strand of DNA [3, 4]. PCNA is a cofactor for DNA synthesis by polymerase δ. Studies have shown that removal of PCNA prevents 3' directed mismatch excision and limits 5' directed excision in human cell extracts [5]. Although PCNA interacts with multiple MMR proteins, it may

not play an essential role because mutations within the interaction domain only show a moderate increase in mutability [5].

Along with the method of strand discrimination, it is also not known how MutS and MutL homologs can signal the downstream excision events. There are four models currently being studied of different mechanisms to accomplish the signaling of downstream events [5]. The first two models involve searching along the DNA in either an ATP hydrolysis dependent or independent fashion. In the ATP dependent model, MutS and MutL homologs bind to ATP after recognizing a mismatch and then slide along the DNA looking for the strand discrimination signal. In the ATP independent model, MutS and MutL homologs form a sliding clamp that diffuses along the DNA looking for the strand discrimination signal. The third model does not involve ATP hydrolysis but rather the polymerization of MMR components along the DNA between the mismatch and the strand discrimination signal. The last model involves the looping of the DNA to search for the strand discrimination signal. In this model, MutS and MutL homologs stay near the mismatch and the DNA loops around allowing MutS and MutL homologs to search through space for the signal.

Additional repair roles have been attributed to the MMR pathway, such as the repair of a variety of base pair anomalies resulting from DNA damage and preventing non-homologous recombination. Base pair damage is typically repaired by the BER and NER pathways but MMR has been implicated in the repair of DNA damage caused by UV radiation, reactive oxygen species, and alkylating agents. While NER is the predominant

pathway for repairing UV damage, the MMR protein MSH2 interacts with Rad1-3, Rad10, Rad14, and Rad25 in yeast and cells deficient in both NER and MMR show a larger decrease in survival than cells deficient in just one of the pathways [4, 54]. The major oxidative damage is the formation of 8-oxo-guanine (8-GO). This will mispair with an adenine, which the MMR pathway can remove. This is supported by studies in yeast showing interactions between MSH2 or MSH6 and OGG1, a MutM glycosylase homolog, and by the high affinity of MSH2 and MSH6 to 8-GO/A mismatches [3, 55, 56]. The MMR pathway has also been shown to be involved in cellular responses to DNA alkylation damage [3, 4, 57-59]. Alkylation damage is characterized by the generation of an O^6-methylguanine, which will mispair with a thymine. The MMR proteins are unable to correct the damage since the methylated guanine is on the template strand initiating futile cycles of repair. Deactivation of the MMR system allows bypass of this lesion in bacteria and mammals but not in yeast [3, 60]. Heteroduplexes formed during homologous recombination can also be corrected by MMR proteins similarly to postreplicative repair [4, 61, 62]. Studies in yeast have implicated MSH2 and MSH3 along with Rad1 and Rad10 in the removal of non-homologous single strand tails [3, 63, 64].

1.3 Archaeal MMR

The MMR pathway has not been confirmed in the Archaea but there exists striking evidence that the Archaea have a low incidence of mutation [65, 66]. The genomic mutation rate, a measure of genomic stability, has only been measured in two archaea, the thermophilic acidophile *Sulfolobus acidocaldarius* and the halophile *Haloferax volcanii*.

The genomic mutation rate in *S. acidocaldarius* was based on the rate of forward mutation at the *pyrE* gene, an orotatephosphoribosyl transferase [67]. A spectrum of mutation was determined by sequencing 101 mutants and the rate was calculated to be 1.8×10^{-3} per genome per replication, which is close to the genomic rate average for other DNA-based microorganisms [65, 67]. In *H. volcanii,* the genomic mutation rate was also determined by measuring the rate of mutation at the *pyrE2* gene, an orotatephosphoribosyl transferase [68]. Resultant mutants were sequenced and a spectrum of mutation determined. The genomic mutation rate was calculated at 4.5×10^{-4} per genome per replication, which is 7.5-fold lower than the average genomic rate for both bacterial and eukaryotic organisms [65, 68]. This study was done on a much smaller scale than *S. acidocaldarius,* with only 23 mutants sequenced. A larger study could allow more insight into the spectrum of mutation and refine the mutation rate calculation. Nonetheless, efficient DNA repair pathways such as MMR or a decreased amount of mutations, resulting from a high fidelity polymerase, must be present in these organisms for them to maintain such a low spontaneous genomic mutation rate.

The replicative polymerases in the Archaea are members of the B-family and are more similar to their eukaryal counterparts than the bacterial replicative polymerases. There are two types of replicative polymerases in the *Euryarchaeota,* the B-family and the D-family, which both contain strong exonuclease activity suggesting alternative replicative polymerases [69-71]. In contrast, the *Crenarchaeota* only contain members of the B-family of replicative polymerases [69]. Studies in the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus litoralis* looking at DNA polymerase fidelity

demonstrated an approximately 10-fold higher fidelity, 1 x 10^{-6} mutation frequency rate in a forward mutaiton assay, than that of the bacterial *T. aquaticus* polymerase, 2 x 10^{-5} mutation frequency rate, which does not contain a proofreading exonuclease [72-75]. The DNA polymerase III holoenzyme in *E. coli* has a fidelity of approximately 5 x 10^{-6}, which is 5 times lower than that of the hyperthermophilic archaea [76]. It is possible that in the absence of MMR homologs, such as in most thermophilic archaea including *S. acidocaldarius,* a specific pathway is present to correct these mutations or that other known DNA repair proteins are playing that role along with a higher fidelity polymerase resulting in a decreased amount of mutation [77]. In the Archaea with MMR homologs, such as *H. volcanii,* they could either correct mismatches via the bacterial-like pathway, an archaeal-specific pathway, or a combination of the two.

Only eleven archaeal genomes out of the 49 sequenced to date contain homologs of the MutS1 protein subfamily found in bacteria and eukaryotes [6]. These homologs are mainly based on protein sequence comparisons since the cellular and biochemical roles of archaeal MutS proteins have not been investigated. Archaea with MutS1 homologs include halophiles and methanogens, all part of the domain *Euryarchaeota*. These MutS proteins share identical domain structure with their bacterial counterparts likely due to a lateral gene transfer event (See Figure 1-2). Two of the thermophilic archaea, *Ferroplasma acidarmanus* and *Thermoplasma volcanium,* encode members of the MutS4 subfamily whose function is not known. Also detected in the Archaea is a novel subfamily erroneously classified as MutS2 subfamily, which was renamed as the MutS5 subfamily based on phylogenetic analysis [6]. Fourteen Archaea, including the

hyperthermophile, *Pyrococcus furiosus,* have these MutS5 family genes. In *P. furiosus,* this protein has been shown to have ATPase and DNA binding activity but no specific mismatch binding activity [78]. The other subfamilies are not similar in sequence, except for the ATPase domain, to the MutS1 subfamily proteins.

1.4 *Halobacterium* sp. strain NRC-1

Halobacterium is a member of the *Euryarchaeota,* one of the two kingdoms in the domain Archaea (see Figure 1-1). *Halobacterium* is a good model system for studying DNA damage repair due to the presence of genetic tools such as shuttle vectors and targeted gene replacement systems that are not found in many other archaea. *Halobacterium* has a fully sequenced genome which includes a major chromosome and two mini-chromosomes all of which are GC-rich [79]. *Halobacterium* cells contain multiple copies of the major chromosome and the two mini-chromosomes averaging between 15-25 copies depending on growth phase and exhibit a 6-8 hour doubling time [80]. Genetic systems readily available in *Halobacterium* allow gene expression patterns and gene regulation studies in response to various DNA damages. *Halobacterium* is found in hypersaline environments characterized by elevated temperatures, dessicating conditions leading to cycles of rehydration and dessication, and differing concentrations of oxygen and nutrients [81-83]. To maintain osmotic balance with the external hypersaline environment *Halobacterium* cells contain a high intracellular salt environment. *Halobacterium* has also been shown to be highly resistant to dessication and UV-C and gamma irradiation [84-86].

The *Halobacterium* genome contains homologs to genes present in many eukaryotic and bacterial DNA repair pathways including NER, BER, MMR, photoreactivation, and recombinational repair [79]. Proteins from these pathways were identified by genome comparison and by biochemical characterization of proteins from other archaea. Only three studies of DNA repair genes have been characterized in the halophilic archaea: construction of a *radA* mutant to study homologous recombination in *H. volcanii,* the *rad50* and *mre11* genes to study homologous recombination in *Halobacterium,* and the *uvrA/B/C* mutants in *Halobacterium* involved in NER [87-89]. *Halobacterium* contains proteins homologous to both bacterial and eukaryal NER proteins. Deletion mutant analyses of the bacterial homologs, *uvrA*, *uvrB,* and *uvrC,* resulted in no survival to UV-C radiation demonstrating that these genes are essential for NER [88]. Similarly to MMR, homologs of the bacterial *uvr* genes have not been identified in many archaea, however alternate methods must be employed to remove DNA damage caused by UV-C radiation and other DNA damaging treatments. Genomic sequencing of *Sulfolobus solfataricus* reveals the presence of homologs to the eukaryal NER proteins and studies suggest that these homologs may perform NER in this organism [90]. Even though these pathways are still putative and many key proteins are missing, the presence of a mixture of eukaryal-like and bacterial-like DNA repair pathways makes *Halobacterium* a good model system to study repair mechanisms in the Archaea.

Homologs of the bacterial MutS and MutL proteins have been found in the genome of *Halobacterium*. Through computational analysis we have found that *Halobacterium* has a *zim* gene, which encodes a CTAG methylase, 3 bacterial-like *mutS* genes, a bacterial-like

mutL gene, 4 bacterial-like *recJ* exonuclease genes, 1 eukaryotic-like *rad2* 5'-3' exonuclease, and a bacterial-like *uvrD* helicase (See Table 1-1). The *rad2* exonuclease is homologous to the Exo1 protein in yeast and humans [91]. *Halobacterium* MutS1, MutS2, and MutS3 proteins are members of the MutS1 subfamily although MutS3 protein is not homologous to the MMR MutS protein in Bacteria [6]. The function of MutS3 is not known but studies in *H. pylori* suggest a role in homologous recombination.

The protein sequences of MutS1 and MutS2 in *Halobacterium* are 43% identical to one another and are more closely related to bacterial MutS than the eukaryal homologs. They share 39-44% similarity with *E. coli* and *T. aquaticus* but only 21-22% with *S. cerevisiae.*The domain organization of MutS1 and MutS2 in *Halobacterium* is similar that that of other MutS proteins (See Figure 1-2 and Figure 1-7). There are 5 domains. Domains I and IV are involved in DNA binding and domain V contains the ATPase activity. This ATPase domain also contains the 4 nucleotide binding sites found in *E. coli* and *T. aquaticus* as well as the helix-u-turn-helix (HuH) motif which is essential for MutS dimerization, mismatch binding, and ATP hydrolysis [15, 16, 92]. Also conserved is the Phe-X-Glu (Phe36 in *E. coli*) positioned in domain I. This is required for binding of MutS to DNA mismatches and substitutions at this position render the enzyme defective for MMR *in vivo* [93]. One thing of significance to note is that the PCNA binding motif located at the N-terminus of yeast MSH3 and MSH6 is not found in *Halobacterium*. This is important for the functional role of MutS1 and MutS2 since the replication machinery of *Halobacterium* is similar to that of eukaryotes and includes PCNA [53, 94]. PCNA confers processivity to the DNA polymerase by acting as a sliding clamp. The

Table 1-1. MMR protein homologs from the three domains of life with *Halobacterium* gene numbers.

Function	*S. cerevisiae*	*E. coli*	*Halobacterium*
Mismatch recognition	MSH2/3/6	MutS	MutS1 (VNG0163G) MutS2 (VNG0172G)
Meiosis/Unknown	MSH4/5		MutS3 (VNG2270G)
Binds MutS homologs	MLH1 PMS1	MutL	MutL (VNG0159G)
Endonuclease		MutH	
Exonuclease 3'-5'	DNA polymerases δ and ε	RecJ ExoVII	RecJ (VNG0650G, VNG2333C, VNG2519H, VNG6183C)
Exonuclease 5'-3'	Exo1 Rad27	ExoI ExoX	Rad2 (VNG1359G)
DNA helicase		UvrD	UvrD (VNG2620G)
GATC methylase		Dam	

```
HaloMutS1       ------------------------------------------------------------
HaloMutS2       ------------------------------------------------------------
EcoliMutS       ------------------------------------------------------------
ScerMSH2        ------------------------------------------------------------
ScerMSH6        MAPATRKTSKTAHFENGSTSSQKKMKQSSLLSFFSKQVPSGTPSKKVQKPTPATLENTAT 60
                                          ::::::::::::::::::
HaloMutS1       ------------------------------------------------------------
HaloMutS2       ------------------------------------------------------------
EcoliMutS       ------------------------------------------------------------
ScerMSH2        ------------------------------------------------------------
ScerMSH6        DKITKNPQGGKTGKLFVDVDEDNDLTIAEETVSTVRSDIMHSQEPQSDTMLNSNTTEPKS 120

HaloMutS1       ------------------------------------------------------------
HaloMutS2       ------------------------------------------------------------
EcoliMutS       ------------------------------------------------------------
ScerMSH2        ------------------------------------------------------------
ScerMSH6        TTTDEDLSSSQSRRNHKRRVNYAESDDDDSDTTFTAKRKKGKVVDSESDEDEYLPDKNDG 180

HaloMutS1       ------------------------------------------------------------
HaloMutS2       ------------------------------------------------------------
EcoliMutS       ------------------------------------------------------------
ScerMSH2        ------------------------------------------------------------
ScerMSH6        DEDDDIADDKEDIKGELAEDSGDDDDLISLAETTSKKKFSYNTSHSSSPFTRNISRDNSK 240

HaloMutS1       ------------------------------------------------------MG- 2
HaloMutS2       ----------------------------------------------------MDAALGP 7
EcoliMutS       ----------------------------------------------------MSAIEN- 6
ScerMSH2        ------------------------------------------------------MSSTR 5
ScerMSH6        KKSRPNQAPSRSYNPSHSQPSATSKSSKFNKQNEERYQWLVDERDAQRRPKSDPEYDPRT 300

HaloMutS1       ------------IVDEFQALKAETDADLLAMQVGDFYEFFAADARTVASVLDLQVSEK- 48
HaloMutS2       PDAMAASEGDLTPMMSQYFELTRRYDDALVLFQVGDFYELFCAAAETAARICEVTLTAR- 66
EcoliMutS       -------FDAHTPMMQQYLRLKAQHPEILLFYRMGDFYELFYDDAKRASQLLDISLTKRG 59
ScerMSH2        PELKFSDVSEERNFYKKYTGLPKKPLKTIRLVDKGDYYTVIGSDAIFVADSVYHTQSVLK 65
ScerMSH6        LYIPSSAWNKFTPFEKQYWEIKSKMWDCIVFFKKGKEFELYEKDALLANALFDLKIAGGG 360
                                              ***
HaloMutS1       ---------SNHGSSYPMAGVPVDDLTPYLAA-LVERGYRVAVAEQS---ETDAGD---- 91
HaloMutS2       ---------EDSTGQYPMAGVPIDTAEPYIEA-LLDAGYRVAVADQVQDPDEVSGV---- 112
EcoliMutS       ---------ASAGEPIPMAGIPYHAVENYLAK-LVNQGESVAICEQIGDPATSKGP---- 105
ScerMSH2        NCQLDPVTAKNFHEPTKYVTVSLQVLATLLKLCLLDLGYKVEIYDKGWKLIKSASPGNIE 125
ScerMSH6        ------------RANMQLAGIPEMSFEYWAAQ-FIQMGYKVAKVDQRESMLAKEMREGSK 407

HaloMutS1       --IEREIERVVTPGTLLAST---DADPRYLAAVV------------------REAGGDW 127
HaloMutS2       --VDRAVTRVVTPGTVTEDELLGGADNNFVAALAGG-----------------RDADAGF 153
EcoliMutS       --VERKVVRIVTPGTISDEALLQERQDNLLAAIW-------------------QDSKGF 143
ScerMSH2        -QVNELMNMNIDSSIIIASLKVQWNSQDGNCIIGVA-----------------FIDTTAY 167
ScerMSH6        GIVKRELQCILTSGTLTDGDMLHSDLATFCLAIREEPGNFYNETQLDSSTIVQKLNTKIF 467

HaloMutS1       GLAFVDVTTGQFRVTRG---ADRADAVTELYRFAPAEVLPGPALRGDDD--------FLG 176
HaloMutS2       GLALLDVSTGDCYATRL---GDEARVRDELGRFTPAELVVGPGVDAD------------- 197
EcoliMutS       GYATLDISSGRFRLSEP---ADRETMAAELQRTNPAELLYAEDFAEMS------------ 188
ScerMSH2        KVGMLDIVDNEVYSN-----LESFLIQLGVKECLVQDLTSNSNSNAEMQ--------KVI 214
ScerMSH6        GAAFIDTATGELQMLEFEDDSECTKLDTLMSQVRPMEVVMERNNLSTLANKIVKFNSAPN 527

HaloMutS1       VLRERTDATLTLHDAGAFDAGRATHRVREQFG--DGVIESLGVAADGPVVRAAGAAVGYI 234
HaloMutS2       --RFADEAFVAAYDDDAFEPAAARERVADYFGGEDVLPTTAELRACGALLSYAEYTRG-- 253
EcoliMutS       --LIEGRRGLRRRPLWEFEIDTARQQLNLQFGTRDLVGFGVENAPRG--LCAAGCLLQYA 244
ScerMSH2        NVIDRCGCVVTLLKNSEFSEKDVELDLTKLLG--DDLALSLPQKYSKLSMGACNALIGYL 272
ScerMSH6        AIFNEVKAGEEFYDCDKTYAEIISEEYFSTEEDWPEVLKSYYDTGKKVGFSAFGGLLYYL 587

HaloMutS1       AAADEGV----LASVSRIQPFGGGDHVELDATTQRNLELTET----------------- 272
HaloMutS2       GAGDSQR----LTYLNHVTRYSPTEHLQMDAVALRSLELFEQR---------------- 292
EcoliMutS       KDTQRTT----LPHIRSITMEREQDSIIMDAATRRNLEITQN----------------- 282
ScerMSH2        QLLSEQD----QVGKYELVEHKLKEFMKLDASAIKALNLFPQGPQNPFGSNNLAVSGFTS 328
ScerMSH6        KWLKLDKNLISMKNIKEYDFVKSQHSMVLDGITLQNLEIFSN----------------- 629

HaloMutS1       MTGGSDG-SLLATIDHTASAAGGRRLAAWVTRPTRDRAELDRRQAAVGALADAALARDAL 331
HaloMutS2       SVHGTDGTALVDVLDETACALGRRKLTDWLRRPLVDSDAIAARHDAVGELVADPLSREEL 352
EcoliMutS       LAGGAEN-TLASVLDCTVTPMGSRMLKRWLHMPVRDTRVLLERQQTIGALQDFTAG---L 338
ScerMSH2        AGNSGKVTSLFQLLNHCKTNAGVRLLNEWLKQPLTNIDEINKRHDLVDYLIDQIELRQML 388
ScerMSH6        SFDGSDKGTLFKLFNRAITPMGKRMMKKWLMHPLLRKNDIESRLDSVDSLLQDITLREQL 689
```

```
HaloMutS1       GDVLGEIYDLERLASRAASGRADATDLLRVRDTLAALPDVADALT-----TTPELAESPA 386
HaloMutS2       HEHLRDVYDIERLVSRVSRGRANARDLRALADTLAVVPEVRGLLA-----DADARKLQSL 407
EcoliMutS       QPVLRQVGDLERILARLALRTARPRDLARMRHAFQQLPELRAQLE-----TVDSAPVQAL 393
ScerMSH2        TSEYLPMIPDIRRLTKKLNKRGNLEDVLKIYQFSKRIPEIVQVFTSFLEDDSPTEPVKEL 448
ScerMSH6        EITFSKLPDLERMLARIHSRTIKVKDFEKVITAFETIIELQDSLK-------NNDLKGDV 742

HaloMutS1       RDVLARVDRAAAADVRAELADALADDPPKTLSEGG---LLQAGYDEALDELLAAHDEHRA 443
HaloMutS2       REALDDLP-----EIRGLLDRAIVADPPQELTDGG---VIRDGYDERLDDLRATERAGKQ 459
EcoliMutS       REKMGEFA-----ELRDLLERAIIDTPPVLVRDGG---VIASGYNEELDEWRALADGATD 445
ScerMSH2        VRSVWLAPLSHHVEPLSKFEEMVETTVDLDAYEENNEFMIKVEFNEELGKIRSKLDALRD 508
ScerMSH6        SKYISSFP-----EGLVEAVKSWTNAFERQKAINENIIVPQRGFDIEFDKSMDRIQELED 797

HaloMutS1       WLDGLADREKDRLGITH---LQVDRNKTDGYYIQVGNSETDAVPDGEDGAYRRIKQLKNA 500
HaloMutS2       WVDDLEASERERTGVDS---LKVGQNSVHGYYIEVTKANMDAVPED----YQRRQTLKNA 512
EcoliMutS       YLERLEVRERERTGLDT---LKVGFNAVHGYYIQISRGQSHLAPIN----YMRRQTLKNA 498
ScerMSH2        EIHSIHLDSAEDLGFDPDKKLKLENHHLHGWCMRLTRNDAKELRKHKK--YIELSTVKAG 566
ScerMSH6        ELMEILMTYRKQFKCSN---IQYKDSGKEIYTIEIPISATKNVPSN----WVQMAANKTY 850

HaloMutS1       TRYTMAELDSHEREVLRIEAERAELERELFAALRERVGERA-AVLQDVGRALAEVDALVS 559
HaloMutS2       ERYVTPELKEREEEIVRAEQRAQDLEYELFVGIRERVAEAA-ERMQAVARALAAVDALAS 571
EcoliMutS       ERYIIPELKEYEDKVLTSKGKALALEKQLYEELFDLLLPHL-EALQQSASALAELDVLVN 557
ScerMSH2        IFFSTKQLKSIANETNILQKEYDKQQSALVREIINITLTYT-PVFEKLSLVLAHLDVIAS 625
ScerMSH6        KRYYSDEVRALARSMAEAKEIHKTLEEDLKNRLCQKFDAHYNTIWMPTIQAISNIDCLLA 910

HaloMutS1       LAEHA--AANQWVRPELVAGDG----------LDIDAGRHPVVEQ--TTS--FVPNDARF 603
HaloMutS2       FAAVA--AAHDYTKP-VMGGDG----------IHIEGGRHPVVER--TESG-FVPNDTTL 615
EcoliMutS       LAERA--YTLNYTCPTFIDKPG----------IRITEGRHPVVEQ--VLNEPFIANPLNL 603
ScerMSH2        FAHTSSYAPIPYIRPKLHPMDSERR-------THLISSRHPVLEM--QDDISFISNDVTL 676
ScerMSH6        ITRTSEYLGAPSCRPTIVDEVDSKTNTQLNGFLKFKSLRHPCFNLGATTAKDFIPNDIEL 970

HaloMutS1       DASR-RFQVVTGPNMSGKSTYMRQVAVIVLLAQVGSFVPADAARIGLVDGIYTRVGALDE 662
HaloMutS2       NDDR-RVAVITGPNMSGKSTYMRQVAVIVVLAQAGCFVPAAAAELRVVDRVFTRVGASDD 674
EcoliMutS       SPQR-RMLIITGPNMGGKSTYMRQTALIALMAYIGSYVPAQKVEIGPIDRIFTRVGAADD 662
ScerMSH2        ESGKGDFLIITGPNMGGKSTYIRQVGVISLMAQIGCFVPCEEAEIAIVDAILCRVGAGDS 736
ScerMSH6        GKEQPRLGLLTGANAAGKSTILRMACIAVIMAQMGCYVPCESAVLTPIDRIMTRLGANDN 1030

HaloMutS1       LAGGRSTFMVEMEELSRILHAATSDSLVVLDEVGRGTATYDGISIAWAATEYLHNEVRAT 722
HaloMutS2       IAGGRSTFMVEMTELASILRAATDESLVLLDEVGRGTATTDGLAIARAVTEHIHDAVGAT 734
EcoliMutS       LASGRSTFMVEMTETANILHNATEYSLVLMDEIGRGTSTYDGLSLAWACAENLANKIKAL 722
ScerMSH2        QLKGVSTFMVEILETASILKNASKNSLIIVDELGRGTSTYDGFGLAWAIAEHIASKIGCF 796
ScerMSH6        IMQGKSTFFVELAETKKILDMATNRSLLVVDELGRGGSSSDGFAIAESVLHHVATHIQSL 1090
                ••••••••••••••••           ------
HaloMutS1       TLFATHYHELTALADHLDAVVNVHVAAEERD---------GAVTFLRTVRDGATDRSYGV 773
HaloMutS2       TLFATHHHELTADADRLPDALNLHFAATRGD---------DGVAFEHAVRAGAATASYGV 785
EcoliMutS       TLFATHYFELTQLPEKMEGVANVHLDALEHG---------DTIAFMHSVQDGAASKSYGL 773
ScerMSH2        ALFATHFHELTELSEKLPNVKNMHVVAHIEKNLKEQKHDDEDITLLYKVEPGISDQSFGI 856
ScerMSH6        GFFATHYGTLASSFKHHPQVRPLKMSILVDEAT-------RNVTFLYKMLEGQSEGSFGM 1143

HaloMutS1       HVAALAGVPEPVVDRARGVLDRLREENAVEAKGSAGESVQAVFDVDSGGFVDDAGDDGEA 833
HaloMutS2       EVARTAGVPEPVVDRARELLD-----------------APATADGGDGGTTPTADANGQR 828
EcoliMutS       AVAALAGVPKEVIKRARQKLR----------------ELESISPNAAATQVDGTQMSLLS 817
ScerMSH2        HVAEVVQFPEKIVKMAKRKAN--------ELDDLKTNNEDLKKAKLSLQEVNEGNIRLKA 908
ScerMSH6        HVASMCGISKEIIDNAQIAADN--------------LEHTSRLVKERDLAANNLNGEVVS 1189

HaloMutS1       DDPEAAAVLDELRTVELAETSPVELLGTVQAWQDRLED------------------ 871
HaloMutS2       G--AAAGIVAELRDVSVAELTPIEALNVLNDLASRVD------------------- 863
EcoliMutS       VPEETSPAVEALENLDPDSLTPRQALEWIYRLKSLV-------------------- 853
ScerMSH2        LLKEWIRKVKEEGLHDPSKITEEASQHKIQELLRAIANEPEKENDNYLKYIKALLL 964
ScerMSH6        VPGGLQSDFVRIAYGDGLKNTKLGSGEGVLNYDWNIKRNVLKSLFSIIDDLQS--- 1242
```

Figure 1-7. Sequence alignment of MutS proteins from the three domains of life using ClustalW [95]. Halo refers to *Halobacterium* and Scer refers to *S. cerevisiae*. The PIP box is marked with (::); the DNA binding Phe-X-Glu motif is marked with (**); the

Walker A/P-loop motif is indicated by (__); the ABC transporter motif is indicated by (••); and the Walker B motif is marked as (--).

Euryarachaeota contain only one PCNA whereas the *Crenarchaeota* contain mutiple PCNAs [69, 96].

Also of interest is studying the response of *Halobacterium* to oxidizing agents. Genomic integrity is critical for survival but the genetic material is constantly being challenged by intracellular and extracellular stresses. A broader understanding of the DNA repair proteins and/or mutation avoidance properties can be achieved by studying organisms living in extreme environments, such as *Halobacterium*. *Halobacterium* is an aerobic organism and lives in an environment subjected to cycles of desiccation and rehydration. Studies looking at dessication and gamma irradiation have revealed that the major type of damage is caused by oxidative stress [86]. Oxidative stress is characterized by the presence of reactive oxygen species (ROS). ROS consist of unstable oxygen ions, free radicals, and peroxides, which include superoxide radical, hydroxyl radical, and hydrogen peroxide [97]. A common feature between the different ROS types is their capability of damaging different molecules in the cell including DNA, proteins, membrane lipids, and carbohydrates. Lesions to the DNA include single strand breaks, nucleotide modifications, and cross-linking of DNA to proteins [97, 98]. ROS are generated during normal aerobic metabolism but increased ROS can lead to oxidative stress. Studies looking at the effects of gamma irradiation show that *Halobacterium* is resistant to high doses of gamma irradiation (>5000 Gy) [85, 86]. Most of the deleterious effects of gamma-ray are a result of hydroxyl production by the radiolysis of water. Chromosome fragmentation after gamma irradiation was repaired within several hours of incubation [86]. Studies showing the high resistance of *Halobacterium* to dessication and gamma

irradiation, both of which lead to oxidative stress make *Halobacterium* a good model system to study the response of oxidative stress from other oxidizing agents.

1.5 Objectives and Aims

A fundamental issue in molecular biology is understanding how organisms maintain genomic stability. DNA MMR plays a key role in the recognition and repair of errors made during replication and other processes. The key proteins, MutS and MutL, are conserved from Bacteria to Eukarya. Comparative sequence analyses reveal that only 11 of the 49 completely sequenced archaeal genomes encode homologs of these proteins, however, the spontaneous mutation rate measured in the archaeon *S. acidocaldarius* is comparable to that of other DNA-based microorganisms [67, 68]. This suggests that some form of MMR exists in the Archaea. It is likely that if an archaeal-specific MMR system exists, it is also present in *Halobacterium*, but in addition *Halobacterium* has, along with other archaea including *H. volcanii*, bacterial MutS and MutL homologs. The MMR genes found in the few archaea containing MutS and MutL homologs are canonical bacterial *mutS* and *mutL* genes, suggesting that they might be the result of a lateral gene transfer event [6, 84]. The purpose of this study is to determine if the bacterial-like MMR genes in *Halobacterium* are essential for MMR in this organism. The general hypothesis is that *Halobacterium* carries out postreplicative removal of mismatches by a bacterial-like MMR pathway. We also chose to look at the response of *Halobacterium* to oxidative stress. *Halobacterium* lives in a high salt environment and its genetic material is constantly being challenged by environmental stress. Examining the stress response in an

organism living in an extreme environment will result in a broader understanding of mutation avoidance and repair pathways.

The specific aims of this project were:

1. To determine the genomic mutation rate and the spectrum of spontaneous mutations in the halophilic archaeon *Halobacterium*. We used fluctuation tests targeting genes of the UMP biosynthesis pathway and sequence analysis of the mutated genes.
2. To characterize the cellular role of MutS1, MutS2, MutL, and UvrD by in-frame targeted gene knockout and by the analysis of mutant phenotypes.
3. To elucidate the oxidative stress response to hydrogen peroxide and paraquat in *Halobacterium* using a whole genome transcriptional array, in-frame targeted gene knockout, and analysis of the mutant phenotypes.

Archaea are useful as model organisms because they have a simplified version of eukaryal DNA replication and repair systems. Determining the functions of the different proteins in the MMR pathway could help determine how *Halobacterium* is able to repair replication errors while living in an extreme environment. Only a few archaea have homologs of the key proteins involved in MMR from *E. coli*, including MutS, MutL, RecJ, and UvrD. They could repair these errors by using repair systems established in both Bacteria and Eukarya or use a combination of repair genes from different systems. In Bacteria and Eukarya, MMR-deficient cells display a significant increase in spontaneous mutation resulting in a mutator phenotype. In mammals, this loss of function can lead to a predisposition to cancer. By studying MMR in *Halobacterium*, the potential

for discovery of a new pathway based on the recruitment of other repair enzymes could be used to understand pathway interactions, genomic stability processes, and mutation avoidance in eukaryotes.

Chapter 2: Genomic mutation rate and mutational spectrum of *Halobacterium*

2.1 Introduction

Genomic integrity is critical for cell survival. The low spontaneous mutation rates found in organisms from the three domains of life underscore this point. Rates of spontaneous mutation have been calculated for both Bacteria and Eukarya including *E. coli* and *S. cerevisiae*. The average genomic mutation rate in *E. coli* using the *lacI* gene was 3.26 x 10^{-3} per replication and using the *hisGDCBHAFE* operon was 2.38 x 10^{-3} per replication [66, 99-101]. In *S. cerevisiae,* the average mutation rate using the *ura3* gene was 3.81 x 10^{-3} per genome per replication while the *can1* gene gave a similar value of 2.38 x 10^{-3} per genome per replication [66, 102, 103]. Analyses of spontaneous mutation rates in DNA-based microorganisms showed similarities between the Bacteria and Eukarya including an average genomic mutation rate of 3.4 x 10^{-3} per replication and a high occurrence (approximately 70%) of base pair substitutions (BPS) [65-67, 99, 100, 104, 105]. Only two analyses of genomic mutation rates in the Archaea have been undertaken, one in a thermophilic acidophile and one in an extreme halophile. Grogan *et al.* calculated the genomic mutation rate in *S. acidocaldarius,* a thermophilic acidophile, and found it to be 1.8 x 10^{-3} per replication, demonstrating that this archaeon has a similar mutation rate to the other domains of life [67]. Unlike other DNA-based microorganisms, the frequency of BPS in *S. acidocaldarius* was only 33% [67]. In *H. volcanii*, an extreme halophile, the genomic mutation rate was calculated to be 4.5 x 10^{-4} per replication, 7.5-

fold lower than the average genomic mutation rate calculated in the Bacteria and Eukarya [68]. In addition to a low genomic mutation rate, the proportion of BPS was 12% [68]. It is important to note that this study had a very small sample size, possibly introducing bias in the data.

Halophilic archaea grow optimally at moderate temperature in 2-4M salt and maintain osmotic balance by a high concentration of intracellular potassium chloride [106]. *Halobacterium* is an extreme halophile requiring 3.5 to 5M NaCl for growth. Genetic tools are available for this organism. DNA repair studies on the bacterial-like *uvrABC* genes in *Halobacterium* have shown that they are essential for repair of UV damage [88]. Other studies looking at homologous recombination using the *rad50* and *mre11* genes have been done in *Halobacterium* as well as a *radA* gene in *H. volcanii* [87, 89]. A large scale study of the accuracy of genomic replication has not yet been completed in *Halobacterium*.

The question we are asking in this chapter is whether *Halobacterium* has a functioning MMR repair system. If a MMR system is present in this organism, we should observe a low genomic mutation rate, whereas if there is no MMR system, we should observe a high genomic mutation rate. To distinguish between these two possibilities, we measured the rate of spontaneous genomic mutation in *Halobacterium* at the *pyrF, pyrE1,* and *pyrE2* loci and analyzed the spectrum of mutation through DNA sequencing. The *pyrF, pyrE1,* and *pyrE2* genes are part of the pyrimidine biosynthesis pathway (UMP pathway) [107]. If an orotate analog, 5-fluoroorotic acid (5-FOA) is present in the medium it can be

metabolized by these genes, which create a toxic byproduct that binds irreversibly to thymidylate synthase inhibiting pyrimidine synthesis. Mutations in these genes confer resistance to 5-FOA and pyrimidines are synthesized through an alternate pathway [67, 68, 107]. The *pyrF* gene encodes an orotidine 5'monophosphate decarboxylase (ODCase) and the *pyrE1* and *pyrE2* genes both encode an orotate phosphoribosyl transferase (OPRTase) (See Figure 2-1). There are two *pyrE* genes in several of the halophilic archaea, however studies in *H. volcanii* demonstrated that *pyrE2* encodes the physiological OPRTase of the cell with *pyrE1* showing partial resistance to 5-FOA [108]. The Luria-Delbruck experiment, also called a fluctuation test, demonstrated that genetic mutations arise in the absence of selection rather than as a response to selection [109, 110]. It is a commonly used method for measuring mutations rates in microorganisms. We used a fluctuation test selecting for mutations in the UMP biosynthetic pathway and our data showed a low mutation rate similar to that of other DNA-based microorganisms.

2.2 Methods and Materials

Organism and Growth Conditions:

Halobacterium sp. strain NRC-1 (ATCC number 700922) was grown in GN101 medium [250g/L NaCl, 20g/L $MgSO_4$, 2g/L KCl, 3g/L sodium citrate, 10g/L Oxoid brand bacteriological peptone] with the addition of 1 mL/L trace elements solution [31.5mg/L $FeSO_4 \cdot 7H_2O$, 4.4mg/L $ZnSO_4 \cdot 7H_2O$, 3.3mg/L $MnSO_4 \cdot H_2O$, 0.1mg/L $CuSO_4 \cdot 5H_2O$] at 42°C shaking in a Gyromax 737 shaker (Amerex Instruments, LaFayette, CA) at 220rpm. The GN101 media was supplemented with 50mg/L uracil and 20g/L agar for solid plates.

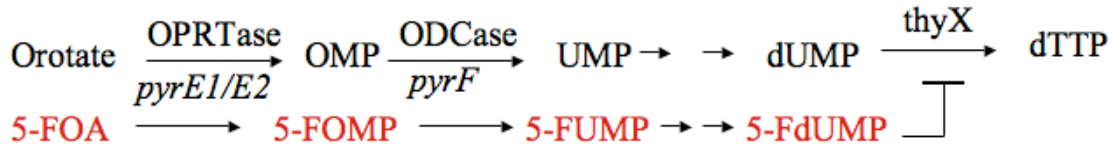


Figure 2-1. Pyrimidine biosynthesis pathway using the *pyrF*, *pyrE1*, and *pyrE2* genes. In the absence of 5-FOA, orotate is converted into OMP through the *pyrE1* and/or *pyrE2* gene. OMP is the converted into UMP by the *pyrF* gene and UMP is eventually converted into dUMP to form dTTP through the thymidylate synthase (*thyX)*. When 5-FOA is present, it can be taken up by the *pyrE1/E2* genes to form 5-FOMP, which will be converted into 5-FdUMP. 5-FdUMP binds irreversibly to thymidylate synthase blocking pyrimidine synthesis.

Basal salts solution (BSS), same composition as GN101 without the peptone, was used for culture dilutions.

For mutant selection, 350mg/L 5-FOA or 50mg/L 5-fluorouracil (5-FU) purchased from Sigma Aldrich (St. Louis, MO) were added. 5-FOA and 5-FU are taken up by the cells and converted into 5-FdUMP, which binds irreversibly to the thymidylate synthase inhibiting pyrimidine synthesis [67]. Loss of the UMP biosynthetic enzymes, *pyrF* (orotidine 5'-monophosphate), *pyrE1* and *pyrE2* (orotate phosphoribosyl transferases), renders cells resistant to 5-FOA with the addition of exogenous uracil and loss of *udp1* and *udp2,* uridine phosphorylases, renders cells resistant to 5-FU. 5-FU was discontinued in favor of 5-FOA since mutation rate studies have been done using 5-FOA and enable comparison between other organisms.

Mutant Isolation and Mutation Rate Assay:

Fluctuation tests were as previously described [67, 111] with some modifications. Fluctuation tests were started with a culture containing one colony of wildtype *Halobacterium* in GN101/50mg/L uracil media grown at 42°C shaking for two days. The innoculum was diluted back to 1 x 10^2 cells/mL and 150µL aliquots dispensed into a 96 well flat-bottomed plate. The resulting cultures were incubated an additional three days at 42°C without shaking until cell density reached approximately 1 x 10^5 cells/mL. Cell counts were determined by plating dilutions of culture on GN101/50mg/L uracil plates. Entire contents from the well were spread on GN101 plates with 50mg/L uracil and 350mg/L 5-FOA and resistant colonies were counted 11 days after incubation at 42°C.

To calculate the spontaneous mutation rate for wildtype *Halobacterium,* the relationship $\mu=\ln(m/N_{av})$ was used. The mutation rate, μ, is equal to the natural log of m, number of mutational events per culture, divided by N_{av}, average number of cells per culture. The m value was calculated using the MSS Maximum-Likelihood method as previously described [109]. Comparison of mutation spectrums was feasible using 5-FOA since it has been used to calculate mutation rates in *S. acidocaldarius* and *H. volcanii* [67, 68]. Mutants from the 5-FOA plates were chosen for sequence analysis by picking the colony closest to the middle of the plate. Picking the colony closest to the middle of the plate ensured no bias on colony size. Colonies were clonally purified by restreaking onto GN101 with 50mg/L uracil and 350mg/L 5-FOA plates twice and stored at -80°C.

DNA Sequencing:

The UMP biosynthesis genes, *pyrE1, pyrE2,* and *pyrF*, were amplified using colony PCR with the following primers: pyrE1-F (5'CCTCGTCCTGGAGAACAAAG3'), pyrE1-R (5'ATCGAAGGCCATGTCCCACCGT3'), pyrE2-F (5'GGTTCATACCGACCACACG3'), pyrE2-R (5'TCGGCGACACCTTCGGGCTG3'), pyrF-F (5'GCGCGCCTCGTGGTGTTCGT3'), and pyrF-R (5'AGCGTCGTCTGTGACACCCA3'). These primers allowed for approximately 200 bases before the start codon and 100 bases after the stop codon to be amplified along with the gene. PCR conditions for all three genes consisted of an initial 2 minute denaturation at 94°C then 30 cycles of 35 seconds at 94°C, 40 seconds at 56°C, and 60 seconds at 72°C, with a final extension time of 5 minutes at 72°C using FastTaq DNA polymerase (Roche, Indianapolis, IN). PCR products were purified using ExoSap-IT (GE Healthcare,

Piscataway, NJ) and sequencing was performed using an Applied Biosystems 3730xl DNA Analyzer with BigDye Terminator reagents (Applied Biosystems, Foster City, CA). Each colony was sequenced using both forward and reverse primers for each gene.

2.3 Results

The objective of this study was to calculate the genomic mutation rate and to characterize the mutational spectrum for *Halobacterium* to determine whether this organism has a functioning MMR system. We performed six independent fluctuation tests, a commonly used method to calculate mutation rates in microorganisms, from which 149 5-FOA-resistant mutants were recovered [66]. Of these, we sequenced 80 mutants using primers for the *pyrF* and *pyrE1* genes and 83 mutants using primers for the *pyrE2* gene. We found mutations in the *pyrF* gene for 50 mutants, in the *pyrE1* gene for 24 mutants, and in the *pyrE2* gene for 46 mutants. Several of the mutants had multiple mutations within each gene. Five of the mutants sequenced using the *pyrF* and *pyrE2* primers had no mutations, 48 of the mutants sequenced using the *pyrE1* primers had no mutations, and several mutants had mutations in more than one gene (See Figure 2-2 and Table 2-1).

The average rate of mutation at the gene level for the *pyrF*, *pyrE1*, and *pyrE2* genes was 2.79×10^{-3}, 3.52×10^{-3}, and 4.25×10^{-3} per replication, respectively. To correct this rate for undetected mutations, we estimated the total number of BPS using published information on BPS detection efficiency (approximately 0.2) [67, 68]. The resulting rate estimate per base pair was calculated as follows:

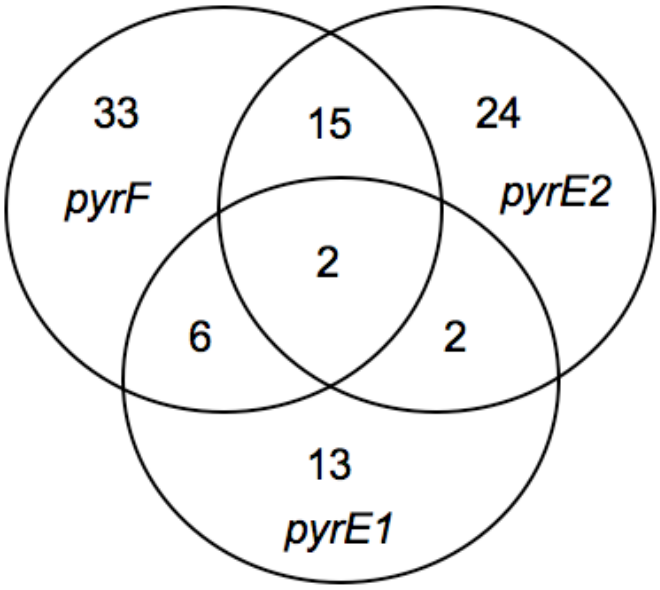


Figure 2-2. Ven diagram showing the distribution of mutations in the three UMP biosynthesis genes, *pyrF*, *pyrE1*, and *pyrE2*. There were 56 mutants with mutations in the *pyrF* gene including 21 double and 2 triple mutants. The *pyrE1* gene had mutations in 23 mutants including 8 double and 2 triple mutants. Forty-three mutants had mutations in the *pyrE2* gene including 17 double and 2 triple mutants.

Table 2-1. Spontaneous mutations in the *pyrF, pyrE1,* and *pyrE2* genes. Mutants were sequenced using primers 200 base pairs upstream of the putative promoter region and 100 base pairs after the stop codon for each gene. We noted three different types of mutations, insertions, deletions, and BPS. Gene position refers to the nucleotide position within each of the UMP biosynthesis genes.

Gene Position	Mutant Clone Number	Mutation
pyrF		
-87	131, 149, 163	Deletion (G)
-86	148	Insertion (C)
-86	131, 126, 121, 123, 155, 127, 156, 154, 142, 138, 149, 160, 158, 137, 161, 140, 159, 152, 134, 110, 150, 111, 144, 163, 120, 141, 139	T→ G
-85	121, 123, 155, 127, 156, 154, 142, 138, 137, 159, 152, 134, 110, 111, 144, 120	T → G
-84	126, 131, 123, 155, 127, 156, 154, 142, 138, 149, 137, 159, 152, 134, 110, 111, 144, 120, 141, 146, 160, 158, 161, 140, 150, 163, 157, 139	Deletion (G)
-84	121, 133	T →C
-83	146, 161, 59, 148, 107, 126, 124, 147	Insertion (A)
-83	121, 123, 155, 127, 156, 154, 142, 138, 149, 160, 137, 140, 159, 152, 134, 110, 150, 111, 144, 163, 133, 148	Deletion (C)
-83	120	T →G
-83	131, 141, 139	G →T
-81	107	Insertion (A)
-75	120	G →C
-73	107	Insertion (C)
-62	120	Insertion (A)
-56	147	Insertion (T)
-51	147, 107	Insertion (T)

-45	126, 16, 121, 123, 155, 146, 127, 156, 154, 142, 142, 138, 149, 158, 137, 161, 140, 159, 152, 134, 110, 111, 144, 120, 139, 131, 160, 150, 163	Deletion (G)
-45	147, 157, 148, 107	G →T
-43	147, 141	Insertion (T)
22	5, 55, 79, 22, 54	Insertion (C)
23	5	G →T
23	79, 22, 54, 62	Deletion (G)
24	5	Insertion (T)
24	62, 55, 79, 22, 54, 27, 86	Insertion (G)
28	9	Deletion (C)
81	120	Insertion (A)
81	107	G →T
95	120	Insertion (G)
145	157	G →T
148	141	G →T
155	131	A →T
163	148	G →T
175	139	A →T
196	148	G →T
233	148	T →C
276	148	G →C
287	107	A →T
289	147	A →T
292	147	G →T
294	148	G →T
325	148	G →T
340	148	A →T
354	55, 54, 56, 53	Insertion (C)
354	148	G →T
356	55, 54, 56, 53	Insertion (C)
356	139	A →T
357	55, 54, 56, 53	Insertion (A)
358	148	C →T
359	55, 54, 56, 53	Insertion (C)
360	147	Insertion (A)
361	55, 54, 56, 53	G →T
362	55, 54, 56, 53	Insertion (C)
365	55, 54, 56, 53	Insertion (G)
393	141	G →T
397	126, 109, 121, 124, 123, 127, 110, 111, 120	G →T

412	148	G →T
417	148	Insertion (T)
423	148	C →T
439	131	Deletion (G)
440	131	Insertion (A)
440	148	Insertion (G)
450	148	C →T
452	133	A →T
463	160	C →A
484	141	Insertion (G)
582	133	C →T
642	140	C →A
670	142	C →T
705	140	G →A
754	109, 158	G →A
756	131, 126, 109, 142, 149, 158, 161, 134, 120	Insertion (A)
756	140	C →G
777	120	Insertion (T)
783	109, 142	Insertion (A)
788	109, 134	Insertion (C)
788	120	Insertion (T)
788	131	T →C
790	126, 146	Insertion (C)
790	134	Insertion (A)
790	142, 158	Insertion (T)
797	40	A →T
798	5, 55, 30	Insertion (C)
798	40	Insertion (T)
799	95	Insertion (A)
802	81	T →G
pyrE1		
-100	98, 181	G →T
-99	98, 181	Insertion (T)
-61	181	Insertion (T)
-60	181	C →T
-43	104	C →A
-40	104	Insertion (G)
1	27, 42, 55	A →T
1-5	30	Deletion (ATGAA)
3	54	Insertion (T)
3	27, 31	G →A
4	56	Insertion (T)
4	30	A →C

4	66	Deletion (A)
6	31, 43	G →A
6	58	Deletion (G)
9	30	Insertion (T)
9	43	Insertion (A)
11	31	Insertion (G)
12	30, 31	C →T
12	2, 3, 9, 11, 23, 81	Deletion (C)
13	2, 3, 9, 11, 23, 30, 31, 81	Insertion (C)
15	30	C →A
16	31	G →A
21	31	Insertion (G)
21	30	C →T
24	31	Insertion (T)
28	30	G →T
52	31	C →A
53	30, 31	C →A
55	31	C →T
57	30, 56	C →T
73	98	G →C
549	181	C →A
550	106	C →A
552	106	C →G
573	106	C →A
581	3	A →C
584	106	G →C
586	3, 9	Insertion (C)
586	106	A →T
587	3	Deletion (C)
587	181	C →A
588	2, 104, 106	Insertion (T)
588	3	G →C
588	181	G →A
589	104, 106, 181	Insertion (A)
589	180	Insertion (C)
590	2, 56	Deletion (C)
590	29	Insertion (A)
590	3, 104, 106	C →T
590	181	C →A
591	2, 29, 56	Insertion (A)
591	3	Insertion (T)
592	56, 181	A →G
592	104, 106	Deletion (A)
593	2, 56	T →G
593	3	T →A

593	29	Deletion (T)
594	27	Deletion (C)
594	29, 42	C →G
594	3	C →A
595	3	G →A
599	3	Insertion (G)
602	106	T →A
620	181	Deletion (T)
621	139	Insertion (T)
621	181	Insertion (A)
625	181	Insertion (A)
625	106	Insertion (G)
627	106	Insertion (A)
630	106	Insertion (G)
632	2	Insertion (A)
632	106	Insertion (G)
634	106	G →C
635	2	Insertion (G)
636	2, 106, 181	Deletion (G)
637	106	Deletion (A)
638	17	Insertion (T)
pyrE2		
-72	30, 95	Insertion (C)
-71	30	Insertion (T)
-9	33	C →T
381	156, 153, 181, 150, 182, 132, 164, 178, 163, 170, 177, 169, 168, 152, 158, 157, 166, 180, 151, 146, 140, 171, 149, 176, 162, 165, 167, 161, 135, 179, 183, 143, 155, 172, 160, 141	Deletion (G)
382	65, 95, 93, 94, 79, 81	Deletion (C)
383	65, 95, 93, 94, 79, 81	Deletion (G)
383	156, 153, 181, 150, 182, 132, 164, 178, 163, 170, 177, 169, 168, 152, 158, 157, 166, 180, 151, 146, 140, 171, 149, 176, 162, 165, 167, 161, 135, 179, 183, 143, 155, 172, 160, 141	C →T
384	65, 95, 93, 94, 79, 81	Deletion (C)
385	65, 95, 93, 94, 79, 81, 156,	Deletion (A)

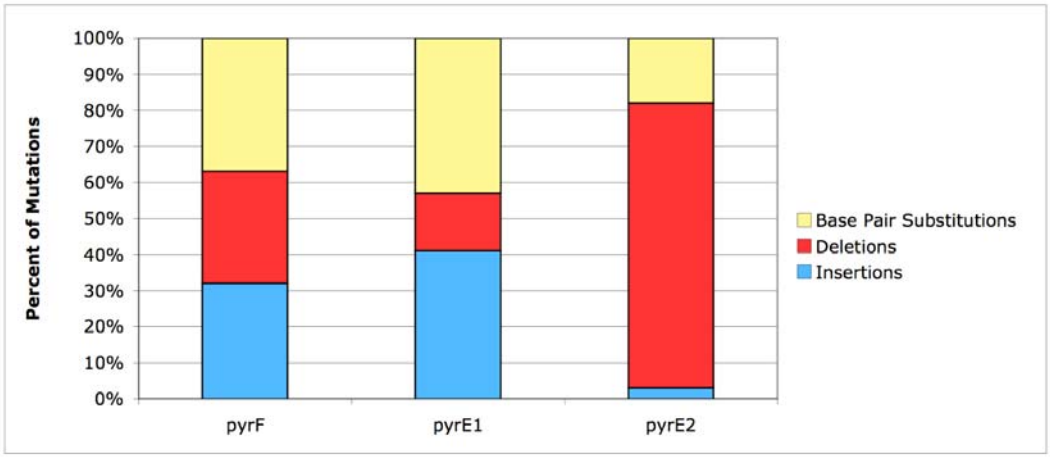
	153, 181, 150, 182, 132, 164, 178, 163, 170, 177, 169, 168, 152, 158, 157, 166, 180, 151, 146, 140, 171, 149, 176, 162, 165, 167, 161, 135, 179, 183, 143, 155, 172, 160, 141	
386	156, 153, 181, 150, 182, 132, 164, 178, 163, 170, 177, 169, 168, 152, 158, 157, 166, 180, 151, 146, 140, 171, 149, 176, 162, 165, 167, 161, 135, 179, 183, 143, 155, 172, 160, 141, 65, 95, 93, 94, 79, 81	Deletion (A)
387	65, 95, 93, 94, 79, 81	Deletion (G)
388	65, 95, 93, 94, 79, 81	Deletion (C)
390	65, 95, 93, 94, 79, 81	C →T
392	65, 95, 93, 94, 79, 81	Deletion (C)
394	65, 95, 93, 94, 79, 81	C →G
397	65, 95, 93, 94, 79, 81	Deletion (G)
471	156	G →A
472	156	Insertion (A)
500	156	Insertion (A)
505	156	Insertion (T)
505	69	Insertion (G)
508	69	Insertion (G)
508	30	Insertion (A)
510	69	C →A

$$\frac{\text{gene rate x (insertions + (BPS/correction factor))}}{\text{insertions + BPS}}$$

This was then converted into a genomic rate by dividing by gene size (*pyrF* = 803bp, *pyrE1* = 638bp, *pyrE2* = 527bp) and multiplying by genome size (2571010bp). This correction factor had a small effect on the genomic mutation rate of *Halobacterium* because of the low frequency of BPS. The average genomic mutation rate was corrected for undetected mutations and calculated to be $3.43 \times 10^{-3} \pm 5.7 \times 10^{-4}$ per replication using the equation, $\mu = \ln(m/N_{av})$ [109].

The spectrum of mutation was characterized for all three genes. The distribution of BPS, insertions, and deletions is summarized in Figure 2-3. The majority of mutations found in the *pyr* genes were deletions, however, the percentage of deletions varied by gene (See Figure 2-3). The *pyrF* gene had equal numbers of BPS, insertions and deletions (approximately 100 each). Deletions were found mainly at the beginning of the gene upstream of the putative promoter region (See Figure 2-4). The *pyrF* gene is not found in an operon and it is possible that a regulatory element might be located further upstream of the gene which could explain why there is a large number of mutations located upstream of the putative promoter region. The BPS and insertions were found throughout the gene (See Figure 2-4). The *pyrE1* gene had similar numbers of BPS and insertions (approximately 50) with fewer deletions (approximately 25) (See Figure 2-3). Almost all of the mutations were found at the beginning and end of the gene (See Figure 2-5). The *pyrE2* gene had an overwhelming majority of deletions (79% or approximately 375)

A.



B.

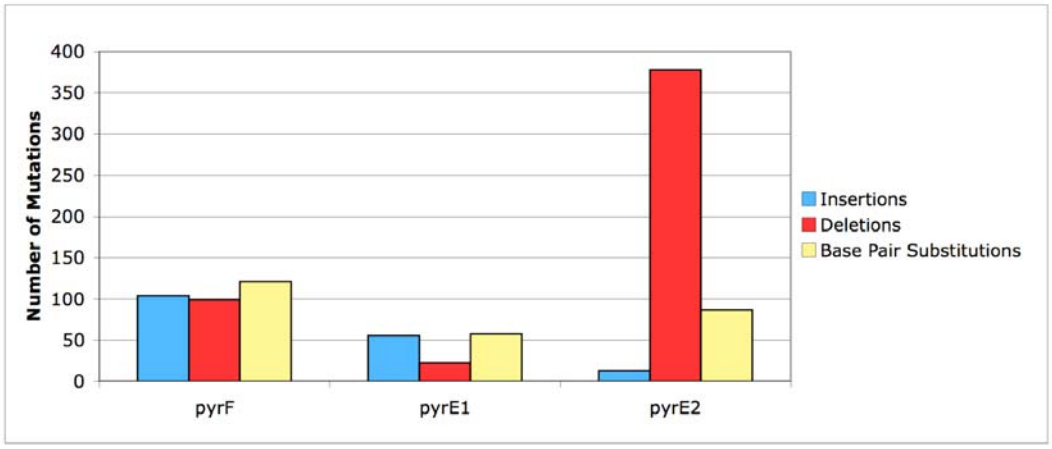


Figure 2-3. Distribution of BPS, insertions, and deletions within each of the UMP biosynthesis genes. (A) Percent and (B) number of BPS, insertions, and deletions in the *pyr* genes. Data was obtained by sequencing each of the genes using primers that started approximately 200 base pairs before the start codon and ended 100 base pairs after the stop codon. Most of the deletions found in the *pyr* genes were in *pyrE2* and were predominately deletions.

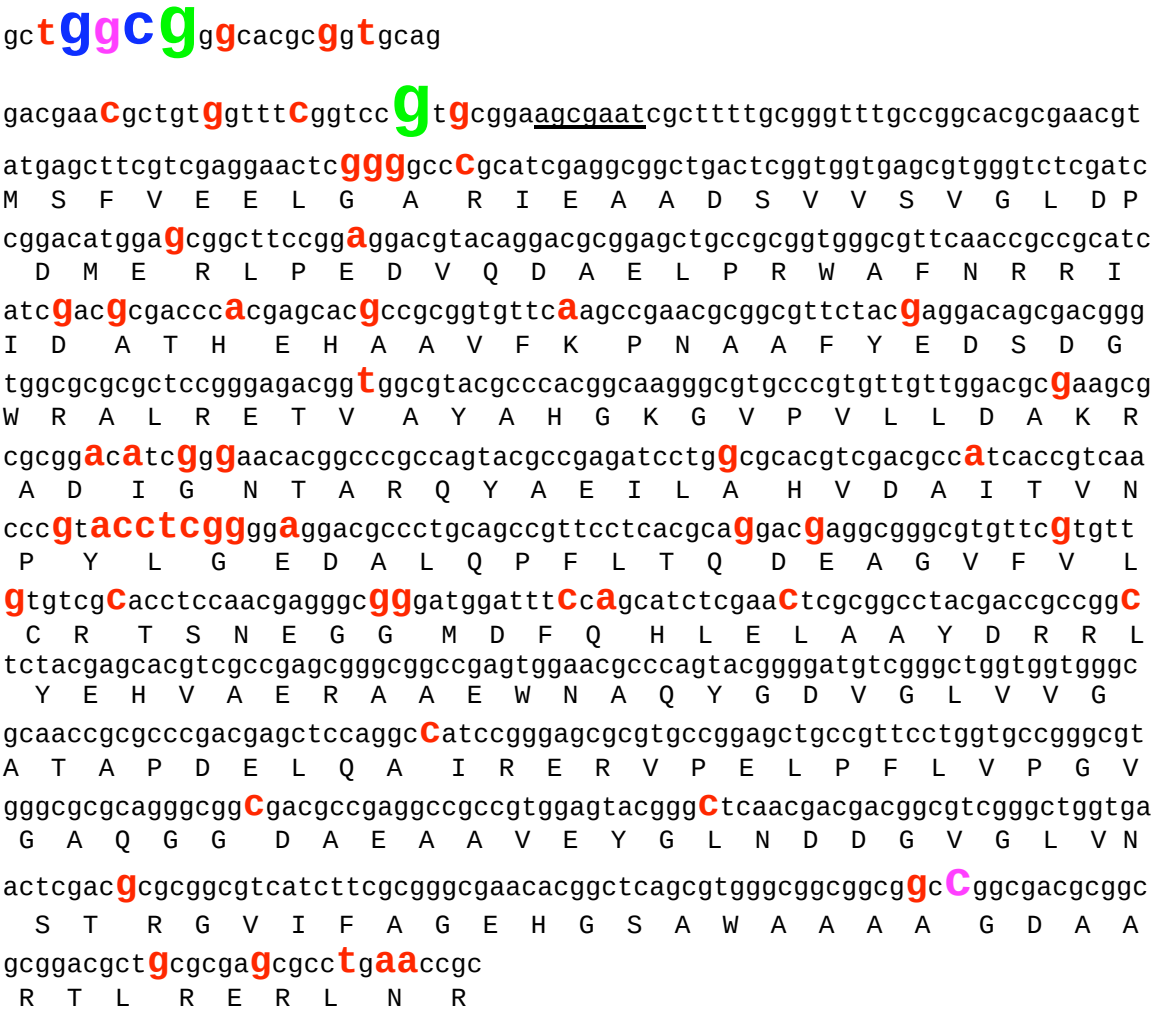


Figure 2-4. Mutational spectrum of the *pyrF* gene for 80 mutants. One letter amino acid abbreviations are listed below the nucleotide sequence and begin at the start codon of the *pyrF* gene. The putative TATA box is underlined. Nucleotide changes are shown in bold; the letter size indicates the number of changes found by sequence analysis of the *pyrF* gene in the mutant strains. Letter sizes are as follows: No change = a; 1-10 changes = a; 11-20 changes = a; 21-30 changes = a; >30 changes = a.

Figure 2-5. Mutational spectrum of the *pyrE1* gene for 80 mutants. One letter amino acid abbreviations are listed below the nucleotide sequence and begin at the start codon of the *pyrE1* gene. The putative TATA box is underlined. Nucleotide changes are shown in bold; the letter size indicates the number of changes found by sequence analysis of the *pyrE1* gene in the mutant strains. Letter sizes are as follows: No change = a; 1-10 changes = **a**.

compared to insertions and BPS. The deletions were all found at the end of the gene (See Figure 2-6). The insertions and BPS were found at both the beginning and end of the gene (See Figure 2-6). The *pyrE2* gene has a sensitive region located towards the end of the gene where most of the deletions were found. The insertions and deletions found in the *pyr* genes were only one or two base pairs long except in the *pyrE2* gene where there was an eight base pair deletion at the latter part of the gene (See Table 2-1).

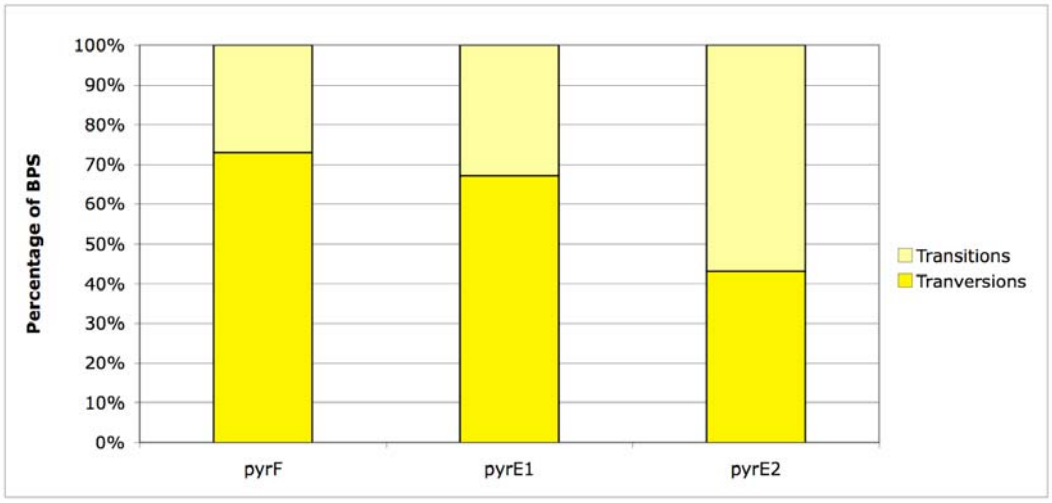
Next, we looked at whether the BPS in the UMP genes were mostly transversions or transitions. In the *pyrF* and *pyrE1* genes, the BPS were mostly transversions (See Figure 2-7). In the *pyrF* gene, BPS were dominated by a GC to TA transversions whereas the *pyrE1* gene did not have any dominant BPS. The *pyrE2* gene had slightly more transition mutations than transversion mutations and similarly to the *pyrE1* gene, there was no bias towards a particular BPS (See Figure 2-7). Of the BPS, approximately 70% led to a nonsynonomous change in amino acid in all three UMP genes (See Table 2-2).

2.4 Discussion

The spontaneous genomic mutation rate of *Halobacterium* was calculated to be 3.43 x 10^{-3} ± 5.70 x 10^{-4} per replication based on the rate of spontaneous mutation for the UMP biosynthetic genes and adjusted for undetected mutations. This rate is within the range of genomic mutation rates, 1.90-4.60 x 10^{-3} per replication, calculated by Drake [66] for DNA-based microorganisms. This indicates that either DNA repair mechanisms to maintain a low spontaneous mutation rate are present in *Halobacterium* or there is a low occurrence of mutations occurring.

Figure 2-6. Mutational spectrum of the *pyrE2* gene for 83 mutants. One letter amino acid abbreviations are listed below the nucleotide sequence and begin at the start codon of the *pyrE2* gene. The putative TATA box is underlined. Nucleotide changes are shown in bold; the letter size indicates the number of changes found by sequence analysis of the *pyrE2* gene in the mutant strains. Letter sizes are as follows: No change = a; 1-10 changes = **a**; 11-40 changes = **a**; >40 changes = **a**.

A.



B.

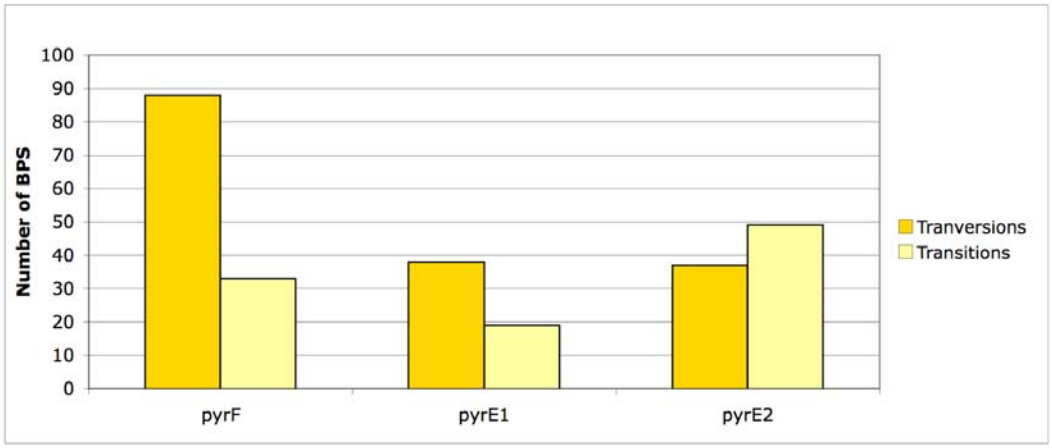


Figure 2-7. Percent (A) and number (B) of transitions and transversions in the UMP biosynthetic genes. (A) The *pyrF* gene had 37% BPS, of which 73% were transversions and 27% transitions. In the *pyrE1* gene, out of the 43% BPS, 67% were transversions and 33% transitions. The *pyrE2* gene had 18% BPS with a 43% occurrence of transversions and 57% occurrence of transitions. (B) The *pyrF* and *pyrE1* genes had more transversion mutations than transition mutations. The *pyrF* gene had 88 transversion mutations and the *pyrE1* gene had 38 transversion mutations. In contrast, the *pyrE2* gene had more transition mutations (49) than transversion mutations.

Table 2-2. Percentage of BPS in *Halobacterium* that led to a nonsynonymous change in amino acid. All three genes showed that approximately 70% of the BPS led to a change in amino acid.

Gene	Nonsynonymous Amino Acid Change
pyrF	71%
pyrE1	70%
pyrE2	73%

The spontaneous genomic mutation rate calculated in *Halobacterium* is also similar to what has been calculated in other archaea, including *H. volcanii,* 4.5×10^{-4} per replication, and *S. acidocaldarius,* 1.80×10^{-3} per replication [67, 68]. No homologs of the MMR pathway have been found in the thermophilic acidophiles, the clade in which *S. acidocaldarius* is found, leading to the puzzling question of what is responsible for maintaining the low genomic mutation rate calculated in the Archaea [77].

One hypothesis to explain the low genomic mutation rate observed in the Archaea is the high genome copy number found in both *Halobacterium* and *H. volcanii*, approximately 15-25 copies per cell depending on growth phase [68, 80]. *Halobacterium* has multiple copies of its major and its two minichromosomes. An advantage of gene redundancy is that it is more difficult to obtain homozygotes for a given mutation. It may take several generations for homozygous mutants to be generated, resulting in mutations not being detected if they only occur at low copy number [68]. However, *S. acidocaldarius* has only two copies of its genome per cell and it still maintains a low genomic mutation rate so this cannot explain the low genomic mutation rate seen in the Archaea as a whole [112].

An alternative hypothesis might come from adaptive mechanisms these organisms have evolved to live in extreme environments. For example, archaeal DNA polymerases that have been characterized from hyperthermophiles exhibit high fidelity when compared with that of Bacteria and Eukarya [72-74, 76]. Both *Pyrococcus furiosus* and *Thermus aquaticus* are used commercially as high fidelity polymerases, *P. furiosus* being a

member of the hyperthermophilic archaea. Studies calculating replication fidelity show that *P. furiosus* polymerase has 10-fold higher replication fidelity than *T. aquaticus* polymerase, which does not have an exonuclease activity, and 5-fold higher replication fidelity than *E. coli* DNA polymerase III holoenzyme [72, 74, 76]. The replicative polymerases in the Archaea are members of the B-family and are more similar to the eukaryotic replicative polymerases than the bacterial ones [69-71]. In the *Euryarchaeota,* the kingdom to which halophiles and methanogens belong, there are two types of putative replicative polymerases: the B-family and the D-family [69]. The D-family is unique to the *Euryarchaeota* and contains an exonuclease activity suggesting it could be an alternative replicative polymerase. The domain *Crenarchaeota,* which contains mostly thermophiles, only has the B-family of replicative polymerases [69]. The higher fidelity of archaeal polymerases along with structural and sequence differences and repair pathways could participate in maintaining the genomic integrity of the Archaea.

In addition to a low genomic mutation rate, *Halobacterium* also exhibited a very high number of deletions and a low percentage of BPS, which is very different than what is seen in the other two domains of life. In *Halobacterium,* there was an 18-43% occurrence of BPS with approximately 70% resulting in a non-synonomous change in amino acid. Similarly in other archaea, *S. acidocaldarius* and *H. volcanii* also exhibited a high number of insertions and deletions but very few BPS. There was a 33% and 12% occurrence of BPS respectively for those organisms [67, 68]. Our data and that of others showed that Archaea exhibit a high occurrence of insertion and deletions, while in other domains of life, BPS make up an overwhelming majority, 60-80%, of mutations (See

Table 2-3 and Figure 2-8) [67, 68]. This may be a result of the structure and function, for example, the higher fidelity, of polymerases in the Archaea compared to those in the Bacteria and Eukarya.

In *Halobacterium* the BPS in the *pyrF* and *pyrE1* genes were comprised of mostly transversions and in the *pyrF* gene they were mostly GC to TA transversions. Similarly, in *S. cerevisiae,* using the *ura3* (*pyrF*) gene, most of the BPS in were transversions [66, 113]. The predominance of transversion mutations seems to be gene-specific for *pyrF*.

Unlike the *pyrF* and *pyrE1* genes in *Halobacterium*, in the *pyrE2* gene, we found mostly transitions but no bias towards a particular transition mutation. This is similar to what is seen in *S. acidocaldarius* where the BPS in the *pyrE* gene were predominately GC to AT transitions [67]. In *E. coli,* using the *lacI* and *rpsL* genes as the target, there was a higher frequency of transition mutations than transversion mutations [66, 99, 100, 104, 105]. There were only two BPS described in the genomic mutation rate study in *H. volcanii* both of which led to a change in amino acid [68]. Also interesting is that in *S. acidocaldarius* and *H. volcanii*, only one of the UMP biosynthesis genes was found to contain mutations whereas we see mutations across all three UMP genes in *Halobacterium.*

One hypothesis for the low occurrence of BPS in the Archaea is that both *Halobacterium* and *H. volcanii* have a high GC content (>60%) which results in long runs of G and C nucleotides [79]. During replication, these GC runs, could cause polymerase slippage

Table 2-3. BPS frequencies in bacteriophage, Bacteria, Eukayra, and Archaea (Table adapted from [67]). The Archaea have a much lower frequency of BPS than the Bacteria, Eukarya, and bacteriophage.

Organism (reporter gene)	% BPS
Bacteriophage M13 (*lacZ*α)	57
Bacteriophage λ	60
Bacteriophage T4	62
E. coli (λ*cI*, *lacI*^{q}, *crp*, *supF*, *rpsL*, *tonB*, *lacI*)	66
S. cerevisiae (*SUP4*, *URA3*, *CAN1*)	87
Mouse, rat, hamster, monkey, human (*lacI*, *gpt*, *hprt*, *aprt*, *supF*, *tk*, *cI*, *cII*)	70
S. acidocaldarius (*pyrE*)	33
H. volcanii (*pyrE2*)	12
Halobacterium (*pyrF*, *pyrE1*, *pyrE2*)	33

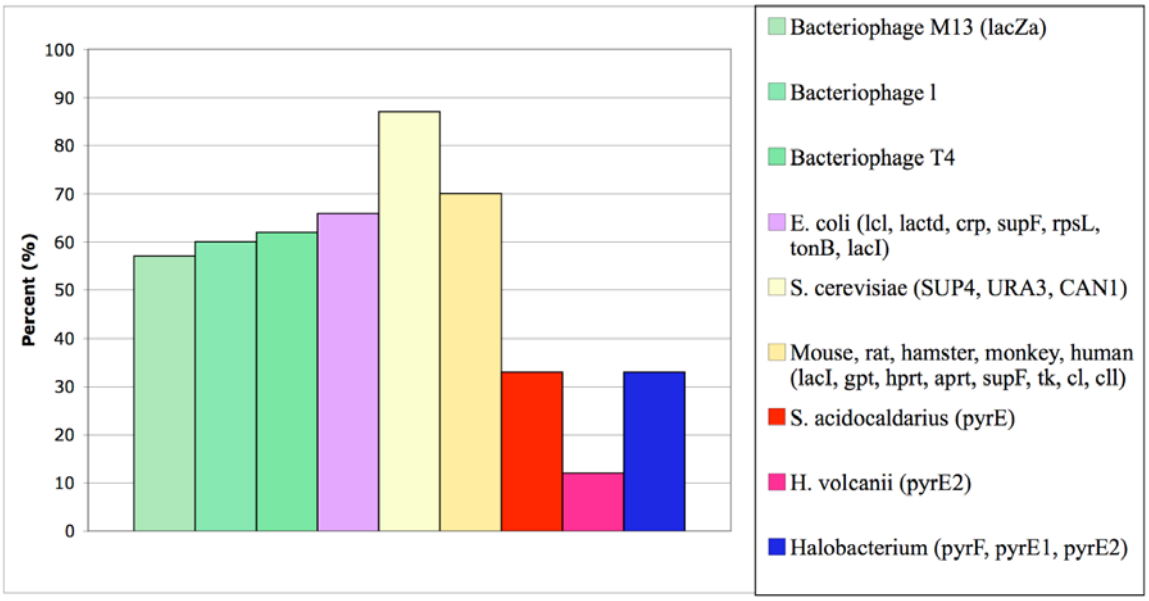


Figure 2-8. Percentage of BPS in bacteriophage, Bacteria, Eukarya, and Archaea (adapted from [67]). The bacteriophage have an approximately 60% occurrence of BPS, *E. coli* has a 66% occurrence of BPS, the Eukarya have an approximately 75% occurrence of BPS, while the Archaea have only a 26% occurrence of BPS.

leading to insertion or deletion of bases. However, *S. acidocaldarius* only has a 36% GC content so this could not be used as an explanation for the low percentage of BPS in the Archaea as a whole [112]. One alternative hypothesis is that the low percentage of BPS is a result of the specificity of the archaeal polymerases described above. These polymerases are high fidelity but it is also possible that structural and functional differences in the polymerase contribute to the high number of insertions and deletions in the Archaea and the low occurrence of BPS. In addition to a high fidelity polymerase, it is possible that there is some error correction system specific to Archaea perhaps utilizing an array of glycosylases to perform mismatch correction. Glycosylases are responsible for removing a damaged base, such as an oxidized guanine (8-oxo-G) [1]. There are four main classes of glycosylases, uracil-DNA, endonuclease III (Nth), formamidopyrimidine-DNA/endonuclease VIII (Fpg/Nei), and others [2]. Of these, the Archaea have members of the uracil-DNA and Nth family of glycosylases. While these glycosylases share the same general function as their family counterparts, they are structurally different and belong in their own group within each family [2].

One of the major pathways used in Bacteria and Eukarya to correct insertions, deletions, and BPS after DNA polymerase proofreading is the DNA MMR pathway [3-5]. *Halobacterium* has homologs of the bacterial-like MutS and MutL proteins responsible for the postreplicative removal of mismatches in bacteria and eukaryotes. Homologs from the MMR pathway have been found in a handful of archaeal genomes based on sequence comparison [77]. These archaea comprise several closely methanogens and halophiles.

Using protein BLAST analysis of the MutS1, MutS2, and MutL proteins in *Halobacterium,* we found homologs in 4 halophiles and 7 methanogens including *Haloarcula marismortui, Halorubrum lacusprofundi, Haloquadratum waisbye, Natronomonas pharaonis, Methanococcoides burtonii, Methanosaeta thermophila, Methanosarcina barkeri, Methanosarcina acetivorans, Methanosarcina mazei, Methanoculleus marisnigri,* and *Methanospirillum hungatei*. We also found, using protein BLAST analysis, homologs of the UvrD protein in *Halobacterium* in all the above archaea with the exception of *Methanosaeta thermophila.* We hypothesized that the bacterial-like MutS and MutL homologs in *Halobacterium* are involved in a MMR pathway similar to that seen in the Bacteria and Eukarya. However, it is likely that *Halobacterium* also has the same type of mismatch avoidance and repair system that might be present in other archaea.

In the Archaea with MMR homologs, MMR could be completed similarly to the Bacteria and Eukarya. Studies looking at the bacterial NER homologs in *Halobacterium* have shown that they are essential for the repair of UV damage [88]. This suggests that the bacterial-like MMR proteins found in *Halobacterium* might also be essential for MMR. We propose to characterize the cellular functions of the MutS1, MutS2, MutL, and UvrD proteins in *Halobacterium* using a targeted gene deletion approach and phenotypic characterization of the resulting mutants.

Chapter 3: Genetic inactivation of MMR homologs in *Halobacterium* to determine their cellular function

3.1 Introduction

MMR is the major pathway responsible for repairing errors during DNA replication, and deletion of genes involved in MMR have been shown to lead to a higher than average spontaneous mutation rate [4]. The MutS and MutL proteins are essential for MMR in Bacteria and Eukarya and are therefore highly conserved [3-5]. The major difference between the two systems is that in Bacteria, MutS and MutL are single proteins that form homodimers whereas in the Eukarya, there are mutiple MutS and MutL homologs that form heterodimers, suggesting a more complex pathway [3, 4]. No MMR pathway has been characterized in the Archaea so far but there is evidence that postreplicative removal of mismatches is occuring [67, 68]. Genomic mutation rate studies in *S. acidocaldarius* and *H. volcanii* show a low genomic mutation rate per replication close to the average (0.0034) calculated for both bacteria and eukaryotes [66-68]. We calculated the rate of spontaneous genomic mutation in *Halobacterium* to be 0.0034 per replication based on the rate of spontaneous mutation for genes of the UMP biosynthesis pathway. This rate is within the range of genomic mutation rates per replication (0.0019-0.0046), previously calculated by Drake for DNA-based microorganisms [66].

Defects in the MMR pathway lead to genomic instability which can cause a 50-1000 fold increase in spontaneous mutability, meiotic defects, and tolerance to several DNA damaging agents [3-5]. In humans, inactivation of the MMR pathway leads to simple

repeat instability resulting in hereditary nonpolyposis colon cancer [4, 5]. Based on sequence comparison only eleven of the fully sequenced Archaea have homologs to the conserved MutS and MutL proteins and these include four halophiles and seven methanogens. No studies investigating these MMR homologs in the Archaea have been carried out.

An increased mutation rate following deletion of MMR genes has been demonstrated both in Bacteria and in Eukarya. In *D. radiodurans* and *E. coli,* the mutation rate was calculated in cells deficient in MutS, MutL, or UvrD proteins and was found to increase 7-1000 fold compared to wildtype cells [34, 114]. In *E. coli* lacking MutH, MutS, and MutL proteins, forward mutation studies in the *lacI* gene shows a 200-fold increase in mutation rate [99, 115, 116]. In *D. radiodurans* cells with inactivated MutS1, MutL, and UvrD proteins, forward mutation frequency studies using rifampicin demonstrated a 7-fold increase in mutation rate [34]. Interestingly, mutants lacking MutS2 did not show an increase in mutation rate, however *D. radiodurans* mutants defective in both MutS1 and MutS2 showed a 7-fold increase in mutation rate [34]. In *S. cerevisiae,* both forward and reverse mutation rate studies have shown that the MutS and MutL homologs are required for base correction [3, 4].

Other processes than the repair of mismatched bases following DNA replication have implicated MMR proteins. For example, MMR proteins have been implicated in the repair of base pair anomalies resulting from damage by alkylating agents and UV light. N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) creates a cytotoxic lesion, O^6-methyl

guanine, which can mispair with a thymine. The MMR proteins cannot correct this mispairing since the damaged guanine is on the template strand, which will initiate futile cycles of repair. Inactivation of the MMR system in *E. coli* and mammals allows bypass of these lesions, therefore, inactivation of the MMR system results in increased tolerance to MNNG in these organisms [57-59]. Interestingly, yeast deficient in MMR do not show an increase in tolerance to alkyating agents [3, 4]. This is likely because there is a O^6-methyl guanine methyltransferase (MGMT) in yeast that is responsible for repairing this type of alkylation modification [3, 4, 117].

While NER is the major pathway for repairing UV damage, MMR proteins have been shown to interact with the NER proteins in several ways [4]. Mammals and *E. coli* with an inactivated MMR pathway display a greater sensitivity to UV light than cells with an active MMR pathway [4, 118-120]. *D. radiodurans* and *S. cerevisiae* do not show an increase in sensitivity to UV light when their MMR machinery is inactivated, however in yeast deficient in both NER and MMR the decrease in survival is larger than yeast deficient in only the NER pathway suggesting that MMR may play an accessory role in NER [3, 4, 34].

Homologs of the bacterial MutS and MutL protein encoding genes have been found in the genome of *Halobacterium.* Through computational analysis we found that *Halobacterium* has a *zim* gene, which encodes a putative CTAG methylase, 3 bacterial-like *mutS* genes, a bacterial-like *mutL* gene, 4 bacterial-like *recJ* exonucleases, a eukaryotic-like *rad2* 5'-3'

exonuclease, and a bacterial-like *uvrD* helicase. These are all located on the main chromosome [79].

We constructed deletion mutants of the *mutL, mutS1, mutS2,* and *uvrD* genes as well as a *mutS1/mutS2* double mutant in *Halobacterium* using an in-frame gene deletion method described in Peck *et al* [121]. The phenotypes of the deletion mutants were characterized to look for tolerance to MNNG alkylating agents. Survival studies using UV-C light and gamma-ray were also conducted to further the phenotypic characterization of the mutants. Lastly, mutation frequencies were calculated for each of the mutant strains to determine if there was an increased rate of mutation that correlated with inactivation of the MMR proteins in *Halobacterium.* No tolerance to MNNG and no decrease in survival to UV-C light or gamma-ray was demonstrated by deleting the MMR homologs in *Halobacterium.* We also noted no increase in mutation frequency for any of the MMR mutants.

3.2 Materials and Methods

Strains and Growth Conditions:

Halobacterium cultures were grown in GN101 media [250g/L NaCl, 20g/L $MgSO_4$, 2g/L KCl, 3g/L sodium citrate, 10g/L Oxoid brand bacteriological peptone] with the addition of 1 mL/L trace elements solution [31.5mg/L $FeSO_4 \cdot 7H_2O$, 4.4mg/L $ZnSO_4 \cdot 7H_2O$, 3.3mg/L $MnSO_4 \cdot H_2O$, 0.1mg/L $CuSO_4 \cdot 5H_2O$] at 42°C shaking in a Gyromax 737 shaker (Amerex Instruments, LaFayette, CA) at 220rpm. The GN101 media was supplemented with 20g/L agar for solid plates. Basal salts solution (BSS), same composition as GN101 without the peptone, was used for culture dilutions. Uracil dropout media was made using

BSS supplemented with 10g/L yeast nitrogen base without amino acids (Sigma, St. Louis, MO) and 1.92g/L yeast synthetic dropout media without uracil (Sigma, St. Louis, MO). When specified, uracil, 5-FOA (Sigma, St. Louis, MO), and mevinolin (Sigma, St. Louis, MO) were added to a final concentration of 50mg/L, 300mg/L, and 50μM respectively.

In-frame gene deletions of *ΔmutL*, *ΔmutS1*, *ΔmutS2*, and *ΔuvrD* were constructed using the protocol of Peck *et al* [121] (See Figure 3-1A). Gene knockout constructs were composed of 500 base pairs upstream and downstream of the target gene. These constructs were ligated into the *Halobacterium* plasmid pNBK07 downstream of the constitutive ferredoxin promoter. The plasmid contains a mevinolin resistance gene as well as a functional *ura3* gene (See Figure 3-1B). The resulting plasmid was transformed into either a *Δura3* or *Δura3Δzim* background strain and insertion of the plasmid into the chromosome is selected for on uracil dropout plates [122]. Plating on uracil dropout plates will select only the cells containing a functional *ura3* gene ensuring plasmid integration. A second crossover event is selected for on 5-FOA-containing media, which will select against a functional *ura3* gene and removal of plasmid. Recombinants are screened by PCR to ensure deletion of the target gene. We designed primers starting at the beginning of the gene and ending 1000 base pairs outside of the gene (See Table 3-1). This ensures not only the deletion of the gene but also the correct orientation of the knockout. All deletions were confirmed by southern blot analysis. The GN101 media was supplemented with 50mg/L uracil for strains made from a *Δura3* background.

A. B.

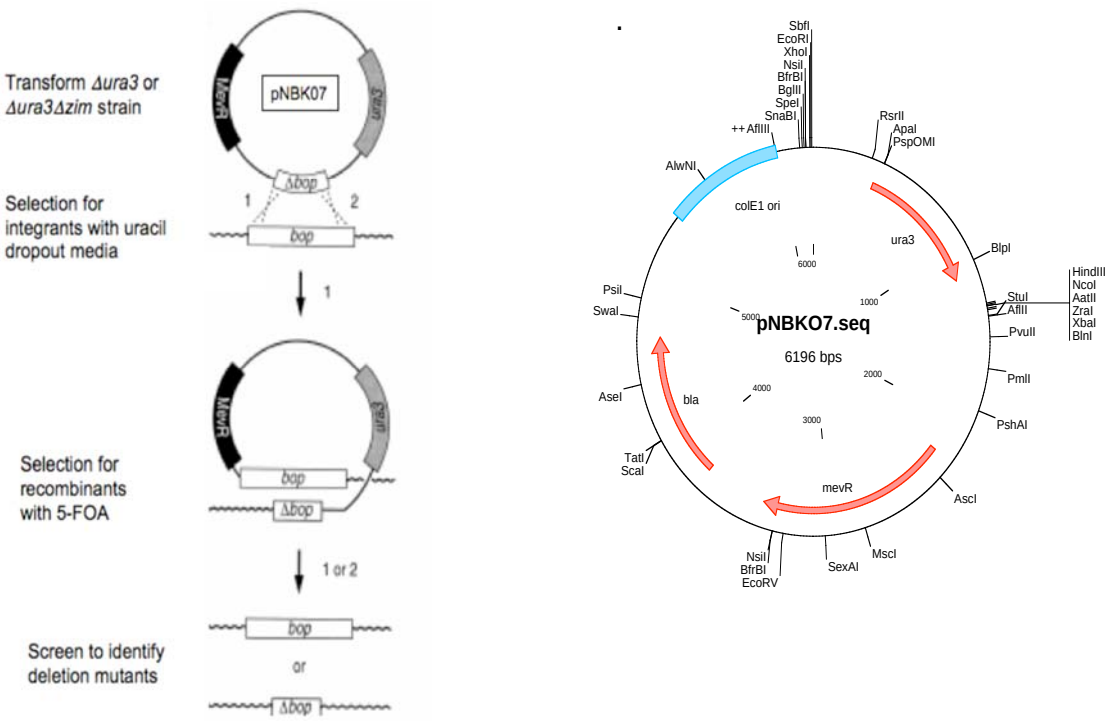


Figure 3-1. (A) Homologous recombination gene deletion scheme for *Halobacterium* [121]. The *bop* gene is used as an example. The *bop* gene was deleted from a *Δura3* background strain using plasmid pNBK07. This plasmid has a *ura3* gene, a mevinolin resistance gene, and the flanking regions (approximately 500 base pairs upstream and downstream) of the *bop* gene. Plasmid integration was selected for on uracil dropout media, which requires a functional *ura3* gene for colony growth. Crossover occurs by homologous recombination between the wildtype gene and deletion construct. Removal of plasmid was achieved by plasmid counterselection using 5-FOA against the *ura3* gene on the plasmid. Colony PCR was used to distinguish between wildtype and recombinant products. (B) Plasmid map of pNBK07.

Table 3-1. Primers used in the construction of in-frame gene deletions for the bacterial MMR homologs genes in *Halobacterium.* The restriction sites are underlined.

Gene	Primers
mutS1	F: 5’ CGCTCTAGATGCTGTTCGGGGACCGCCTCCCGAT 3’
	R: 5’ ACGAGGCCCACGACGTCGCGGCCACT 3’
mutS2	F: 5’ CGCTCTAGATTTCGGCATGCCGGTGGCGAGCT 3’
	R: 5’ TCGCCGCTCGTATCTACTCCTTGT 3’
mutL	F: 5’ CGCTCTAGAGGACGGTGTCGGTCATGTTGA 3’
	R: 5’ CCAGCGCGAACTACGCCCTCCTGTACT 3’
uvrD	F: 5’ CGCTCTAGATCAAAGTGCTGGTGAAGGCCT 3’
	R: 5’ ACGGTGGTGTCCTCGGGCGGGAA 3’

Methylase Activity of Zim:

A *Δzim* deletion mutant was constructed to test if the corresponding protein had a CTAG methylase activity *in vivo*. Wildtype and *Δzim* strains were grown to stationary phase and DNA extracted using phenol/chloroform [123]. Genomic and *Δzim* DNA (1μg) was digested with *BfaI* (New England Biolabs, Ipswich, MA). *BfaI* will only cut unmethylated DNA at the CTAG site. Digested DNA was visualized using a 1% agarose gel run at 100V/cm for 25 minutes.

Survival Assays:

Deletion mutants were tested for growth defects at 37°C, 42°C, and 45°C. Single colonies were grown to midlog phase in GN101/50mg/L uracil. Cultures were diluted to 10^4, 10^3, and 10^2 cells/mL in GN101/50mg/L uracil and 5μL of each spotted on square GN101/50mg/L uracil plates in triplicate. Plates were grown at respective temperatures for seven days and colony were observed for growth.

MNNG survival was calculated as the optical density (OD) at 600nm of the nontreated culture divided by that of the treated cultures. Cells were grown up to an OD_{600} 0.6 in GN101/50mg/L uracil and diluted back to an OD_{600} 0.2. Cultures were divided into 5mL aliquots and treated with the addition of 0, 50, 100, or 400mg/L MNNG in triplicate. Cultures were incubated until the wildtype control reached an OD_{600} 0.8, approximately 21 hours, at 42°C shaking in a Gyromax 737 shaker (Amerex Instruments, LaFayette, CA) at 220rpm in the dark. MNNG (TimTech, New Zealand) stock solution was made at 10g/L in DMSO.

UV survival experiments were done as follows: cultures were grown up to an OD_{600} 0.6 in GN101/50mg/L uracil, 1mL of cells were distributed in a monolayer into 4 well plates (24mm x 67mm), and irradiated with 200J/m^2 UV-C light (UVP Pen-Ray). Cells were then diluted in BSS and plated on GN101/50mg/L uracil plates in triplicate in the dark to prevent photolyase activity. Plates were incubated at 42°C in the dark for 5-7 days and colonies counted. Survival was calculated as Ni/No where Ni is the number of viable cells after UV-C irradiation and No is the number of viable cells without UV-C irradiation.

Gamma survival experiments were done similarly to the UV-C survival experiments. Mutants were grown up to an OD_{600} 0.6 and aliquoted into 1mL amounts in triplicate. Cells were irradiated with 2.5kGy of gamma-ray, measured by a Omega Engineering Model HH611PLA4F Type logging thermometer using a 26,000 curie ^{60}C source located at the University of Maryland College Park Gamma Test Facility at a dose rate of 62.01Gy/min. Both irradiated and control cells were diluted in BSS, plated on GN101/50mg/L uracil plates, and incubated for 5-7 days at 42°C. After incubation, colonies were counted and survival calculated.

Drug Testing for Mutation Frequency:

Various drugs were tested in order to assess their potential for use in a mutation frequency assay similarly to what was done for wildtype *Halobacterium* in Chapter 2. These drugs included G418, canavanine, novobiocin, chloramphenicol, 5-FU,

paromomycin, sparsomycin, anisomycin, coumermycin, aphidicolin, and 5-fluoroanthranilic acid (5-FAA) (Sigma Aldrich, St. Louis, MO). To test if *Halobacterium* was sensitive to these drugs, liquid minimum inhibitory concentrations (MIC) were determined as follows. Single colonies were grown in GN101/50mg/L uracil to midlog phase, diluted back to an OD_{600} 0.10, and a range of drug concentrations added in triplicate. A culture with no drug added was used as the control. Cultures were grown at 42°C shaking until control culture reached an OD_{600} 0.80. OD's were taken every two hours and growth charted. The liquid MIC was determined as the lowest dose of drug that completely inhibited growth of *Halobacterium*. Drug stocks were made as follows: 10% w/v in water for G418, 2g/L in water for canavanine, 2g/L in water for novobiocin, 34g/L in water for chloramphenicol, and 50g/L in ethanol for 5-FAA.

After liquid MIC's were determined for novobiocin and 5-FAA, optimum doses of these drugs for plating assays was assayed. Midlog phase cultures were plated on GN101 plates supplemented with either 400-600mg/L 5-FAA or 1-5mg/L novobiocin. Plates were grown at 42°C for 7 days and colonies observed for growth. Optimization of drug concentration needed for 5-FAA proved difficult using this method, so as an alternative, dilution spotting assays were carried out. Cultures were grown up to midlog, diluted to 10^8, 10^6, and 10^4 cells/mL, and 5μL spotted on plates in triplicate. Plates were made using a range of concentrations from 0-400mg/L 5-FAA. Plates were grown at 42° for 11 days and colony growth observed.

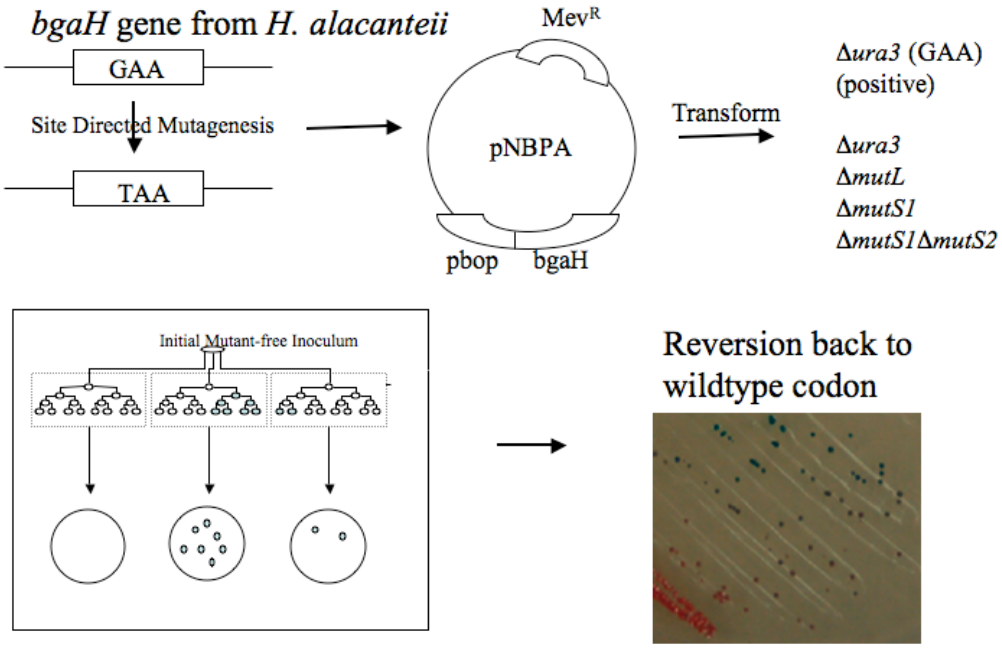
Mutation Frequency Assays:

Mutation frequency assays were originally attempted using novobiocin and 5-FAA following the fluctuation test protocol from Chapter 2. Alternative methods of calculating mutation frequency were developed after novobiocin and 5-FAA proved unreliable. To calculate the mutation frequencies for the mutant strains, all of which are in a *Δura3* background, using 5-FOA was not a feasible option. 5-FOA works through the UMP biosynthesis genes in the pyrimidine metabolism pathway. The *ura3* gene is one of the UMP biosynthesis genes and deletion of this gene renders cells resistant to 5-FOA. Two new *in vivo* mutator assays were developed to work around this problem and to allow comparison between the methods.

β-galactosidase method:

The first mutation frequency assay utilized the β-galactosidase gene for blue/pink selection (Figure 3-2A). A *Δura3* strain of *Halobacterium* containing the PNBpbop_bgaH plasmid was a kind gift from Nitin Baliga's lab at the Institute for Systems Biology in Seattle Washington. This plasmid contains the β-galactosidase (*bgaH*) gene from *Haloferax alacanteii* downstream of the *bop* promoter, a strong promoter activated by light, and a mevinolin resistance gene to ensure maintenance of plasmid in *Halobacterium*. We amplified the *bgaH* gene along with *bop* promoter from this plasmid using PCR (bop F: 5'CGCAAGCTTGACGTGAAGATGGGGCTCCCG3'; bgaH R: 5'CGCGGACTCCGGATCCTCTAGTCCATCGCCG3') and cloned the construct (restriction sites are underlined) into *Halobacterium* plasmid pNBPA in place of the ferredoxin promoter and pA tag (See Figure 3-2B). This plasmid was transformed

A.



B.

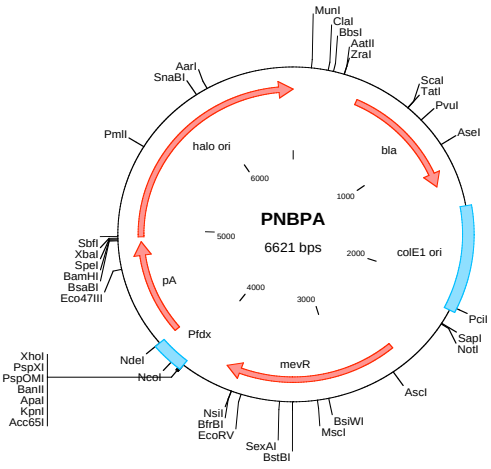


Figure 3-2. (A) Experimental design for the β-galactosidase *in vivo* assay. A construct containing the *bgaH* gene from *H. alacanteii* was subjected to site-directed mutagenesis, changing the catalytic amino acid from a GAA to a TAA. Both constructs were cloned into *Halobacterium* plasmid pNBPA behind the *bop* gene promoter. This plasmid contains a *Halobacterium* origin of replication a mevinolin resistance gene to ensure maintenance of the plasmid in the cells. The plasmids were transformed into the *Δura3*

background strain and the *ΔmutL*, *ΔmutS1*, and *ΔmutS1ΔmutS2* mutant strains. Fluctuation tests were carried out and cells plated on media containing mevinolin. Reversion back to a functional *bgaH* gene product (blue colonies) was scored. (B) Plasmid map of *Halobacterium* plasmid pNBPA.

in the *Δura3* background strain of *Halobacterium. Halobacterium* colonies containing this construct were flashed with light for 4 hours to activate the promoter. Colonies were sprayed with 10g/L IPTG/X-gal mix (Fisher Scientific, Pittsburgh, PA) resulting in a blue color after incubation for 2 days at 42°C. An overview of the *in vivo* mutator assay can be seen in Figure 3-2A.

The *bgaH* gene contains putative catalytic residues at Glu141 and Glu312, which correspond to the *E. coli* catalytic residues from the *lacZ* gene (Glu461 and Glu537) [124, 125]. The *bgaH* gene was subjected to site directed mutagenesis changing the glutamic acid from a GAA to a TAA. We performed site-directed mutagenesis using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA). This kit consists of a mutant strand synthesis protocol, followed by digestion with *Dpn1* to remove parental DNA, and transformation into competent cells for nick repair. The resulting plasmid was transformed into the *Δura3 Halobacterium* strain and *ΔmutS1, ΔmutS1ΔmutS2*, and *ΔmutL* deletion strains and fluctuation tests performed. Colonies containing the mutated base pair were also flashed in light and sprayed with the IPTG/X-gal mix to ensure colonies remained pink in color indicating the catalytic domain was inactivated. The fluctuation test is described in Chapter 2 with the following changes: single colonies are grown in GN101/50mg/L uracil/50μM mevinolin and instead of plating cells on media containing 5-FOA, we plated cells on media containing mevinolin to ensure maintenance of the plasmid. We plated approximately 5000 cells per plate. Plates were incubated for 7-10 days at 42°C until colony growth was observed, sprayed with IPTG/X-gal mix, and incubated for an additional 2 days. Plates were resprayed with

IPTG/X-gal mix and incubated at 42°C for two days and colonies were scored to reversion back to the wildtype codon. Colonies that have reverted back to the wildtype (GAA) codon have active β-galactosidase activity and turn blue, colonies with the mutated codon do not have active β-galactosidase activity and remain pink.

5-FOA Method:

A general overview of this method is seen in Figure 3-3. We PCR amplified the *ura3* (*pyrF*) gene along with its native promoter (pura3) and cloned this construct into *Halobacterium* plasmid pNBPA in the place of the ferredoxin promoter and pA tag. The resulting plasmid was transformed into *Halobacterium* strains *Δura3* (used as wildtype), *ΔmutL, ΔmutS1,* and *ΔmutS1ΔmutS2.* Similarly to above, transformed cells were grown in GN101/50mg/L uracil/50μM Mevinolin and aliquoted into a 96 well plate. After growth for 3 days, cultures were spread *in toto* on GN101/50mg/L uracil/50μM Mevinolin/350mg/L 5-FOA plates. Resultant colonies were screened by PCR to ensure maintenance of plasmid. The addition of the *ura3* gene will render cells sensitive to 5-FOA, which allows determination of the mutation frequency of the mutants using this drug. Resistant colonies should contain a mutation in the *ura3* gene in order to grow in 5-FOA containing media and this was confirmed by DNA sequencing using the primers: ura3 Acc65I (5'GCGGGTACCGTCGGCTGGCGGGCACGCGGT3') and ura3 SpeI (5'GCGACTAGTCTACCGGTGGCGGTTCAGGCG3'). Mutation frequency was calculated as the ratio of resistant colonies to average number of colonies plated.

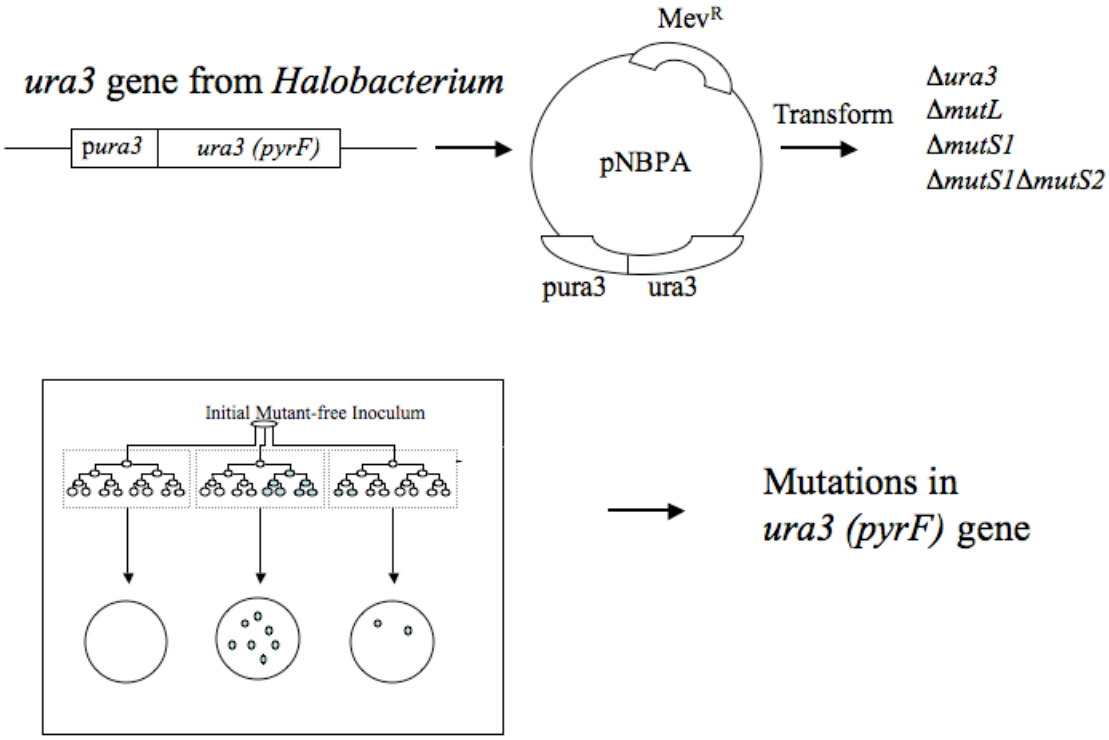


Figure 3-3. Experimental design for 5-FOA *in vivo* mutation frequency assay. The *ura3* gene along with its native promotor was cloned into *Halobacterium* plasmid pNBPA. This plasmid contains a *Halobacterium* origin of replication and the mevinolin resistance gene to ensure maintenance of the plasmid in cells. The resulting plasmid was transformed into the *Δura3* background strain as well as the *ΔmutS1*, *ΔmutS1ΔmutS2*, and *ΔmutL* mutant strains. Fluctuations tests were performed and cells were plated on media containing both 5-FOA and mevinolin. Only cells with mutations in the *ura3* gene will grow in medium containing 5-FOA and the mutation frequency of these cells was calculated as the number of mutant colonies divided by the number of cells plated.

3.3 Results

The objective of these experiments was to determine if the bacterial-like MMR proteins found in *Halobacterium* are essential for the low incidence of mutation observed in this organism. Deletion mutants *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2, ΔmutL,* and *ΔuvrD* were successfully constructed using the method of Peck *et al* [121] in both a *Δura3* and *Δura3Δzim* background. All mutants were checked both with PCR and Southern blot analysis to confirm complete gene deletion (See Figure 3-4). Furthermore, both background strains of *Halobacterium* showed identical phenotypes when the MMR homologs were removed verifying the construction of true gene deletions.

We confirmed the methylase activity of the *zim* gene product by digestion with *BfaI,* a restriction enzyme that will only cut unmethylated DNA at a CTAG site and analyzed the product by agarose gel electrophoresis. The *Δzim* DNA is unmethylated at CTAG sites and hence digested by *BfaI* whereas wildtype DNA is methylated and remained intact after digestion. We showed no tolerance to MNNG, no decreased survival to UV-C and gamma irradiation, and no increase in mutation frequency of the *Δzim* mutant compared to the wildtype strain.

We characterized the phenotypes of the *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2, ΔmutL,* and *ΔuvrD* mutants to compare with other MMR deficient strains of organisms. Deletion mutants were tested for growth defects at 37, 42, and 45°C. No defect in growth was observed. Survival assays were done on the deletion strains using MNNG, UV-C irradiation, and gamma-ray. No change in tolerance to MNNG at 50, 100, and 400mg/L was observed (See Figure 3-5). *Halobacterium* showed an 80, 70, and 25% survival to

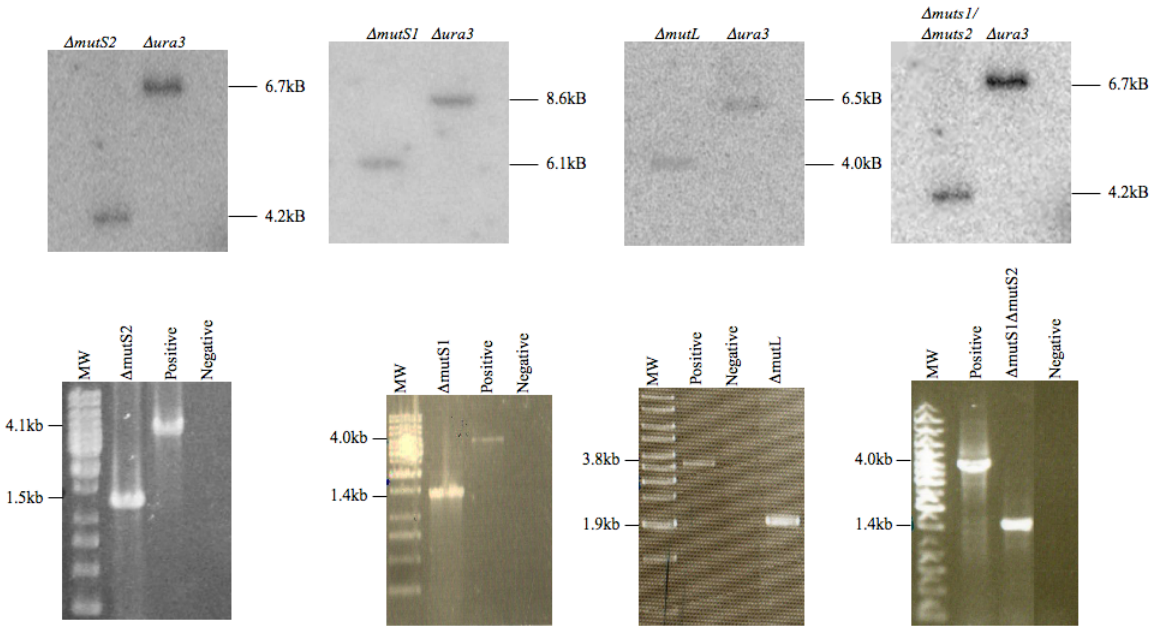


Figure 3-4. Southern hybridizations (top) and PCR analysis (bottom) showing gene deletions for *ΔmutL, ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2,* and *ΔuvrD*. Probes were labeled with ^{32}P and were designed to hybridize to regions 500 base pairs downstream of the target genes coding region. PCR analysis was done for each mutant to ensure complete gene deletion. Primers began at the start codon for the targeted gene and ended 500 base pairs past the stop codon. The positive lane contained wildtype *Halobacterium* DNA, the negative lane contained no DNA, and the sample lanes contained mutant DNA. Agarose gels (1% TAE) were run at 100V/cm for 25 minutes and visualized with ethidium bromide.

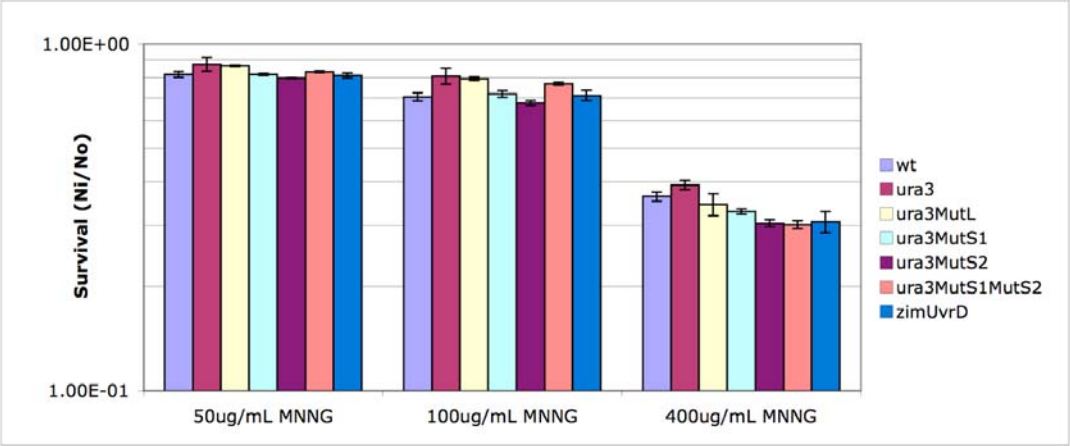


Figure 3-5. Survival of *Halobacterium* (wt), background strain for deletion mutants (*Δura3*), and the mutant strains, *ΔmutL*, *ΔmutS1*, *ΔmutS2*, *ΔmutS1ΔmutS2*, and *ΔuvrD* to 50, 100, and 400μg/mL MNNG. Survival was calculated as the average ratio (Ni/No) of the viable cells in the challenged sample (Ni) compared to the viable cells in the unchallenged sample (No). Error bars represent standard error for three independent replicates.

these doses of MNNG. No decrease in survival was noted with any of the deletion strains using UV-C light at 200J/m^2 or 2.5kGy gamma irradiation (See Figures 3-6 and 3-7). These doses of UV-C light and gamma-ray show a 80% survival in *Halobacterium.*

The frequency of mutation for the deletion strains *ΔmutS1*, *ΔmutS1ΔmutS2*, and *ΔmutL* was determined to further characterize the phenotype of these deletion strains. Drugs were tested for efficacy in a mutation frequency assay in *Halobacterium;* however, *Halobacterium* proved to be resistant to most of the drugs tested (See Table 3-2). *Halobacterium* was sensitive to novobiocin, 5-FU, and 5-FAA but further studies with these drugs proved inconsistent.

Due to the difficulties in finding a suitable candidate drug for a forward mutation assay, we developed two *in vivo* mutator assays to calculate mutation frequency in the *Δura3* background strain, and the *ΔmutL, ΔmutS1, ΔmutS1ΔmutS2* mutant strains. We cloned the *bgaH* gene from *H. alacantei* into the *Halobacterium* plasmid pNBPA. The catalytic amino acid was mutated and we measured the reversion back to the wildtype codon by looking for cells that produced a functional β-galactosidase and therefore blue colonies with the addition of IPTG and X-gal solution. Based on the *Halobacterium* mutation rate, 3.43×10^{-3} per genome per replication, determined in Chapter 2, we expect to see reversion back to wildtype in one out of every $1x10^5$ cells. However, that was calculated for a target size of approximately 600 base pairs, which is the average size for the three UMP biosynthesis genes. In our mutator assay, the target size was only one base pair,

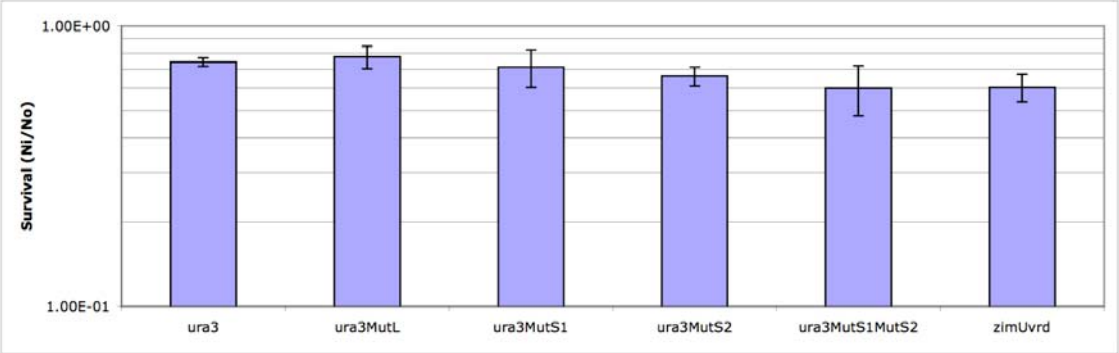


Figure 3-6. Survival of *Halobacterium* background strain (*Δura3*) and the mutant strains, *ΔmutL, ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2,* and *ΔuvrD,* to 200J/m^2 UV-C light. Survival was calculated as the average ratio (Ni/No) of the viable cells in the challenged sample (Ni) compared to the viable cells in the unchallenged sample (No). Error bars represent standard error for three independent replicates.

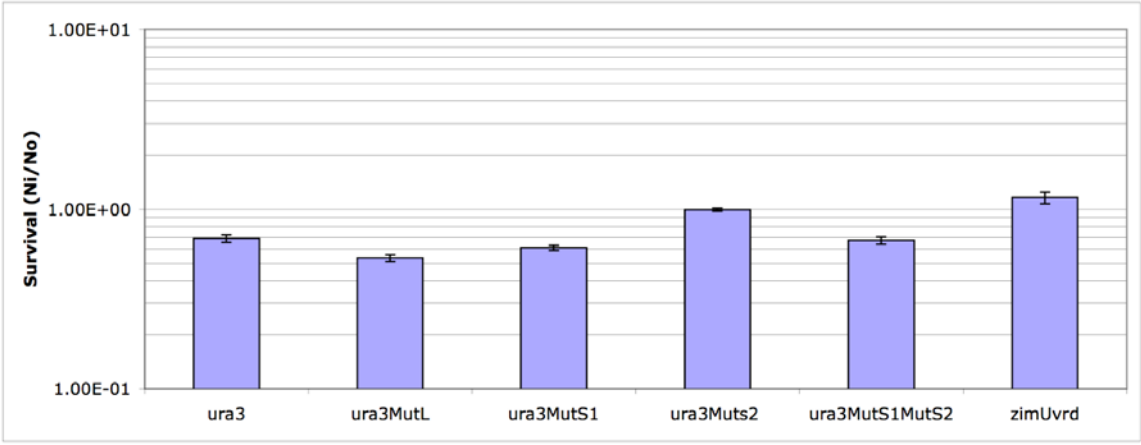


Figure 3-7. Survival of *Halobacterium* background strain (*Δura3*) and the mutant strains, *ΔmutL*, *ΔmutS1*, *ΔmutS2*, *ΔmutS1ΔmutS2*, and *ΔuvrD*, to 2.5kGy gamma-ray. Survival was calculated as the average ratio (Ni/No) of the viable cells in the challenged sample (Ni) compared to the viable cells in the unchallenged sample (No). Error bars represent standard error for three independent replicates.

Table 3-2. Drugs tested for possible use in a mutation frequency assay for *Halobacterium*.

Drug	Concentration Tested	Outcome
G418	50-700mg/L	Resistant
Canavanine	1-50mg/L	Resistant
Chloramphenicol	20-150mg/L	Resistant
Paromomycin	100-850mg/L	Resistant
Novobiocin	0.5-9mg/L	Sensitive – only one base pair change can cause resistance
5-FU	50-75mg/L	Sensitive – inconsistent results
5-FAA	100-600mg/L	Sensitive – inconsistent results

which decreases the expected mutation rate by about 600-fold. However, if deletion of the MMR genes produced a mutator phenotype with an increase of 10-1000-fold as was found in Bacteria, we would expect to see reversion mutants of the *bgaH* gene in one out of every $6 \times 10^6 - 6 \times 10^4$ cells. We plated approximately 1×10^6 cells and did not detect any blue colonies indicating reversion to a functional *bgaH* gene product. These experiments were completed at least 3 times and indicated that the mutation frequency in *Halobacterium* did not drastically increase with the deletion of the bacterial-like homologs of the MutS and MutL MMR genes.

To confirm this observation and calculate the mutation frequency of the *ΔmutS1, ΔmutS1ΔmutS2*, and *ΔmutL* mutants, we developed a second *in vivo* mutator assay. The *ura3* gene along with its native promoter was cloned into the *Halobacterium* plasmid pNBPA. This plasmid has a *Halobacterium* origin of replication and can be maintained in the cells with the addition of mevinolin (gift of Dr. Nitin Baliga). The *ura3+* construct was transformed into the *Δura3* strain (background), and in the *ΔmutL, ΔmutS1,* and *ΔmutS1ΔmutS2* mutant strains. Three independent fluctuation tests were run for *Δura3* and *ΔmutS1ΔmutS2* mutants and four tests were run with *ΔmutL* and *ΔmutS1* mutants. These mutants were plated on 5-FOA containing media to identify mutations in the *ura3* (*pyrF*) gene. Mutations in the *ura3* gene confer resistance to 5-FOA and pyrimidine synthesis is done through the direct uptake of uracil [107]. We saw no increase in mutation frequency for any of the mutants (See Figure 3-8). We also sequenced *ura3* mutants from the *ΔmutL, ΔmutS1,* and *ΔmutS1ΔmutS2* background strains obtained from the assay to (1) ensure that there were mutations on the *ura3* gene in the plasmid and (2)

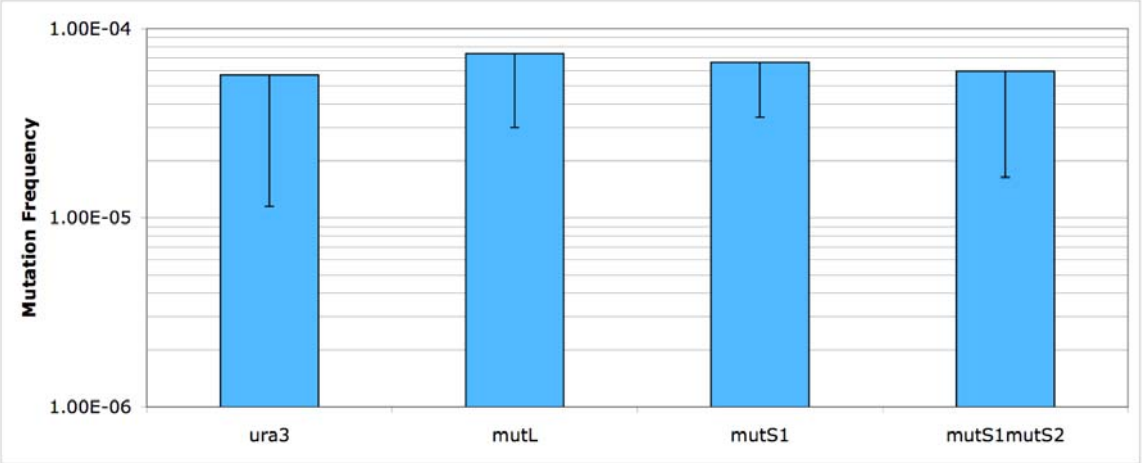


Figure 3-8. Mutation frequencies of the MMR deletion mutants *ΔmutS1, ΔmutL,* and *ΔmutS1ΔmutS2* in *Halobacterium* as well as the *Δura3* background strain. Using the 5-FOA *in vivo* assay, the background and mutant strain were grown up and aliquoted into a 96 well plate. After several days of growth, cells were plated on 5-FOA containing media which selects for mutations in the *ura3* gene. The mutation frequency was calculated as the ratio of mutant cells compared to the average number of cells plated. Standard error was calculated from three independent replicates.

to compare the mutational spectrum with that of *Halobacterium.* The sequencing data correlated very nicely with what was seen for the *pyrF* (*ura3*) gene in *Halobacterium* discussed in Chapter 2. We sequenced 20 mutants in the *Δura3* background strain, 33 mutants in the *ΔmutL* deletion strain, 22 mutants in the *ΔmutS1* deletion strain, and 19 mutants in the *ΔmutS1ΔmutS2* deletion strain. Approximately 60% of the mutants sequenced had changes in the *ura3* gene and they had very similar numbers of BPS, insertions, and deletions to the *ura3* gene in *Halobacterium* (See Figure 3-9).

3.4 Discussion

The objective of this study was to characterize the cellular function of the MutS1, MutS2, MutL, and UvrD proteins in *Halobacterium.* These genes are located on the main chromosome of *Halobacterium* [79]. If these bacterial-like MMR homologs are essential to maintain a low genomic mutation rate in *Halobacterium,* we should see a decrease in survival to UV-C irradiation, an increase in tolerance to MNNG, and an increase in mutation frequency, similar to the phenotypes seen in Bacteria and Eukarya.

In-frame deletions of the bacterial homologs of MMR genes were constructed in *Halobacterium* using both *Δura3* and *Δura3Δzim* background strains. From previous studies investigating the global transcriptional responses of *Halobacterium* to UV-C light, we showed a downregulation of the *zim* gene possibly indicating an undermethylation of DNA [84]. We originally hypothesized that the Zim protein could be the strand discrimination signal in *Halobacterium* similar to the *dam* methylase in *E. coli* and since mutant construction in *Halobacterium* can take a significant amount of time,

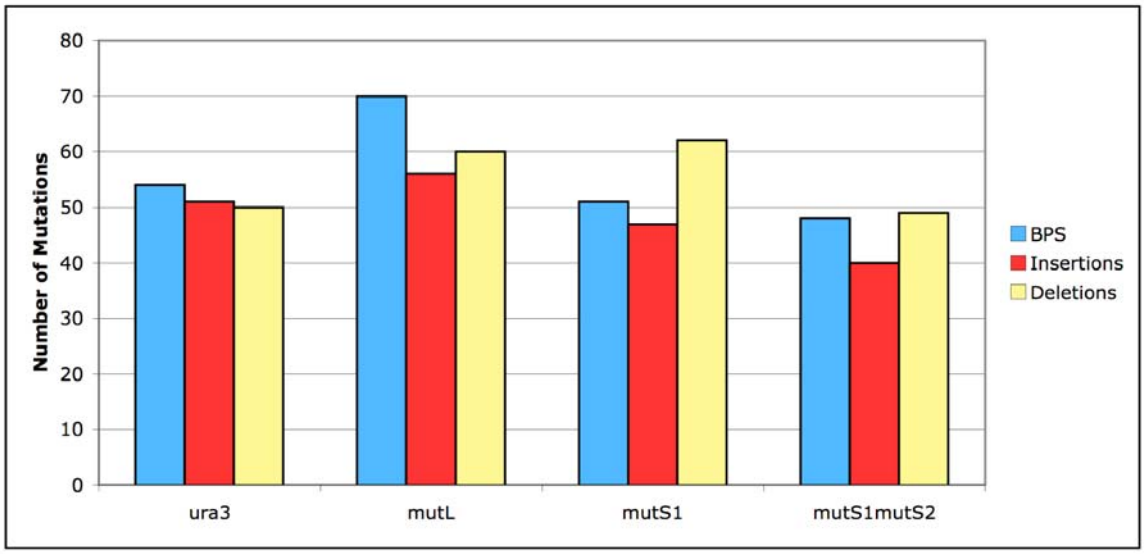


Figure 3-9. Distribution of BPS, insertions, and deletions within the pNBPA encoded *ura3* gene for the background (*Δura3*) strain and the *ΔmutL*, *ΔmutS1*, and *ΔmutS1ΔmutS2* deletion strains. We sequenced 20 *ura3* mutants in the *Δura3* background strain and found 12 mutants with 155 mutations. Of these mutations, 54 were BPS, 51 were insertions, and 50 were deletions. Thirty-three mutants were sequenced in the *ΔmutL* mutant strain and 20 mutants had 186 mutations. Of these mutations, 70 were BPS, 56 were insertions, and 60 were deletions. We sequenced 22 mutants in the *ΔmutS1* mutant strain and found 15 mutants with 160 mutations. Of these mutations, 51 were BPS, 47 were insertions, and 62 were deletions. Lastly, we sequenced 19 mutants in the *ΔmutS1ΔmutS2* mutant strain and found 10 mutants with 137 changes including 48 BPS, 40 insertions, and 49 deletions.

we constructed mutants in both background strains to ensure completeness if Zim was indeed the strand discrimination signal. We concluded that Zim is not the strand discrimination signal for MMR in *Halobacterium* based on mutant analysis. Based on sequence comparison to a methylase in *Methanothermobacter thermautotrophicus*, we concluded that Zim is probably part of a restriction modification system even though the corresponding restriction enzyme was not found in *Halobacterium* [126].

We saw no decrease in survival to UV-C irradiation or gamma-ray and no increase in tolerance to alkylation with MNNG with any of the deletion mutants. These results, while surprising, were not totally unexpected. In *D. radiodurans* and *S. cerevisiae* deficient in MMR, there was no decrease in survival to UV-C irradiation or MNNG. Mennecier *et al* demonstrated that MutS and MutL deficient *D. radiodurans* displayed wildtype levels of resistance up to 1500 J/m^2 UV light and up to 150ng/mL mitomycin C (alkylating agent) [34]. In *S. cerevisiae,* no tolerance to MNNG was demonstrated in MMR deficient yeast strains unless the *mgt1* methyltransferase responsible for correcting O^6-methyl guanine damage was also absent [117]. No decrease in survival to UV light was seen up to 70J/m^2 in MMR deficient yeast, however, there was an additive effect when both the MMR and NER pathways were inactivated [54, 127]. Studies in *Halobacterium* have shown that damage from UV light is repaired mostly by a photolyase in the light and by the bacterial homologs of the UvrA/B/C system in the dark [88]. We were unable to test the effect of UV light on *Halobacterium* strains deficient in both NER and MMR to see if there was an additive effect similar to *S. cerevisiae.* It is possible that there would be a decreased survival to UV light in *Halobacterium* with both the NER and MMR systems removed if

the MMR system is playing a role in NER. Based on computational analysis, *Halobacterium* does not have a homolog to the MGT1 methyltransferase protein found in *S. cerevisiae* but that does not rule out the possibility that *Halobacterium* has other glycosylases that can repair alkylation damage.

In previous studies, proteins involved in MMR in *E. coli* were isolated in screens looking for spontaneous mutator phenotypes. MutS, MutL, and MutH mutants were found to have a 10-1000-fold increase in spontaneous mutation rate [3-5]. Defects in MMR in the Eukarya also lead to increase in spontaneous mutation [4, 5]. Forward mutation rate studies for canavanine resistance in *S. cerevisiae* showed a 18-40-fold increase in spontaneous mutation in strains deficient in MSH2, MSH6, MLH1, and PMS1 [38, 39]. In reversion assays looking for lysine and threonine revertants, there was a 3-55-fold increase for the MutS homologs and a 1000-fold increase for the MutL homologs [4, 38]. To further characterize the MMR mutants in *Halobacterium*, we calculated the frequency of mutation in the *Δura3* background strain, and the *ΔmutL*, *ΔmutS1*, and *ΔmutS1ΔmutS2* mutant strains. We developed two different *in vivo* mutator assays to calculate the mutation frequency in *Halobacterium* MMR mutants.

The first *in vivo* mutator assay using the *bgaH* gene from *H. alacanteii* did not show any reversion from the mutated codon of the *bgaH* gene to the wildtype codon. There are several possibilities why we were unable to see reversion back to wildtype the most obvious reason being that the bacterial MMR homologs in *Halobacterium* are not involved in MMR. We had speculated that a 10-1000-fold increase in mutation rate

would be visible if we plated 1 x 10^6 cells. Since no blue colonies were seen, we concluded that if our mutation frequency increases, it is less than a 10-fold increase. The second problematic area is that it is possible for a mutation to occur that will revert back to wildtype but have another mutation that prevents the colonies from turning blue. This could underestimate the mutation frequency. A third possibility is that *Halobacterium* has an archaeal-specific MMR system that can compensate for the loss of the bacterial-like MMR homologs. This means that the bacterial-like system is not essential but could still be functional. Due to the difficulties with calculating a mutation frequency using the above assay, we developed a second assay.

The second *in vivo* mutator assay was used to calculate the mutation frequency in the background strain, *Δura3,* and the deletion strains, *ΔmutL, ΔmutS1, ΔmutS1ΔmutS2.* In this assay we used a functional *ura3* gene carried on a *Halobacterium* plasmid to select for mutants resistant to 5-FOA. There was no increase in mutation frequency when the MutS and MutL MMR homologs were deleted when compared to the background strain. This correlates with the results from the *in vivo* mutator assay using the *bgaH* gene. Sequencing results of the *ΔmutL, ΔmutS1,*and *ΔmutS1ΔmutS2* mutant show a correlation with mutations seen in the *ura3* (*pyrF*) gene in wildtype *Halobacterium.* In *Halobacterium,* we calculated that approximately 60% of the mutants sequenced had mutations in the *pyrF* gene (See Chapter 2). Looking at mutations in the plasmid encoded *ura3* clones in the *Δura3* background strains and in the MMR deletion strains we also determined that between 50-70% of the mutants sequenced had mutations in the *ura3* gene. The rest of the mutations may have been in the *pyrE1* or *pyrE2* genes on the

chromosome of *Halobacterium*. The mutations seen in the background and deletion strains have about equal numbers of BPS, insertions, and deletions, which is what was seen in wildtype *Halobacterium.*

Our findings suggest that the bacterial-like MMR pathway in *Halobacterium* is not essential. Previous microarray analyses showed that the MMR homologs were transcribed but were not upregulated with DNA damage from UV-C light and gamma irradiation [84, 85]. In light of this and from the genomic mutation rate analysis in Chapter 2 showing *Halobacterium* has a low incidence of mutation, we developed two hypotheses about what is responsible for maintaining a low incidence of mutation in *Halobacterium*: (1) a high fidelity polymerase that could result in few mismatches reducing the need for a MMR pathway or (2) an archaeal-specific MMR system based on the recruitment of other enzymes such as glycosylases.

Most Archaea do not possess homologs to the bacterial MutS and MutL proteins but studies looking at genomic mutation rates have shown that Archaea have a low genomic mutation rate [67, 68]. Many Archaea are found in extreme environments, which lead to a high risk of damage to the cellular components. Because of this many organisms living in extreme environments have evolved adaptive mechanisms to protect and repair damages by the environment. For example, archaeal DNA polymerases that have been characterized from a hyperthermophile exhibit 10-fold higher fidelity when compared with that of Bacteria and Eukarya. *P. furiosus,* a member of the hyperthermophilic archaea, has been shown to have a high fidelity polymerase (See Chapter 2) [72, 74].

Further evidence for this hypothesis is the high frequency of insertions and deletions and the low frequency of BPS seen in Chapter 2. The higher fidelity of archaeal polymerases along with structural and sequence differences in the polymerases could participate in maintaining the low incidence of mutation hence resulting in a lower incidence of mispaired bases following DNA replication.

The second hypothesis is that the Archaea do have some type of MMR but that it could be based on the recruitment of other enzymes rather than the canonical MMR homologs to correct base pair mismatches. The deamination of a cytosine to uracil in a GC base pair is a major mutagenic event, which generates a GC to AT mutation [2]. Polymerases in the Archaea possess the unique ability to stall when a uracil residue is encountered [128]. The uracil is then removed by a uracil-DNA glycosylase [128, 129]. Direct interaction between uracil-DNA glycosylase (UDG) and a PCNA homolog from *Pyrobaculum aerophilum* has been documented [129-131]. This indicates the possibility of glycosylases removing damaged bases that could otherwise result in mispairing. *Halobacterium* has homologs to UDG, uracil-DNA glycosylase, MutY, adenine glycosylase, Ogg, 8-oxo-guanine glycosylase, NtH, endonuclease III glycosylase, XthA, AP endonuclease, and AlkA, 3-methyl adenine DNA glycosylase [79]. Further strengthening this hypothesis is the similar survival seen between wildtype *Halobacterium* and MMR deletion strains after treatment with MNNG. This type of damage creates a base pair mismatch that is normally corrected through MMR but can also be corrected through the actions of an MGT1 methyltransferase. Since strains missing the MMR homologs showed no decrease in survival as compared to the

background strain, it is possible that *Halobacterium* may also have a MGT1 methyltransferase to correct this damage. Recruitment by the DNA polymerase or PCNA of glycosylases to the site of mispaired bases could function as the MMR pathway in *Halobacterium.* In addition to the recruitment of glycosylases to fix mismatched bases, the homologous recombination pathway may also play a role. The homologous recombination pathway is responsible for the repair of double strand DNA breaks and the proteins have been shown to be involved in many other repair activities [132]. The MutS3 protein is thought to play a role in homologous recombination in *H. pylori* and it is possible that the MutS3 protein can interact with the replication machinery in *Halobacterium.*

Chapter 4: Oxidative stress response in *Halobacterium*

4.1 Introduction

In environments characterized by extreme conditions, such as high temperature and salinity, archaea dominate the microbial population [133]. Halophilic archaea possess a range of mutation avoidance and repair systems to endure high levels of solar radiation, extreme salinity (up to 4.5M), and cycles of rehydration and dessication [84, 86]. In previous studies, *Halobacterium* has shown resistance to high levels of UV and gamma radiation [84, 86]. In addition, whole-genome studies of transcriptional responses in *Halobacterium* have been studied using these conditions [84, 85]. Gamma irradiation can induce severe DNA damage, such as nucleotide modification and DNA strand breaks, both directly and indirectly [98, 134]. More than 80% of the damage caused by gamma irradiation is the indirect result of the radiolysis of water into hydroxyl radicals [97]. Only 20% of the damage is caused by the direct effects of the photons [97]. This study showed a downregulation of several dehydrogenases involved in the TCA cycle as well as a putative cell division cycle ortholog [85]. Also shown was an upregulation in the mRNA transcripts of genes involved in homologous recombination and BER [85]. In addition to gamma irradiation, the introduction of hydrogen peroxide (H_2O_2) and paraquat can also cause oxidative stress. Our study looking at the oxidative responses to H_2O_2 and paraquat, along with the gamma study, will allow a broader look at the global response of *Halobacterium* to different types of oxidative stress.

The instability of ROS poses a serious threat to aerobic organisms. ROS are produced through normal aerobic metabolism and environmental stresses. The three most

damaging ROS are superoxide ($O_2^{\bullet}$), H_2O_2, and the hydroxyl radical ($HO^{\bullet}$) [97, 135]. These ROS can damage DNA, proteins, membrane lipids, and carbohydrates through a variety of different reactions [135]. Generation of ROS can be amplified by the presence of transition metal ions, for example Fe^{2+}. Through the Fenton and Haber-Weis reactions, H_2O_2 can be converted into a reactive hydroxyl radical [97, 135].

Fenton Reaction $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + HO^{\bullet} + HO^-$

Haber-Weis Reactions $H_2O_2 + HO^{\bullet} \rightarrow H_2O + O_2^{\cdot} + H^+$

$H_2O_2 + O_2^{\bullet} \rightarrow O_2 + HO^- + HO^{\bullet}$

Another oxidative damaging agent is paraquat. Paraquat, N,N'-dimethyl-4,4'bipyridinium dichloride, is a viologen that generates superoxide radicals. Superoxides can oxidize iron sulfer clusters, which destabilize protein structure and release free Fe^{2+} [97, 136]. This Fe^{2+} can then react with H_2O_2 to form hydroxyl radicals through Fenton chemistry [136].

Aerobic organisms are equipped with mechanisms of defense against ROS including the induction of an SOS response, a upregulation of DNA repair genes and ROS scavenging enzymes, in extreme oxidative stress and the presence of scavenger enzymes such as catalases and superoxide dismutases [97, 135, 137-139]. Catalases convert hydrogen peroxide into oxygen and water whereas superoxide dismutases will convert superoxide

into H_2O_2 and oxygen [140, 141]. The H_2O_2 can then be converted into oxygen and water again by a catalase.

Halobacterium is able to grow under aerobic conditions in hypersaline environments but is able to switch to anaerobic facultative metabolism when the availability of oxygen is reduced [142]. In addition, *Halobacterium* has several mechanisms in place for protection against oxidative damage. The sensory rhodopsins are involved in the movement of cells away from high energy wavelengths [143]. Carotenoids found in the cell membrane of *Halobacterium* have been shown to scavenge free radicals and this is also seen in *Deinococcus radiodurans* [86, 144]. In addition to the sensory rhodopsins and carotenoids, the genome of *Halobacterium* has several catalases, peroxidases, and superoxide dismutases whose products could remove H_2O_2 and superoxide [145]. Studies have also hypothesized about the protection against oxidative damage by the high intracellular concentration of KCl [86, 146]. We will elucidate the oxidative stress response of *Halobacterium* to H_2O_2 and paraquat by whole genome transcriptional analysis. Furthermore, we will analyze gene deletion mutants suggested to participate in ROS scavenging.

4.2 Materials and Methods

Organism and Growth Conditions

Halobacterium was grown in GN101 media [250g/L NaCl, 20g/L $MgSO_4$, 2g/L KCl, 3g/L sodium citrate, 10g/L Oxoid brand bacteriological peptone] with the addition of 1 mL/L trace elements solution [31.5mg/L $FeSO_4 \cdot 7H_2O$, 4.4mg/L $ZnSO_4 \cdot 7H_2O$, 3.3mg/L

$MnSO_4 \cdot H_2O$, 0.1mg/L $CuSO_4 \cdot 5H_2O$] at 42°C shaking at 220rpm (Gyromax737, Amerex Instruments, Layfayette, CA). The GN101 media was supplemented with 50mg/L uracil for strains constructed from a *Δura3* background and with 20g/L agar for solid media. BSS had the same composition as GN101 but without the peptone and was used for culture dilutions.

Halobacterium wild type strains and deletion mutants constructed as described before [121] were used in this study. The rhodopsin-deficient strain, *pho81*, has been characterized previously [147].

Exposure to H_2O_2 and Paraquat

Two time courses were run to determine the transcriptional responses to (1) constant stress and (2) recovery, of wildtype *Halobacterium* to hydrogen peroxide and paraquat (See Figure 4-1). Mid-log phase cultures diluted to OD_{600} 0.10 were grown for 12-14 hours at 42°C with shaking at 220rpm in 125mL baffled flasks. At OD_{600} 0.40, experimental cultures were treated with either 25mM H_2O_2 or 4mM paraquat and incubated at 42°C with shaking up to 240 min. For constant stress time points, samples were taken at 0, 30, 60, 120, and 240 minutes, cells were pelleted (6000 x g, 3 minutes), and flash frozen in dry ice/ethanol bath after decanting the supernatant. For recovery time points, after a 2- hour treatment with either 25 mM H_2O_2 or 4mM paraquat, cells were centrifuged, washed with GN101/50μg/mL uracil and resuspended in GN101/50μg/mL uracil. Cultures were then incubated at 42°C with shaking and time points were taken at

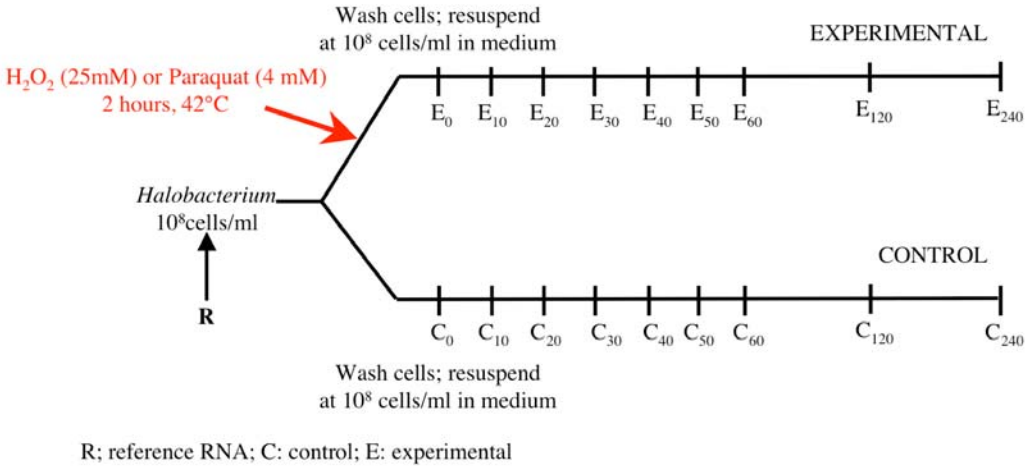


Figure 4-1. Experimental design for the microarray analysis of H_2O_2 and paraquat response in *Halobacterium.* Cells were treated with either 25mM H_2O_2 or 4mM paraquat for 2 hours and then washed to remove H_2O_2 and paraquat. Timepoints were taken every 10 minutes for the first hour and then at 2 hours and 4 hours as cultures recovered. Both control and treated samples were treated in the same manner except that H_2O_2 and paraquat were not added to the control cells.

0, 10, 20, 30, 40, 50, 60, 120, and 240 minutes and processed as above. RNA extractions were performed using the Stratagene Absolute RNA kit (La Jolla, CA) as described previously [148].

Microarray Analyses

Microarrays were designed at the Institute for Systems Biology Microarray Facility in collaboration with Nitin Baliga. The arrays contain four spots for each of the 2400 nonredundant genes in *Halobacterium*. RNA was labeled with Alexa594 and Alexa660 dyes, hybridized to the array, and washed with successive rinses of SSC [148]. A dye-flip was done to account for any bias in dye incorporation. Raw data was processed and converted into log10 ratios with lambda values determined by the maximum likelihood method [149]. Data analysis was performed using the Gaggle program and its coupled programs including Cytoscape and data matrix viewer [150, 151].

Survival Analyses of Mutant Strains

Cultures were grown up to an OD_{600} 0.4 in GN101/50mg/L uracil in baffled flasks shaking at 42°C in a Gyromax 737 shaker (Amerex Instruments, LaFayette, CA). Drug concentrations of 25 mM H_2O_2 (30% stock) or 4mM paraquat (1M stock) were added directly to the cultures in triplicate and incubation was continued for an additional two hours. After incubation, cells were pelleted at 8000 x g for 3 minutes, washed with GN101/50mg/L uracil, and resuspended in 1mL of GN101/50mg/L uracil. Dilutions were made in BSS and cells were plated on GN101/50mg/L uracil plates. Plates were incubated at 42°C for 5-7 days and colonies counted. Survival was calculated as Ni/No

where Ni is the number of viable cells after treatment and No is the number of viable cells without treatment. Three independent measurements were made for each mutant.

4.3 Results

We determined the oxidative stress response of *Halobacterium* to H_2O_2 and paraquat by whole genome transcriptional analysis and analyzed the survival of in-frame gene deletion mutants suggested to participate in ROS scavenging.

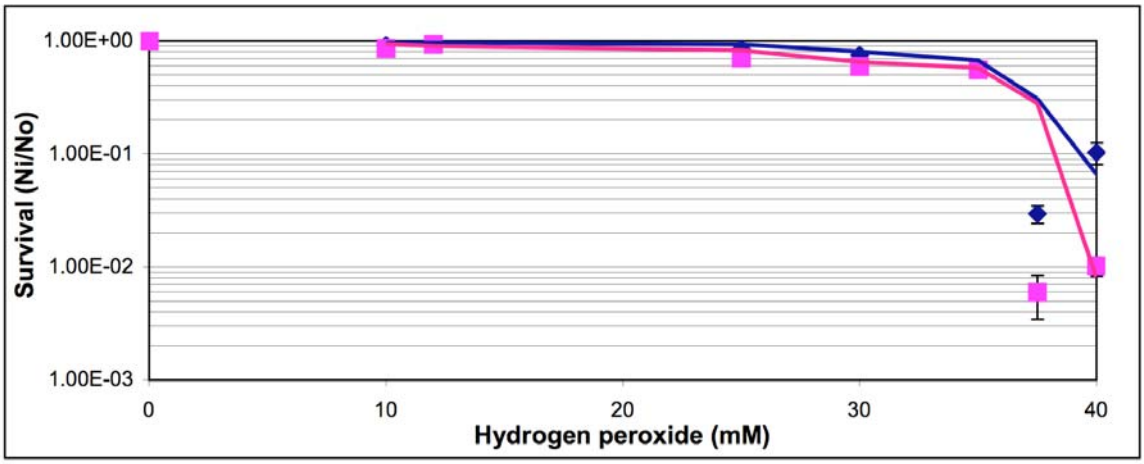
Survival of *Halobacterium* to H_2O_2 and paraquat

Survival curves were completed to determine survival of *Halobacterium* to various concentrations of H_2O_2 and paraquat (See Figure 4-2). Cells were grown to an OD_{600} 0.4 and treated with either H_2O_2 or paraquat for two hours. Survival of treated cells was calculated by counting colonies on plates post incubation at 42°C. *Halobacterium* exhibited survival up to approximately 35mM H_2O_2 and then drops drastically. Survival to paraquat shows a more linear decrease in survival. *Halobacterium* exhibited no decrease into survival up to 20mM H_2O_2 and 0.5mM paraquat. Doses of 40mM H_2O_2 and 10mM paraquat showed less than 10% survival of *Halobacterium.* The 80% survival dose of H_2O_2 and paraquat for *Halobacterium* was 25mM H_2O_2 and 4mM paraquat. These doses of H_2O_2 and paraquat will be used to determine mutant survival.

mRNA level changes

(*This analysis was done jointly by the author and the Baliga lab at the Institute for Systems Biology. Sample treatment and RNA extraction was done by the author.*)

A.



B.

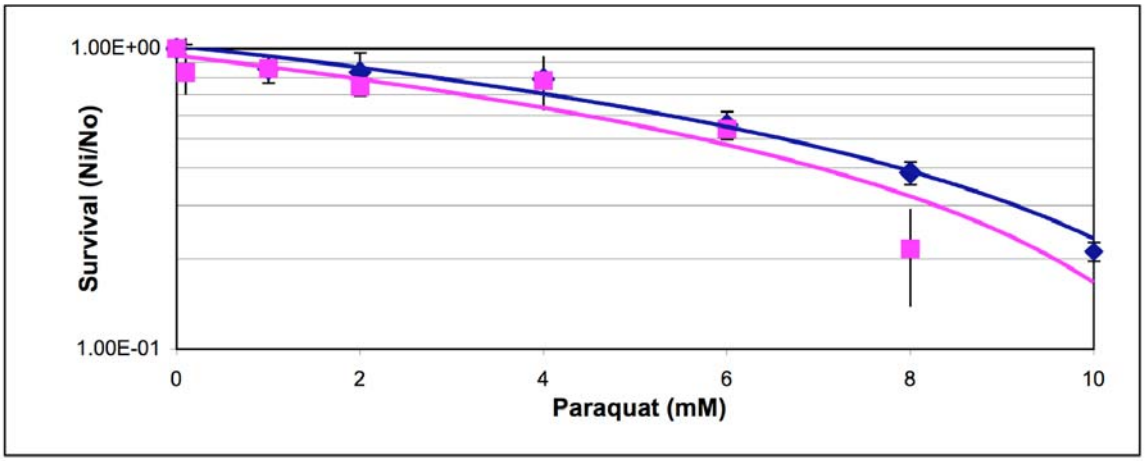


Figure 4-2. Survival of wildtype *Halobacterium* after exposure to increasing concentrations of (A) H_2O_2 and (B) paraquat. Blue diamonds represent survival after a 2 hour treatment with H_2O_2 or paraquat; pink squares represent survival after a 4 hour challenge with H_2O_2 or paraquat. Ni = number of viable cells in challenged samples; No = number of viable cells in control; Error bars represent standard error for three independent replicates. The line represents a best fit line.

The significance of mRNA level changes in the microarray data was estimated with a likelihood ratio test [149]. Comparison of RNA preparations from identically cultured cells, independently processed, yields λ values below 15 for at least 99% of the genes in the array. The results reported below are associated with a λ value above 15 and a confidence level of more than 99%.

Treatment with H_2O_2:

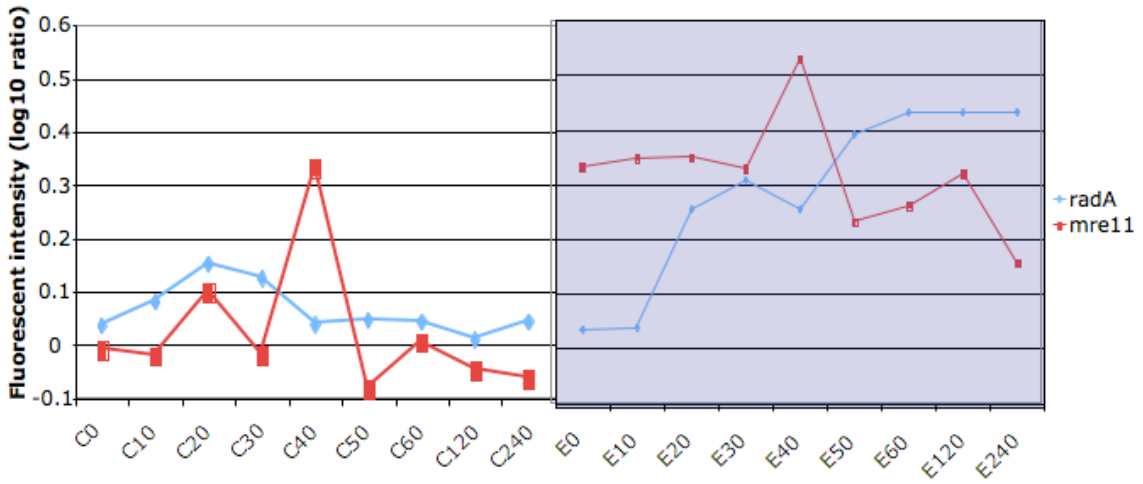
We undertook a microarray analysis of the global mRNA changes occurring during a timecourse of constant H_2O_2 stress for 30, 60, 120, and 240 minutes and during a timecourse of recovery after the removal of H_2O_2 every 10 minutes for the first hour and at 2 and 4 hours. Microarray analysis showed a downregulation of general cell metabolism pathways including the glycolysis and TCA cycles and ATP and nucleotide synthesis pathways for both constant stress and recovery after H_2O_2 treatment. There was also reduction in the mRNA transcript levels of genes involved in RNA polymerase biosynthesis and ribosome biosynthesis. There was an induction of pathways involved in the scavenging and repair of oxidative stress damage for both H_2O_2 constant stress and recovery after H_2O_2 treatment. These include the homologous recombination and BER pathways and ROS scavenging and detoxification systems. We will focus on the transcriptional responses of the DNA repair pathways and ROS scavenging and detoxification systems.

Three DNA repair pathways saw an induction in transcriptional response after H_2O_2 treatment: homologous recombination, BER, and NER. Genes involved in homologous

recombination include *radA,* responsible for strand invasion and exchange, and *mre11,* responsible for DNA double strand break recognition. Genes involved in BER include *mutT,* 8-oxo-dGTPase, *ogg,* uracil glycosylase, *alkA,* 3-methyladenine glycosylase, and *mutY,* an adenine glycosylase. The *rad3* and *rad25* genes are both helicases thought to play a role in NER along with the *uvrA/B/C* genes responsible for pyrimidine dimer excision. The mRNA transcript of *radA* was increased under conditions of H_2O_2 constant stress and recovery and the *mre11* transcript was increased only after recovery from H_2O_2 (See Figure 4-3A). Also increased after recovery from H_2O_2 treatment were mRNA levels of genes of the NER repair pathway, *rad3, rad25,* and *uvrA/B/C,* and BER pathway, *mutT, ogg, alkA,* and *mutY* (See Figure 4-3B).

We saw an induction of the systems responsible for scavenging free radicals during both constant stress with H_2O_2 and recovery after treatment. Carotenoid and bacteriorhodopsin pathways were upregulated after both constant stress with H_2O_2 and recovery after treatment (See Figure 4-4A). The mRNA transcript of the superoxide dismutase, *sod1,* was increased after constant stress with H_2O_2 and recovery after treatment while the other superoxide dismutase, *sod2,* did not have a change in mRNA transcript level (See Figure 4-5A). Several catalases and peroxidases had increased levels of mRNA transcripts including *perA,* catalase, *VNG0018,* putative catalase based on results from Rosetta protein matching programs, and *VNG0798,* predicted peroxidase based on a conserved domain homologous to the dyp-type peroxidase family (See Figure 4-5A). The *perA* mRNA transcript was increased for both constant stress with H_2O_2 and recovery

A.



B.

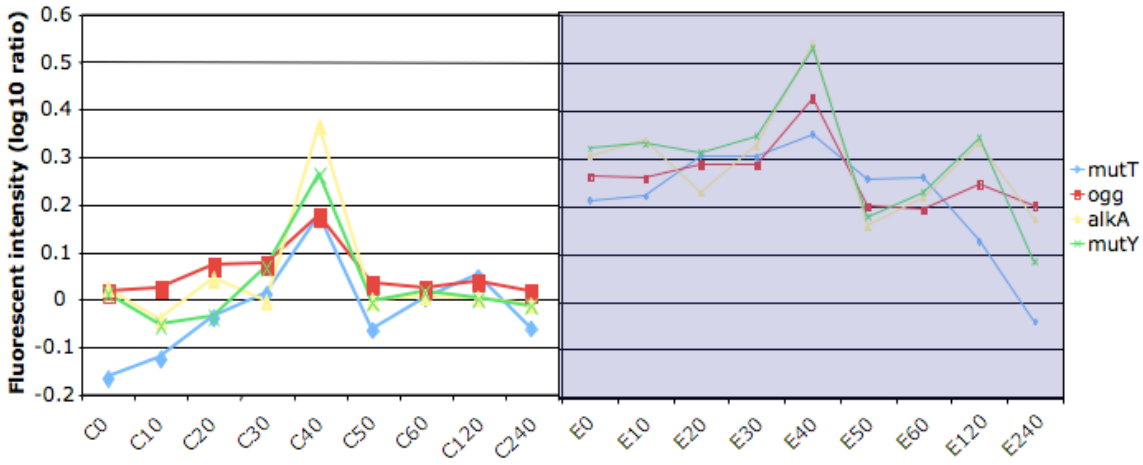
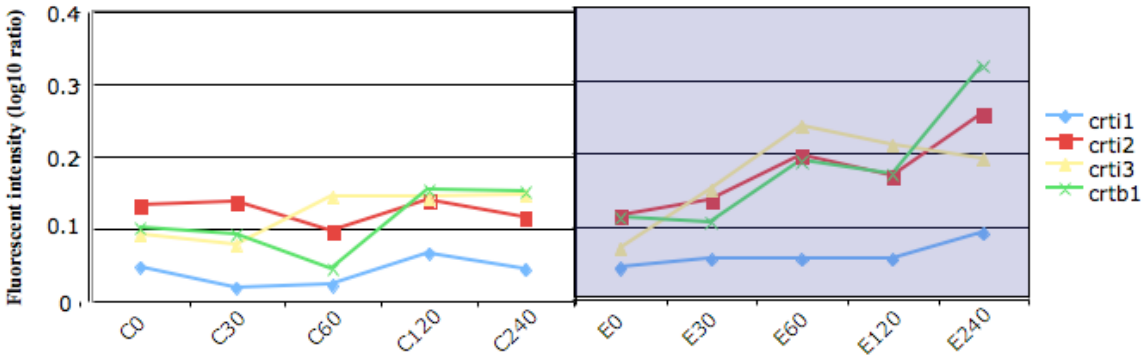


Figure 4-3. mRNA transcript levels of DNA repair genes involved in (A) homologous recombination and (B) BER after recovery from H_2O_2 treatment. The fluorescent intensity value of the mRNA transcriptional responses were expressed as log10 ratios after recovery from treatment at 10, 20, 30, 40, 50, 60, 120, and 240 minutes. C0-C240 represents control cells without treatment and E0-E240 represents treated cells. The dots are connected for visual purposes only.

A.



B.

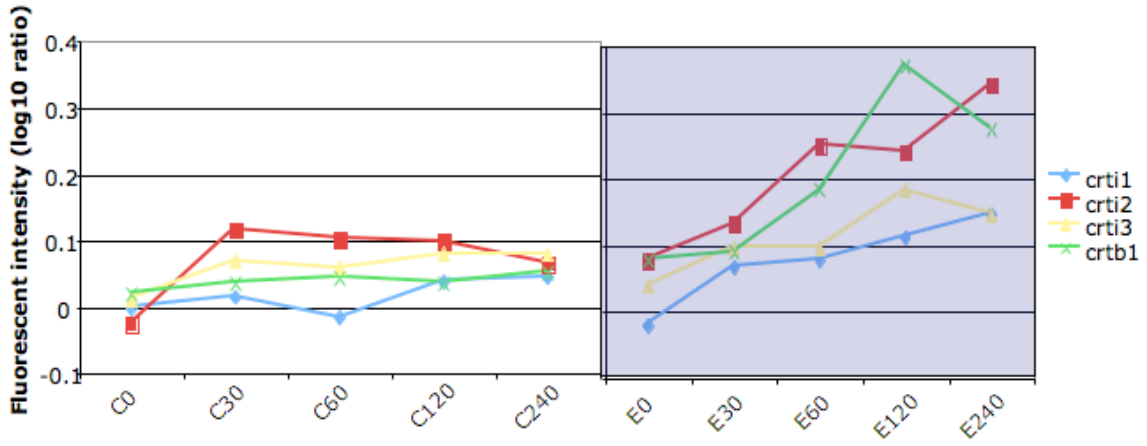
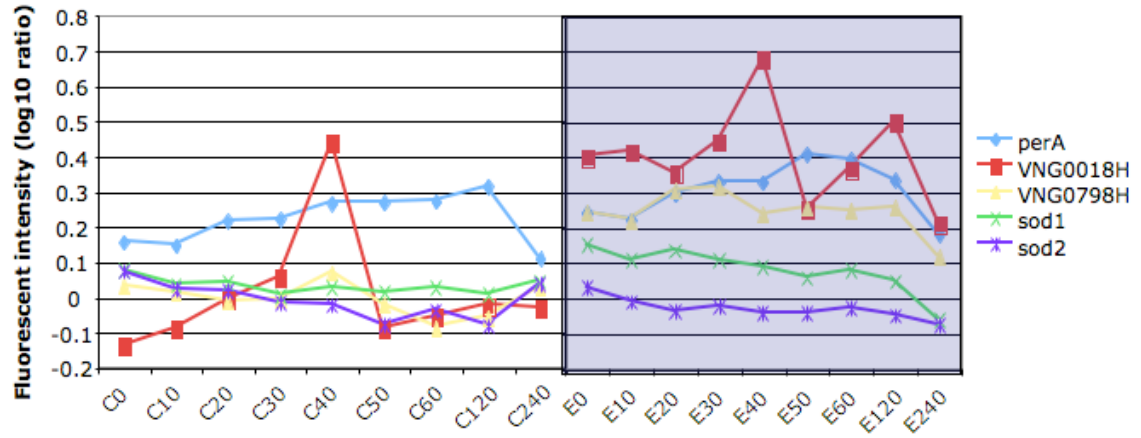


Figure 4-4. mRNA transcript levels of genes involved in carotenoid synthesis during constant stress to (A) 25mM H_2O_2 and (B) 4mM paraquat. The fluorescent intensity of the mRNA transcriptional responses were expressed as log10 ratios after constant stress at 30, 60, 120, and 240 minutes. C0-C240 represent control cells without treatment and E0-E240 represents treated cells. Dots are connected for visual purposes only.

A.



B.

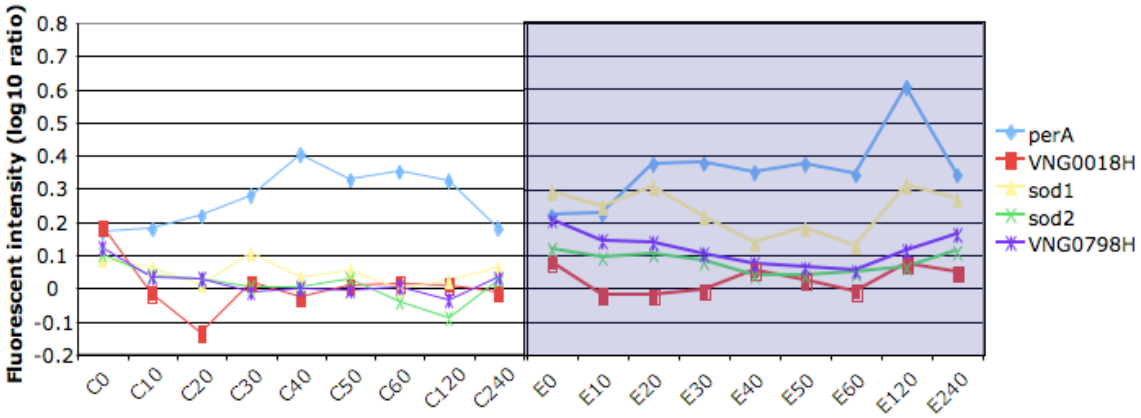


Figure 4-5. Log10 ratios for mRNA transcriptional responses of genes involved in ROS scavenging to (A) 25mM H_2O_2 and (B) 4mM paraquat. The fluorescent intensity of the mRNA transcriptional responses expressed at log10 ratios after recovery from treatment at 10, 20, 30, 40, 50, 60, 120, and 240 minutes. C0-C240 represents control cells without treatment and E0-E240 represents treated cells. Dots are connected for visual purposes only.

after treatment while the *VNG0018* and *VNG0798* mRNA transcripts were increased only during recovery after treatment with H_2O_2. Several other catalases and peroxidases had increased levels of mRNA transcripts along with a ferredoxin, *fer1*, a glutaredoxin system, and several thioredoxin systems.

Treatment with paraquat:

In addition to the transcriptional responses to H_2O_2, we undertook a microarray analysis of the global mRNA changes occurring during a timecourse of constant paraquat stress for 30, 60, 120, and 240 minutes and during a timecourse of recovery after the removal of paraquat every 10 minutes for the first hour and at 2 and 4 hours. Paraquat is a viologen that generates superoxide radicals. We saw a downregulation of general cell metabolism pathways including the glycolysis and TCA cycles and ATP and nucleotide synthesis pathways for both constant stress and recovery after paraquat treatment, which corresponded to what we determined with the H_2O_2 treatment. Again, similarly to the H_2O_2 data, we saw a downregulation of mRNA transcript levels of genes involved in RNA polymerase biosynthesis and ribosome biosynthesis. There was an induction of pathways involved in the scavenging and repair of oxidative stress damage for both paraquat constant stress and recovery after paraquat treatment. These include one DNA repair gene, *radA*, and ROS scavenging and detoxification systems. We will focus on the differences and similarities between the transcriptional responses of the ROS scavenging and detoxification systems between the H_2O_2 and paraquat treatments.

Similar to the transcriptional responses seen in *Halobacterium* for H_2O_2, there was an induction of the systems responsible for scavenging free radicals during both constant stress with paraquat and recovery after treatment. Carotenoid and bacteriorhodopsin pathways were upregulated after both constant stress and recovery after treatment (See Figure 4-4B). The mRNA transcripts of the superoxide dismutase mutant, *sod1,* were increased after constant stress with paraquat and recovery after treatment while the other superoxide dismutase mutant, *sod2,* had an increase in mRNA transcript levels for constant stress only (See Figure 4-5B). Several catalases and peroxidases had increased levels of mRNA transcripts during constant stress including *perA,* and *VNG0018,* and putative peroxidase *VNG0798* under both constant stress and recovery (See Figure 4-5B). Several other catalases and peroxidases had increased levels of mRNA transcripts and induction of glutaredoxin system and several thioredoxin systems was seen.

Survival of mutant strains of *Halobacterium* to H_2O_2 and paraquat

(*Mutant construction was completed at the Baliga lab at the Institute of Systems Biology. Survival analyses to H_2O_2 and paraquat were completed by the author)*

Survival to H_2O_2:

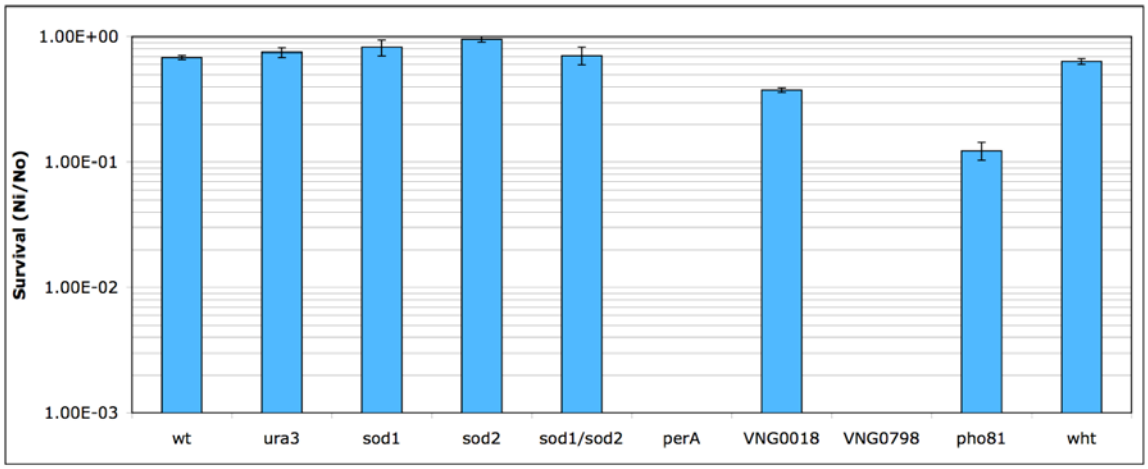
Microarray analysis of transcriptional responses showed an upregulation of mRNA transcripts of genes encoding catalases, peroxidases, and other ROS scavengers to 25mM H_2O_2. In frame gene knockout deletions using the method of Peck *et al* were made of several genes thought to play a role in the detoxification of ROS [121]. Mutant strains were grown to an OD_{600} 0.40 and treated with 25mM H_2O_2 for two hours. After

treatment, cells were washed and survival calculated by counting colony growth post incubation at 42°C.

Exposure to 25mM H_2O_2 led to a decrease in survival for the *ΔperA, ΔVNG0018, ΔVNG0798,* and *pho81* mutants compared to the *Δura3* background strain (See Figure 4-6A). Extremely low survival was shown for the catalase mutant, *ΔperA*, and a predicted peroxidase mutant *ΔVNG0798*. The predicted catalase mutant, *ΔVNG0018,* and *pho81,* a mutant missing the four sensory rhodopsins, showed a 30% and 10% survival after treatment with 25mM H_2O_2 respectively. Mutants missing genes for the superoxide dismutases, *Δsod1, Δsod2,* and *Δsod1Δsod2*, and *wht* mutant, pigment deficient mutant isolated during an EMS screen, did not show a decrease in survival to 25mM H_2O_2 relative to the *Δura3* background strain (Figure 4-6A).

In order to further characterize the *ΔperA, ΔVNG0018, ΔVNG0798,* and *pho81* mutants, we ran a dose response curve with 5, 15, and 25mM H_2O_2 (See Figure 4-6B). This should demonstrate at what concentration of H_2O_2 the cells can survive before becoming overwhelmed by the oxidative damages caused by H_2O_2. The *Δura3* background strain showed greater than 80% survival at 5, 15, and 25mM H_2O_2, while the *pho81* mutant showed 80% survival at 5 and 15mM H_2O_2 and then 1% survival at 25mM H_2O_2. Mutants missing the *perA* gene were very sick, even without H_2O_2 treatment, and showed no survival even at the lower doses of H_2O_2, 5 and 15mM. Putative catalase mutant, *ΔVNG0018,* did not show a decreased survival to 5mM H_2O_2 but showed a 70% decrease

A.



B.

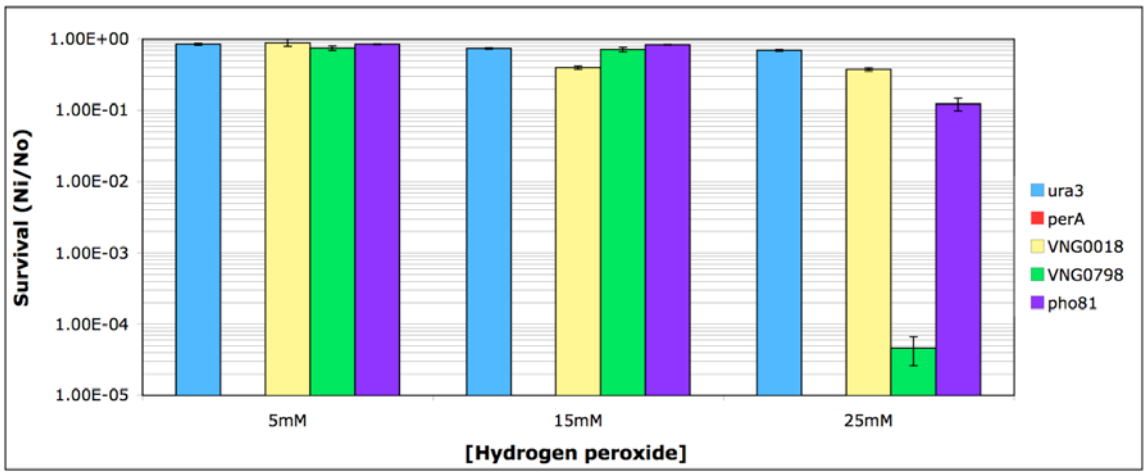


Figure 4-6. *Halobacterium* mutant strain survival after exposure to (A) 25mM H_2O_2 and (B) 5, 15, and 25mM H_2O_2. Ni = number of viable cells in challenged samples; No = number of viable cells in control; error bars represent standard error for three replicates. All mutants were constructed in a Δ*ura3* background strains except for *pho81* and *wht* mutants, which were isolated as previously described [147].

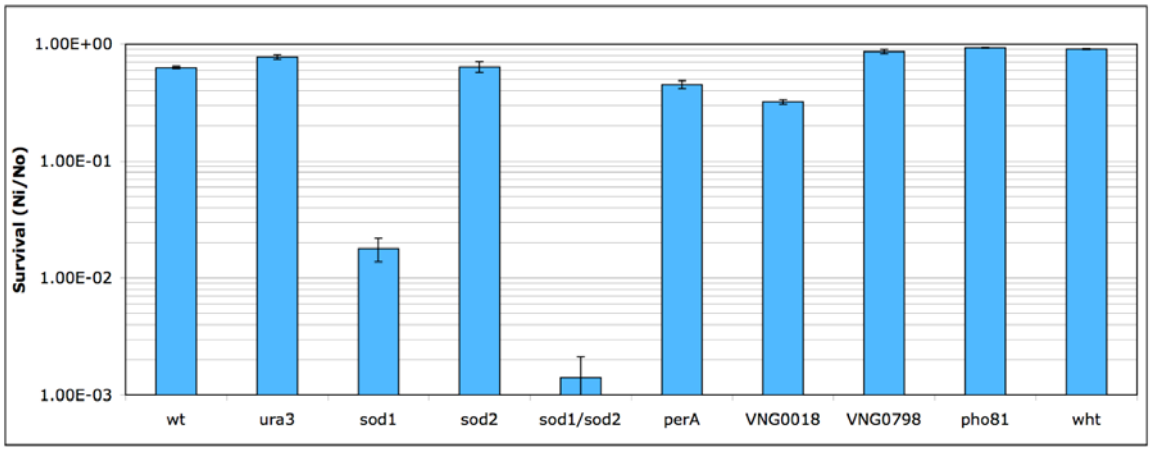
in survival at 15 and 25mM H_2O_2. The putative peroxidase mutant, *ΔVNG0798*, showed similar survival for 5 and 15mM H_2O_2, approximately 80%, but only showed 0.003% survival at 25mM H_2O_2.

Survival to paraquat:

In addition to H_2O_2, microarray analysis of transcriptional responses also showed an upregulation of mRNA transcripts of genes encoding catalases, peroxidases, and other ROS scavengers to 4mM paraquat. Similarly to the survival study with H_2O_2, Mutant strains were grown to an OD_{600} 0.40 and treated with 4mM paraquat for two hours. After treatment, cells were washed and survival calculated by counting colony growth post incubation at 42°C.

At 4mM paraquat, the *Δura3* background strain of *Halobacterium* shows 80% survival (See Figure 4-7A). The superoxide dismutase mutants, *Δsod1* and *Δsod2*, showed 0.1% and 70% survival respectively. The *Δsod2* mutant did not show a large decrease in survival by itself but showed an additive effect when combined with the *Δsod1* mutant as evidenced by the decreased survival (0.01%) in the *Δsod1Δsod2* mutant. The two catalase mutants, *ΔperA* and *ΔVNG0018*, also showed decreases in survival, 40% and 30% respectively. The peroxidase mutant, *ΔVNG0018*, and the pigment deficient mutants, *pho81* and *wht*, showed the same survival to 4mM paraquat as the *Δura3* background strain.

A.



B.

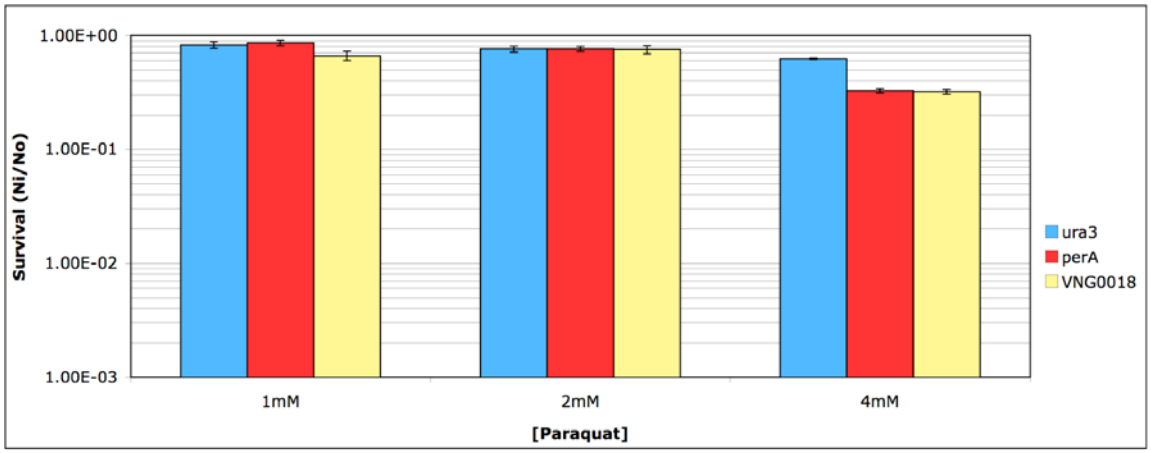


Figure 4-7. *Halobacterium* mutant strain survival after exposure to (A) 4mM paraquat and (B) 1, 2, and 4mM paraquat. Ni = number of viable cells in challenged samples; No = number of viable cells in control; error bars represent standard error for three replicates. All mutants were constructed in a Δ*ura3* background strains except for the *pho81* and *wht* mutants, which were isolated as previously described [147].

We ran a dose response curve to characterize survival of the catalase mutants at two lower concentrations of paraquat, 1 and 2mM. The *ΔperA* and *ΔVNG0018* mutants showed approximately 80% survival at both 1 and 2mM paraquat. At 4mM paraquat approximately 20% survival is shown (See Figure 4-7B). This survival is slightly lower than what was seen in the mutant survival study just using 4mM paraquat; however the survival of the *Δura3* background strain was also slightly lower (70%) in the dose response study, likely due to perturbations in the assay.

4.4 Discussion

From this study, we have identified the stress response of *Halobacterium* to H_2O_2 and paraquat. This is the first type of study using whole genome transcriptional analysis to elucidate the stress response of *Halobacterium,* an aerobic archaeon, to H_2O_2 and paraquat. We also determined the survival of mutants for selected in-frame gene knockouts to further characterize the oxidative stress response. Below is a discussion of genes involved in ROS scavenging and DNA repair and their response to oxidative stress by H_2O_2 and paraquat.

The production of ROS lead to damages in DNA, proteins, membrane lipids, and carbohydrates in aerobic organisms [135]. Aerobic organisms are able to deal with ROS through the presence of scavenger enzymes such as catalases and superoxide dismutases and the presence of repair pathways, such as BER and homologous recombination, to mediate repair of DNA nucleotides after oxidation [137-139]. In *Halobacterium* there are several mechanisms in place for protection against oxidative damage including phototaxis

away from high energy wavelengths, scavenging of free radicals by carotenoids, and the presence of catalases, peroxidases, and superoxide dismutases in the cells [86, 143, 145].

The encoding genes in the catalases in *Halobacterium* were upregulated under both H_2O_2 and paraquat conditions. In *E. coli*, upregulation of two transcriptional factors, SoxR and OxyR, which include superoxide dismutases and catalases, was seen in response to elevated superoxide and H_2O_2 [152]. In *Methanosarcina barkeri*, an anaerobic archaeon that is tolerant to oxygen exposure, the genes encoding catalases and superoxide dismutases were upregulated after exposure to 30 hours of oxygen [153]. In *Halobacterium*, the upregulation of the gene encoding *perA* during both contant stress and recovery after treatment to H_2O_2 and paraquat along with the high sensitivity of the gene deletion mutant to H_2O_2 suggest that PerA is the major catalase in the cells. We also saw an upregulation of the gene encoding a putative catalase, *VNG0018H*, during recovery from H_2O_2 and decreased survival of the gene deletion mutant to H_2O_2, however *VNG0018H* was not able to rescue the *ΔperA* mutant indicating that it may be an accessory catalase. Interestingly, the catalase mutants also demonstrated a decrease in survival to paraquat, which may be due to the production of H_2O_2 by the superoxide dismutases during the conversion of superoxide to H_2O_2 and oxygen. This has been seen in *E. coli* where high concentrations of paraquat induced the OxyR transcription factor normally only induced by H_2O_2 [154]. The mRNA transcript level of a putative iron-dependent peroxidase, *VNG0798H* was increased under both recovery from H_2O_2 and constant stress to and recovery from paraquat and the deletion mutant showed limited survival to H_2O_2 but only at the higher concentrations. This mutation was not

compensated by PerA suggesting that they play different roles in the conversion of H_2O_2 to less reactive products.

The level of *sod1* gene mRNA was increased with both H_2O_2 and paraquat and exposure to paraquat led to a decreased survival of the deletion mutant. The mRNA level of another superoxide dismutase gene, *sod2,* was increased under constant stress to paraquat only and while the *Δsod2* mutant by itself did not show a decrease in survival, the double mutant, *Δsod1Δsod2,* had a larger decrease in survival than the *Δsod1* mutant alone. This suggests that Sod1 and Sod2 work together to remove superoxides from *Halobacterium,* with Sod1 being the major player. Exposure to H_2O_2 did not result in a decreased survival for the superoxide mutants even though the mRNA level of *sod1* was increased. The increase in mRNA level could be result of superoxide dismutase activity needed to remove small levels of superoxide that result from H_2O_2 accumulation and its subsequent transformation into superoxide by the Fenton/Harber-Weiss reaction [97, 135].

The carotenoid and bacteriorhodopsin biosynthetic pathways are highly upregulated for for H_2O_2 and paraquat constant stress and recovery. The scavenging of ROS by pigments play a major role in the oxidative stress response of cells seen in both *Halobacterium* and *D. radiodurans* and an upregulation of their encoding genes mRNA is seen in transcriptional analysis to gamma irradiation, a producer of free hydroxyl radicals, indicating that this is a global response to different types of oxidative stress [85, 86, 143, 144, 155]. The two membrane deficient mutants tested in this study, *pho81,* missing the rhodopsins, and *wht,* missing the rhodopsins and carotenoids, had conflicting results. The

pho81 mutant was sensitive to H_2O_2 suggesting that the rhodopsins may play a role in scavenging some of the ROS, providing some protection to the cells. However, the *wht* mutant did not show a sensitive phenotype to H_2O_2. This suggests that if any scavenging of ROS is occurring, as shown with gamma and *in vitro* studies, it is not detectable at this concentration of oxidant [85]. This also suggests that the decrease in survival of the *pho81* mutant is not related to the absence of rhodospin in the membrane. This discrepancy might be the result of an insertion sequence element in the transducer protein for the sensory rhodopsin gene which results in an inhibition of phototaxis and a change in photochemistry properties [156, 157].

We saw an upregulation of the mRNA levels of the genes encoding the thioredoxin and glutaredoxin systems, involved in the removal of peroxides, electron transfer, and the control of redox reactions, with both H_2O_2 and paraquat [158]. Also upregulated was the mRNA levels of *fer,* a ferredoxin, under H_2O_2 stress. This is similar to responses after gamma irradiation where the glutaredoxin and ferredoxin systems are upregulated [85]. This suggests that these systems may be playing a role in scavenging free radicals in *Halobacterium*. The thioredoxin and glutaredoxin mutants did not show a decrease in survival to either H_2O_2 or paraquat likely because of redundancy in these systems. This is different than what is seen in *E. coli,* where deletion of the thioredoxin genes leads to a greater decrease in survival to H_2O_2 then the wildtype strain [159-161].

DNA repair to correct oxidative damages caused by H_2O_2 and paraquat seems to be mediated by the homologous recombination and BER pathways. After treatment to both

H_2O_2 and paraquat, there was an increase in mRNA transcript level for *radA*, responsible for catalyzing strand exchange in homologous recombination. This upregulation of mRNA transcript levels is also seen after gamma irradiation in *Halobacterium* and *D. radiodurans* [85, 162]. Similarly to the gamma study, there were increased mRNA transcript levels of the BER enzymes, in particular ones used in the removal of oxidized bases such as *mutT,* 8-oxo-GTPase, and also of *mre11,* involved in homologous recombination after recovery from H_2O_2 [85].

During constant stress by H_2O_2 and paraquat, we saw increased levels of mRNA transcripts for the major catalase, *perA*, and superoxide dismutase, *sod1,* in the cells along with genes involved in the thioredoxin and glutaredoxin systems indicating their major roles in the scavenging of oxidative stress agents. During conditions of recovery from H_2O_2 and paraquat, we noted additional upregulation in the mRNA transcripts of alternate catalases and peroxidases, and DNA repair enzymes. This suggests that during recovery, *Halobacterium* induces DNA repair machinery to fix the damages caused by oxidative damage whereas in constant stress, the upregulation of scavenging enzymes is mainly used to remove the oxidative stress from cells.

In general, the microarray analysis showed a downregulation of general cell metabolism pathways such as glycolysis and the TCA cycle, ATP and nucleotide synthesis, and RNA polymerase and ribosome biosynthesis during both constant stress by H_2O_2 and paraquat and recovery after stress. This seems to be a general response of *Halobacterium* to slow down the cells and prepare for DNA repair after damage by UV and gamma and other

oxidative stresses [84, 85]. We saw an upregulation of the ROS systems and through in-frame gene deletions, we were able to determine which catalases and superoxide dismutases were essential in the oxidative stress response in *Halobacterium.* Future work would be construction in-frame gene deletion mutants of the glycosylases thought to play a role in the repair of oxidized bases and characterization of their survival to H_2O_2 and paraquat. During DNA repair, the oxidized base, 8-oxoguanine, is removed by a glycosylase specific for this type of damage [163]. In the Archaea and Eukarya an 8-oxoG glycosylase, OGG, is responsible for removing the oxidized base [163-165]. Bacteria also use a 8-oxoG glycosylase to remove the oxidized base called FPG [166]. These glycosylases are functionally similar but substantially different in sequence. Determining the role these enzymes play in the repair against oxidative damage in *Halobacterium* will expand our understanding of DNA repair processes in all the domains of life.

Chapter 5: Conclusions

DNA mismatch repair plays a major role in correcting errors made after DNA replication. This pathway is highly conserved and the key proteins, MutS and MutL, are found in both Bacteria and Eukarya [3, 4]. This pathway is critical for maintaining genome integrity and defects in this pathway can lead to a 50-1000 fold increase in spontaneous mutability, meiotic defects, and tolerance to several DNA damaging agents [3-5]. Only four halophiles and seven methanogens have homologs of the MutS and MutL proteins but mutation rate studies in other archaea have shown a low genomic mutation rate indicating that either the Archaea have effective repair systems or they have a low incidence of mutation likely due to a high fidelity polymerase [67, 68, 74]. Homologs of the bacterial MutS and MutL proteins have been found in the genome of *Halobacterium.* *Halobacterium* has 3 bacterial-like *mutS* genes, of which *mutS1* and *mutS2* are homologous the MMR *mutS* gene in *E. coli,* 1 bacterial-like *mutL* gene, 5 exonucleases, 4 bacterial-like *recJ* genes, and 1 eukaryotic-like *exoI* gene, and 1 bacterial-like *uvrD* gene. We hypothesized that the bacterial-like MutS and MutL homologs in *Halobacterium* are involved in a MMR pathway similar to that seen in the Bacteria and Eukarya. If not, it is likely that *Halobacterium* also has the same type of mismatch avoidance and repair systems found in other archaea.

We calculated the spontaneous genomic mutation rate in *Halobacterium* to determine if there was a low incidence of mutation, which would suggest that either postreplicative repair of DNA is taking place or there is very little mutation present. We calculated a low genomic mutation rate in *Halobacterium,* which is within the range of genomic mutation

rates calculated for organisms in the other domains of life [66]. We also characterized the spectrum of mutation to allow comparison between other organisms. Unlike what is seen in Bacteria and Eukarya, the Archaea have a high occurrence of insertions and deletions. This may be a result of the structural properties of polymerases in the Archaea compared to those in the Bacteria and Eukarya. There are several hypotheses for the low incidence of mutation (1) the MutS and MutL homologs in *Halobacterium* are functioning in a MMR pathway similar to what is seen in the other domains of life; (2) the Archaea have a high fidelity polymerase resulting in a decreased occurrence of base pair mismatches; and (3) there is an error correction system specific to the Archaea that does not utilize the MutS and MutL homologs.

To test the hypothesis that the MutS and MutL homologs in *Halobacterium* are essential for the low incidence of mutation and thus possibly acting in a bacterial-like MMR pathway, we constructed in-frame targeted gene deletion mutants of the *mutS1, mutS2, mutS1mutS2, mutL,* and *uvrD* genes and characterized the mutant phenotypes with regards to tolerance to DNA alkylating agents (MNNG), survival to UV-C and gamma-ray, and mutation frequency. If the bacterial-like MMR homologs are essential to maintain a low genomic mutation rate in *Halobacterium,* we should see an increase in tolerance to MNNG, a decrease in survival to UV-C, and an increase in mutation frequency, similar to the phenotypes seen in Bacteria and Eukarya. We did not see a phenotype in the mutant strains of *Halobacterium* that was similar to that seen in other organisms. This suggests that the MutS and MutL homologs in *Halobacterium* are not essential for maintaining the low genomic mutation rate. While this is surprising, since

studies looking at the bacterial-like UvrA/B/C homologs in *Halobacterium* showed they were essential in NER, it is possible that the MutS and MutL homologs in *Halobacterium* are acting to correct mismatches caused by DNA replication but that there are also other proteins that can fix this damage in their absence such as DNA glycosylases perhaps acting in conjunction with the homologous recombination pathway [88, 132].

These results leave us to differentiate between the other two hypotheses: (2) the Archaea have a high fidelity polymerase resulting in a decreased occurrence of base pair mismatches or (3) there is an error correction system specific to the Archaea. Benefits of having a high fidelity polymerase would be a decreased incidence of replication error resulting in a low requirement for MMR and studies looking at polymerase fidelity in the hyperthermophiles have shown that they have a 10-fold higher fidelity than that of bacteria [72-74]. Alternatively, if there is a requirement for MMR, recruitment of an archaeal-specific system is likely. Archaea have two glycosylases that can are specific for mismatched bases, a uracil-DNA glycosylase, UDG, and a thymine-DNA glycosylase, TDG [128, 167-169]. The spontaneous deamination of cytosine to uracil can result in a mispair with adenine. Archaeal DNA polymerases are unique in their ability to stall at a uracil residue to allow removal of the uracil by UDG and incorporation of the correct nucleotide by DNA polymerase [128, 130, 167]. Direct interaction between UDG and PCNA has been documented in the Archaea suggesting recruitment of the glycosylases by PCNA to the site of the mispair [129]. Homologs of TDG have also been found in the Archaea. This glycosylase is responsible for the removal of thymine from GT mismatches and can also remove thymine from CT and TT mismatches albeit at lower

efficiency [168, 169]. The possibility of other glycosylases that are specific for mismatched base pairs cannot be ruled out. In addition to the recruitment of glycosylases to fix mismatched bases, the homologous recombination pathway may also play a role. The homologous recombination pathway is responsible for the repair of double strand DNA breaks and the proteins have been shown to be involved in many other repair activities [132].

Future work would be to develop an *in vitro* assay, similar to the one based on the methods by Thomas *et al* [170], to determine the repair efficiency of mismatched heteroduplexes by *Halobacterium* and strains without the bacterial-like MMR proteins. Using this assay, we should be able to determine if there is MMR activity present in the cell extracts of wildtype *Halobacterium* and if this activity varies in *Halobacterium* MMR gene deletion mutants. This assay would have to be optimized for high salt condition since the proteins in *Halobacterium* are not active in low salt concentrations. Other work would be to look for mutator phenotypes in *Halobacterium* and elucidate the genes responsible for causing the mutator phenotype. Bacteria and eukaryotes deficient in MMR are characterized by a mutator phenotype and determining genes responsible for mutator phenotypes in *Halobacterium* could lead to an alternate pathway for MMR in the Archaea [3, 4]. Assays for mutator phenotypes have not been described in *Halobacterium* and development of an assay would be required to look for mutant phenotypes.

Deciphering the DNA repair pathways in *Halobacterium* allows a broader view of DNA repair mechanisms that have evolved in extreme environments. The possible recruitment

of repair enzymes to the site of DNA mismatches could potentially lead to the discovery of new repair pathways or the interactions between different repair pathways. This would further our understanding of pathway interactions, genomic maintenance, and mutation avoidance in the other domains of life.

The global stress response to hydrogen peroxide and paraquat was measured in *Halobacterium* using whole genome transcriptional arrays. Studies of transcriptional responses in *Halobacterium* to gamma irradiation have been done and results showed a downregulation of general cell metabolism and an upregulation in the mRNA transcripts of genes involved in homologous recombination and base excision repair [85]. More than 80% of the damage caused by gamma irradiation is the indirect result of the radiolysis of water into hydroxyl radicals [97]. Hydrogen peroxide and paraquat also produce ROS. Hydrogen peroxide, in the presence of Fe^{2+}, can be converted into a hydroyxl radical through Fenton chemistry [97, 139]. Paraquat is an intracellular generator of superoxide, which can react with iron sulfer clusters releasing free Fe^{2+}[136].

Transcriptional responses of *Halobacterium* to hydrogen peroxide and paraquat showed a global downregulation of cell metabolism, indicating the cells slowing down and preparing to repair damages caused by oxidative stress as well as an upregulation of mRNA transcripts of genes involved in reactive oxygen species scavenging, membrane pigments, catalases, peroxidases, superoxide dismuatases, and thioredoxin and glutaredoxin systems, and DNA repair including homologous recombination, base excision repair, and nucleotide excision repair. Based on the transcriptional responses, we

were able to identify key genes involved in the protection and repair of oxidative stress in *Halobacterium*. The upregulation of membrane pigments and carotenoids is unique and only seen in a few organisms and seems to provide protection against ROS. Deciphering the transcriptional responses of the catalases, peroxidases, and superoxide dismutases, allowed insight into what genes may be responsible for removing these damages in *Halobacterium.* Through in-frame gene deletions of these ROS scavengers and characterization of their survival to H_2O_2 and paraquat, we were able to determine which genes were essential for cell survival. Our study, along with results from the gamma irradiation and dessication studies, allowed a broader look at the global response of *Halobacterium,* and organisms living in environments characterized by oxidative stress, to different types of oxidative stress. Future work would include creating in-frame gene knockout constructs of the glycosylases thought to be involved in the removal of oxidized bases and characterize their mutant phenotypes to hydrogen peroxide and paraquat to address the role of repair after oxidative stress. This could result in a broader view of BER in the other domains of life and in other organisms living in extreme environments.

Appendix A: Overexpression of *Halobacterium* MutS1 protein in *E. coli* and *Halobacterium*

A.1 Introduction

The three dimensional structure of the MutS protein has been resolved in *Escherichia coli* and *Thermus aquaticus* [15, 16]. It is a 95kDa protein that functions as a dimer *in vivo* [5, 17]. MutS is an ATPase with Walker A/B sequence motifs and a highly conserved Phe-X-Glu motif responsible for binding DNA [18]. MutS forms a homodimer in bacteria when binding to the DNA but the asymmetry of the two subunits bound to the mismatched DNA is similar to that of the MutS heterodimers in the eukaryotes [4]. Crystal structures reveal the two subunits form a channel in MutS, one which contains the phenylalanine responsible for binding DNA with the other subunit contacting the DNA to form a clamp [3, 15, 16]. Su *et al* showed that *E. coli* MutS protein is capable of binding several DNA mismatches using purified MutS protein from *E. coli* in a mismatch correction assay [171, 172]. Biochemical characterization of the MutS homologs in *S. cerevisiae* demonstrated that MSH2/MSH3 and MSH2/MSH6 display different preferences for the binding of mismatches in DNA. Using mobility shift experiments, Habraken *et al* looked at the binding of MSH2/3 complex to DNA duplexes containing loops and found that MSH2/3 preferentially binds to loops formed after DNA replication but not DNA mismatches [37]. Experiments looking at the binding of overexpressed MSH2/6 complexes in yeast demonstrated that these complexes preferentially bind to DNA mismatches, especially G/T mismatches [40, 173].

Halobacterium has two homologs of the bacterial-like MutS protein involved in mismatch binding. The presence of two MutS proteins in *Halobacterium* suggests the formation of MutS heterodimers as in the eukaryotes. The protein sequences of MutS1 and MutS2 in *Halobacterium* are 43% identical to one another and share 39-44% and 21-22% similarity at the amino acid level with *E. coli* and *T. aquaticus* and *Saccharomyces cerevisiae* respectively. The domain organization of the MutS1 and MutS2 proteins in *Halobacterium* is similar to that of other MutS proteins (See Chapter 1, Figure 1-2). Conserved in the *Halobacterium* MutS proteins is the Phe-X-Glu motif responsible for the binding of MutS to mismatches [93].

In this study we intended to biochemically characterize the MutS1 protein in *Halobacterium.* We attempted the expression of *Halobacterium* MutS1 in *E. coli* and also the overexpression of the MutS1 protein in *Halobacterium* to test the recombinant protein for its binding ability to mismatched and perfectly matched DNA.

A.2 Material and Methods

E. coli expression

Using the Champion pET Directional TOPO Expression Kit (Invitrogen, Carlsbad, CA), a blunt-end PCR product of the *mutS1* gene from *Halobacterium* was TOPO cloned into a pET100/D/*lacZ* plasmid. This plasmid contains a *lacZ* gene that has been cloned in frame with the N-terminal peptide containing the 6xHis tag in addition to the *lacI* gene encoding the lac repressor (See Figure A-1). We produced PCR products of both the full

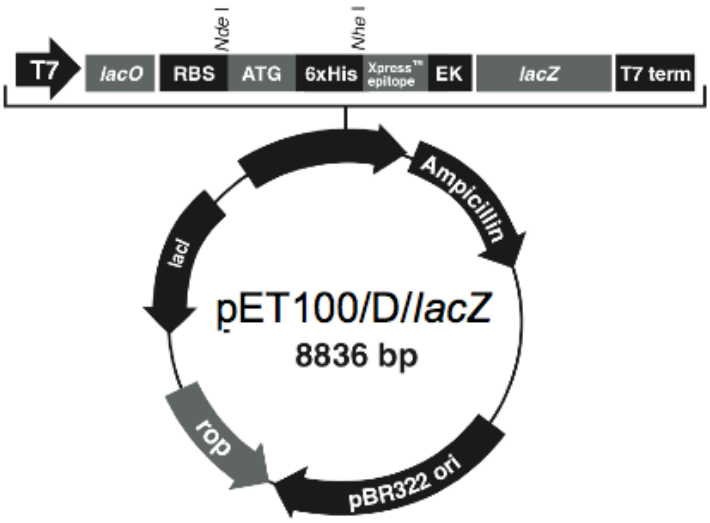


Figure A-1. Plasmid map of pET100/D/*lacZ* [174]. This plasmid contains an ampicillin resistance gene for selection of insert and a *lacI* gene that encodes the lac repressor. This lac repressor binds to the lacUV5 promoter in T7 expression systems and prevents transcription of the gene unless IPTG is present. When IPTG is present, IPTG will bind to the lac repressor and allow transcription of the gene.

length *mutS1* gene (Full Pro Ci: 5'TCAGTCCTCCAGTCGGTCCTGCCA3'; Full Pro Ni: 5'CACCATGGGGATCGTAGACGAGTTC3') and truncated version of the *mutS1* gene (Pro Ci Trunc 5'TCACCGCCGCTGTCCACGTCGAAGACG3'). The resultant MutS1 full length and truncated pET TOPO plasmid was transformed into Top10 competent *E. coli* cells. Clones were analyzed to confirm insert by restriction analysis and PCR. Plasmid DNA was isolated from Top10 *E. coli* cells and transformed into BL21 Star (DE3) cells (Invitrogen, Carlsbad, CA). The BL21 Star (DE3) cells are part of a T7 expression system. BL21 Star (DE3) cells possess a copy of the T7 polymerase under the control of a lacUV5 promoter. When cells are grown without IPTG, the lac repressor in the pET TOPO plasmid will bind to the lac operator in the lacUV5 promoter and prevent transcription. The addition of IPTG results in the binding of IPTG to the lac repressor, which turns on transcription from the lacUV5 promoter. After transformation, cells are incubated overnight at 37°C. The overnight culture was diluted back in LB/50mg/L ampicillin and grown until midlog phase. Expression of the MutS1 protein was induced with the addition of 1mM IPTG and incubated for an additional 2.5 hours. Overexpression of recombinant protein was checked on 8% SDS-PAGE at 10-20mA for 2 hours. The protein extract was incubated with the His·Bind resin (Probond Invitrogen, Carlsbad, CA) and recombinant MutS1 was eluted from the nickel resin with a high salt (4M NaCl) low pH imidazole containing buffer [175]. Purification of the protein is monitored using SDS-PAGE.

Halobacterium expression

In addition to the overexpression of the *Halobacterium* MutS1 protein in *E. coli,* we also overexpressed the MutS1 protein in *Halobacterium.* We PCR amplified a full length (Full Pro Ni NdeI 5'GCGCATATG*CATCATCATCATCATCAT*ATGGGGATCGTAGACGAGTTC3'; Full Pro Ci SpeI 5'GCGACTAGTTCAGTCCTCCAGTCGGTCCTGCCA3') and truncated (Pro Ci Trunc SpeI 5'GCGACTAGTTCACCGCCGCTGTCCACGTCGAAGACG3') *mutS1* gene from *Halobacterium* including a polyhistine tag (italicized) and restriction sites (underlined). The constructs were cloned into *Halobacterium* plasmd pNBPA downstream of the ferredoxin promoter. The pNBPA is a plasmid with a strong constitutive promoter, a copy of the mevinolin resistance gene to ensure maintanence of plasmid in cells, and a *Halobacterium* origin of replication. The resulting plasmid was transformed into *Halobacterium. Halobacterium* containing the plasmid construct was grown to exponential phase and overexpression checked using SDS-PAGE using the same conditions as above.

A.3 Results

The objective of this experiment was to overexpress and purify *Halobacterium* MutS1 protein from either *E. coli* or *Halobacterium* for use in a mismatch binding assay. Overexpression of the *Halobacterium* MutS1 protein was successful in *E. coli* for both the full length and truncated protein (See Figure A-2). However, further analysis showed that the MutS1 protein was in the insoluble fraction (See Figure A-3). Soluble and

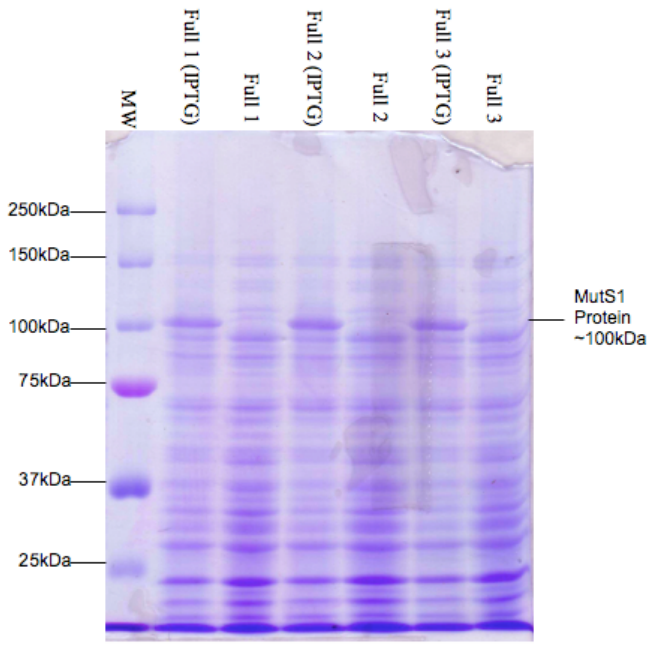


Figure A-2. 8% SDS-polyacrylamide gel showing overexpression of the *Halobacterium* MutS1 protein in *E. coli*. Full length MutS1 protein was expressed in *E. coli* using the T7 expression system. Induction with 1mM IPTG for 2 hours resulted in the overexpression of the 100kDa MutS1 protein. MW = BioRad Precision Plus Protein Standard. Full 1 IPTG, Full 2 IPTG, and Full 3 IPTG are three different *E. coli* extracts with *Halobacterium* MutS1 expressed showing overexpression after addition of IPTG. Full 1, Full 2, and Full 3 are the same as above without IPTG induction. *E. coli* cultures were centrifuged and pellets resuspended in SDS loading buffer. Extracts were boiled for 2 minutes and 10μL of each sample loaded onto gel.

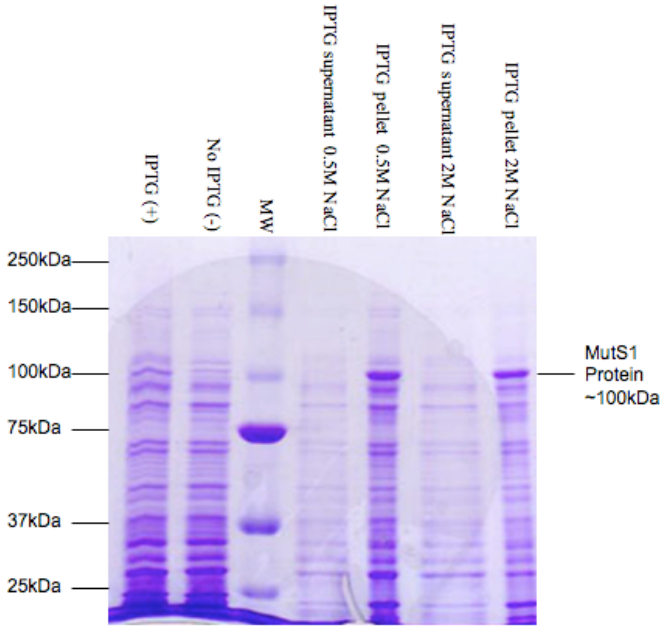


Figure A-3. 8% SDS-polyacrylamide gel showing solubility of *Halobacterium* MutS1 protein in *E. coli* in 0.5M and 2.0M NaCl buffer. Full length MutS1 protein was expressed in *E. coli* using the T7 expression system, centrifuged, and the cell pellet resuspended in 0.5M, 2M or 4M NaCl buffer. Cells were disrupted with sonication and centrifuged to separate the soluble and insoluble fractions. Soluble fraction is found in the supernatant and insoluble fraction is found in the cell pellet. Samples were boiled for two minutes and 10μL loaded on gel. MW = BioRad Precision Plus Protein Standard. IPTG (+) and No IPTG (-) are *E. coli* cell extracts of overexpressed *Halobacterium* MutS1 protein without sonication. IPTG soluble 0.5M and IPTG insoluble 0.5M are supernatant (soluble) and pellet (insoluble) resuspended in 0.5M NaCl buffer while IPTG soluble 2M and IPTG insoluble 2M were resuspended in 2M NaCl. Solubility was also checked in 4M NaCl buffer and results were the same (data not shown).

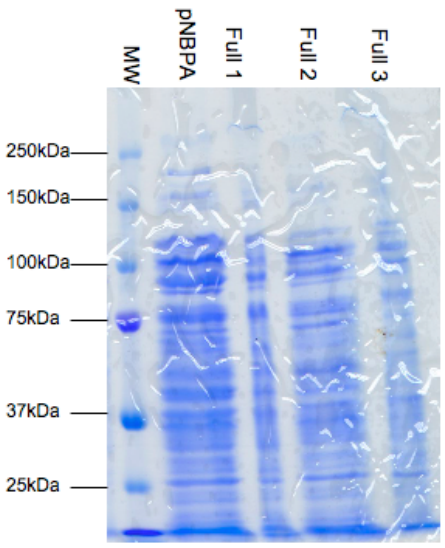


Figure A-4. 8% SDS polyacrylamide gel showing lack of overexpression of full length MutS1 protein in *Halobacterium.* The full length *mutS1* gene was cloned into the *Halobacterium* overexpression plasmid pNBPA behind a strong constitutive promoter. Resultant plasmid was transformed into *Halobacterium* and cells grown to three different ODs. Cells were centrifuged, samples boiled for two minutes, and 10μL loaded onto gel. MW = BioRad Precision Plus Protein Standards. Full 1, Full 2, and Full 3 indicate *Halobacterium* cell extracts at different ODs: 1 = 0.6, 2 = 0.8, and 3 = 1.0.

insoluble fractions were obtained by centrifugation of the overexpressed culture, resuspension of the pellet in buffer containing increasing concentrations of NaCl, sonication, and centrifugation to separate the soluble (supernatant) and insoluble (pellet) fractions. Overexpression of the full length and truncated MutS1 protein in *Halobacterium* was not seen (See Figure A-4). We were unable to purify the MutS1 protein from *E. coli* or *Halobacterium.*

A.4 Discussion

We attempted to overexpress and purify the *Halobacterium* MutS1 protein both in *E. coli* and in *Halobacterium.* Attempts were ultimately unsuccessful. Polyhistidine tagging and nickel affinity chromatography was successfully used with functional sensory rhodopsins from *Halobacterium salinarium* and *Natronobacterium pharaonis* overexpressed in *E. coli* [176, 177]. The sensory rhodopsins are phototaxis receptors found in the cell membrane of *Halobacterium*. The MutS1 protein is found in the cytosol and this may have hindered our ability to solubilize the protein for future testing in a mismatch binding assay. *Halobacterium* maintains osmotic balance in a high salt environment by accumulating a high intracellular concentration of KCl. The proteome of *Halobacterium* is highly acidic and most proteins denature when suspended in low salt environment [178]. It is likely that the low salt environment in *E. coli* led to improper folding of the *Halobacterium* MutS1 protein preventing attempts at solubilization. More puzzling is the lack of overexpression of MutS1 in *Halobacterium.* It is possible that this protein is expressed at too low of levels to elucidate on a SDS-PAGE gel. Western blots using an antibody against the histidine tag were attempted but no detection of the histidine tagged

MutS1 full length or truncated was seen. Previous microarray analyses showed that MutS1 is transcribed but it does not appear that the MutS1 tagged protein is being expressed in *Halobacterium* [84, 85]. This could be because a high expression of MutS1 protein in cells could cause binding to not only mismatched DNA but also perfectly matched DNA.

Appendix B: *in vitro* assay to test the capability of *Halobacterium* wildtype and MMR gene deletion strains to repair mismatches

B.1 Introduction

The MMR proteins, MutS and MutL, play a key role in repairing errors made during replication in Bacteria and Eukarya [3, 4]. This repair is essential for maintaining genomic stability and defects in this pathway can result in a 50-1000 fold increase in spontaneous mutability [5]. MutS homologs initiate MMR by recognizing the mismatched base and recruiting MutL homologs. This MutS/MutL complex activates downstream processes, namely excision of the mismatch and repair of the DNA [3-5].

Bacteriophage, more commonly referred to as phage, are viruses that infect bacteria [1]. Direct evidence of mismatch correction has been seen in *E. coli* by transfection with phage containing mismatches [13, 179]. After transfection with *E. coli*, the phage plaques can be isolated and tested for mismatch correction. Lu *et al* [13] developed an *in vitro* assay to analyze MMR in crude extracts of *E. coli.* A heteroduplex molecule, made from f1 R229 phage DNA, containing a mismatched base located within a single restriction site was incubated with cell extracts of *E. coli* wildtype and MMR deficient strains. Cell extracts from wildtype strains were able to correct the mismatch, thus restoring the restriction site, whereas extracts of the MutS and MutL deficient homologs were not [13].

Similarly, using human cell extracts and heteroduplex substrates, Thomas *et al* [179] were able to elucidate factors needed for MMR activity. M13mp2 DNA substrates containing base mismatches and insertions within the *lacZ* gene were incubated with

human cell extracts. After incubation, the resulting heteroduplexes were transfected into bacterial cells lacking the MutS protein. These bacterial cells were MMR deficient so correction of the duplex could only be achieved by proteins in the cell extracts. Repair was scored by examining plaque color. If the mismatch is corrected, the *lacZ* gene is functional and will result in blue colored plaque, if the mismatch is uncorrected, the *lacZ* gene is not functional and will result in a white colored plaque.

Only eleven species of fully sequenced Archaea have homologs to genes of the MMR pathway including *mutS1*, *mutS2,*and *mutL*. These include 4 halophiles, *Haloarcula marismortui, Halorubrum lacusprofundi, Haloquadratum waisbye,* and *Natronomonas pharaonis,* and several closely related methanogens from the genera *Methanococcoides, Methanosaeta, Methanosarcina, Methanoculleus,* and *Methanospirillum*. The identification of these MutS1, MutS2, and MutL homologs was based on sequence comparison and the cellular and biochemical functions of the MutS and MutL archaeal proteins have not been characterized. A study of the MMR pathway in the Archaea has not been undertaken and it is not known whether the Archaea use a MMR pathway similar to that of Bacteria and Eukarya or if there is an archaeal-specific pathway for correcting errors from DNA replication. Here, we propose to carry out an *in vitro* characterization of the MutS and MutL homologs in *Halobacterium* using a mismatch assay and *Halobacterium* cell extracts of mutants *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2*, and *ΔmutL*. This assay utilizes a circular plasmid, constructed of M13mp2 phage DNA, containing a mismatch within the *lacZα*-complementation gene. After incubation with cell extracts, purified DNA is transfected into a MMR-deficient *E. coli* strain. The assay

is scored as described above. Using this assay, we should be able to determine (1) is there is a MMR activity present in the cell extracts of wildtype *Halobacterium* and (2) if this activity varies in *Halobacterium* MMR gene deletion mutants.

B.2 Materials and Methods

Description of Assay:

The following assay from Thomas *et al* provides a method for determining the repair of mismatches produced during DNA replication (See Figures B-1 and B-2) [170]. A circular double stranded DNA heteroduplex is prepared using wildtype and mutant M13mp2 phage derivatives. Linear replicative form (RF) DNA is digested with an endonuclease that cuts only once. The digested RF DNA is then hybridized to a M13mp2 viral strand. These strands are complementary except for one mutation in the *lacZ*α-complementation gene. The hybridization of these strands forms a heteroduplex containing a nick in the RF strand and a mismatch in one location on the *lacZ*α-complementation gene. This mismatch confers a blue plaque phenotype to one strand and a colorless phenotype to the other strand. A blue plaque phenotype will occur if the *lacZ* gene is functional, if the *lacZ* gene is not functional, *i.e.* because of an incorrect nucleotide, the plaques will show a colorless phenotype. The heteroduplexes containing a GT and CA mismatch, along with a homoduplex as a control, were incubated with both *E. coli* and *Halobacterium* cell extracts and purified. The resulting DNA is transfected into *E. coli* strain NR9162, which lacks the MMR protein MutS, and plaque colors are scored. M13 plaques can be blue, colorless, or mixture of the two. Mixed plaques are a result of the different phenotype between the strands in the heteroduplex. Repair efficiency is calculated as (1 – (treated/untreated) x mixed bursts) x 100.

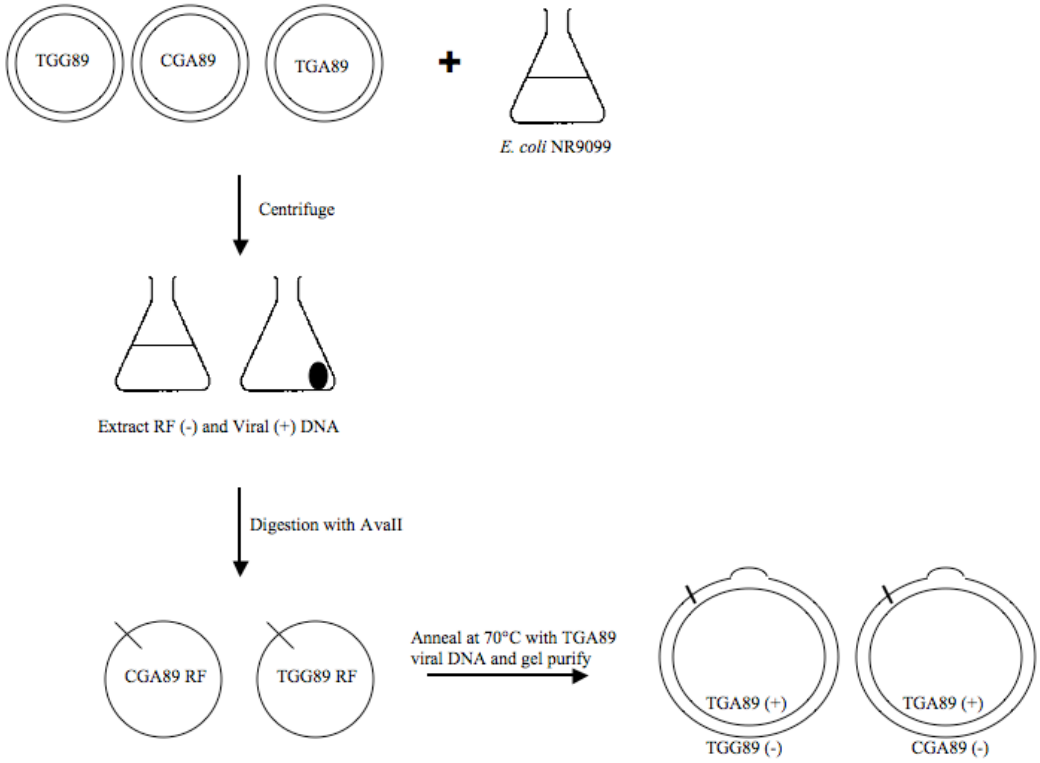


Figure B-1. Flow diagram for constructing the heteroduplexes used in the *in vitro* mismatch repair assay. Phage plaques and *E. coli* NR9099 were mixed and grown overnight at 37°C. Cells were centrifuged and replicative form (RF) DNA from the pellet (black dot) and viral form from the supernatant were extracted using Qiagen kits. RF DNA was digested with *AvaII* endonuclease and heated to 70°C. Viral DNA was added and heated an additional 2 minutes. Heteroduplexes were run out on a 1% agarose gel at 100V/cm for 25 minutes and gel purified.

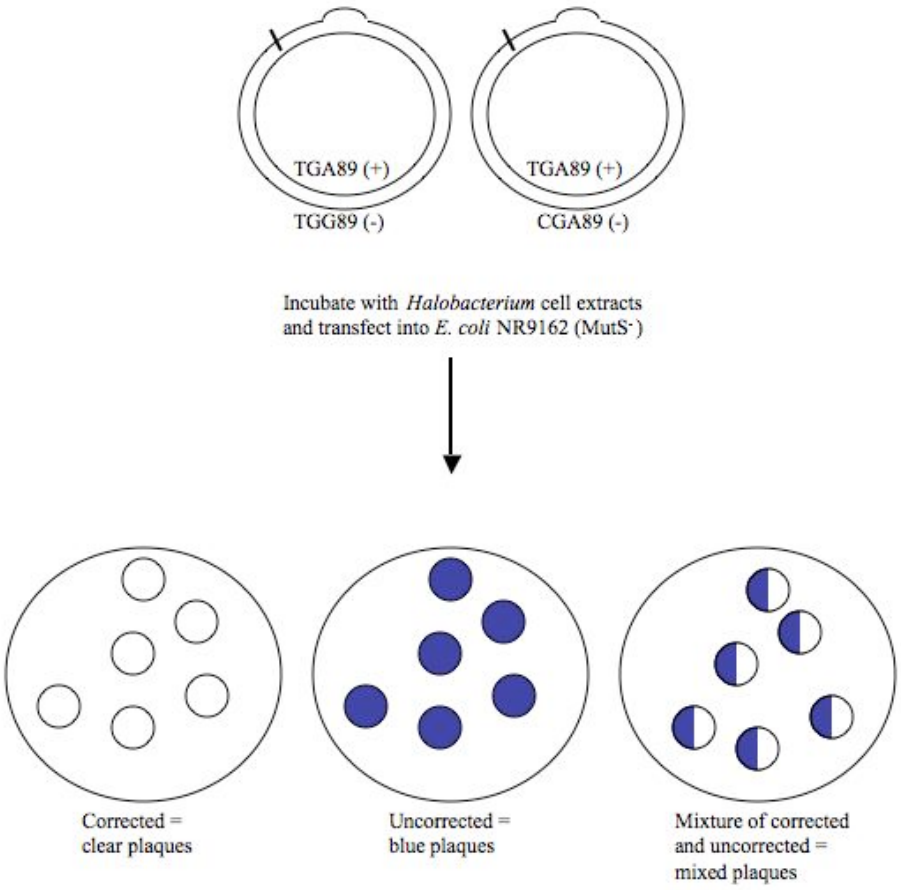


Figure B-2. Flow diagram for measuring heteroduplex repair in *Halobacterium* cell extracts. Heteroduplexes were incubated with mutant cell extracts of *Halobacterium* including *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2,* and *ΔmutL.* Resulting reactions were precipitated, extracted with phenol and chloroform, and resuspended in DEPC treated water. Repair reactions were transfected into *E. coli* NR9162, a MutS$^-$ strain, treated with IPTG and X-gal, and mixed with soft agar and *E. coli* CSH50 cells. Mixture was plated on minimal media plates, incubated overnight at 37°C, and phage plaques were scored for color and repair efficiency.

Organisms and Growth Conditions:

Halobacterium mutant strains *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2,* and *ΔmutL* and the background strain *Δura3* were used in this study. Mutant strains were grown in GN101 media [250g/L NaCl, 20g/L $MgSO_4$, 2g/L KCl, 3g/L sodium citrate, 10g/L Oxoid brand bacteriological peptone] with the addition of 1 mL/L trace elements solution [31.5mg/L $FeSO_4 \cdot 7H_2O$, 4.4mg/L $ZnSO_4 \cdot 7H_2O$, 3.3mg/L $MnSO_4 \cdot H_2O$, 0.1mg/L $CuSO_4 \cdot 5H_2O$] shaking in a Gyromax 737 shaker (Amerex Instruments, LaFayette, CA) at 220rpm supplemented with 50mg/L uracil.

Competent cells of *E. coli* strains NR9162 ($MutS^-$) and MC1061 ($MutS^+$) were made using a previously described protocol [170]. Five milliliters of an overnight culture was added to 500mL 2X YT media and grown to an OD_{600} 0.60 at 37°C shaking at 220rpm. The flasks were iced for 30 minutes and cultures centrifuged in 50mL centrifuge tubes at 4000 x g for 30 minutes at 4°C. Pellets were resuspended in 500mL cold sterile water and centrifuged again at 2200 x g for 20 minutes at 4°C. Resulting pellets were resuspended in 250mL cold sterile water and centrifuged again at the previous conditions. After centrifugation, pellets were resuspended in 10mL 10% cold glycerol and centrifuged at 3000 x g for 15 minutes at 4°C. Lastly, the pellets were resuspended in 1.5mL cold 10% glycerol, flash-frozen in dry ice/ethanol, and stored at -80°C.

E. coli strains and M13mp2 phage derivatives were a gift from Tom Kunkel at Research Triangle Park in North Carolina and genotypes are referenced in [170]. *E. coli* strains NR9162 and MC1061 were cultured on LB plates (10g/L tryptone, 10g/L NaCl, 5g/L

yeast extract, pH 7.0, 15g/L agar) whereas strains CSH50 and NR9099 were cultured on minimal media plates (16g/L agar, 0.3mL/L 0.1M IPTG, 20mL/L 50X VB salts, 20mL/L 60% glucose, 5mL/L thiamine HCL) to maintain the F' plasmid. Strains NR9162 ($MutS^-$) and MC1061 ($MutS^+$) were used for heteroduplex transfection, strain CSH50 was used for *lacZα*-complementation, and strain NR9099 was used for phage preparation.

Stock solutions were as follows. The 50X VB salts were prepared as 10g/L $MgSO_4 \bullet 7H_2O$, 100g/L citric acid, 500g/L K_2HPO_4, and 175g/L $Na(NH_4)HPO_4 \bullet 4H_2O$. Soft agar was made using 8% agar in distilled water. The 2X YT media consisted of 16g/L tryptone, 10g/L yeast extract, and 10g/L NaCl at pH 7.4. Qiagen kits, Qiagen plasmid mini kit and Qiagen QIAprep M13 kit, were used to extract RF and viral form DNA respectively.

Phage Growth and DNA Preparation:

Phage derivatives, TGA89, TGG89, and CGA89 were diluted in LB and 100μL was mixed with 100μL fresh overnight culture of *E. coli* NR9099. Mixture was incubated at room temperature for 10-15 minutes, added to 3mL of soft agar (heated to 50°C), and plated on minimal media plates. Plates were incubated overnight at 37°C. A single plaque was added to 2X YT medium containing 1/10 volume of *E. coli* NR9099 and grown overnight at 37°C. The culture was centrifuged at 5000 x g for 30 minutes and RF and viral DNA was extracted using Qiagen kits. The pellet will contain the RF DNA and the viral DNA is in the supernatant. DNA is quantified on a 1% agarose gel at 100V/cm for 25 minutes.

Preparation of Heteroduplex:

RF DNA (6μg) from TGG89 and CGA89 phage derivatives was digested with AvaII endonuclease. AvaII will cut the DNA to the left of the mismatch. Digest was confirmed by removing 1μg of digested DNA to check on an agarose gel. The remaining 5μg (DNA concentration < 100ng/μL) was heated at 70°C for 15 minutes. Viral DNA (15μg), TGA89 phage derivative, was added to mixture and heated an additional 2 minutes. This will form two different mismatches; TGG89 RF and TGA89 viral DNA will create a C/A mismatch and CGA89 RF and TGA89 viral DNA will form a G/T mismatch. The solution placed on ice for 2 minutes and centrifuged to ensure all DNA is at bottom of tube. SSC (20X) was added to a final concentration of 2X and incubated on ice for an additional 15 minutes. Solution was heated at 65°C before loading on gel. Gel was 1% agarose and includes a molecular weight marker, viral DNA, uncut RF DNA, cut RF DNA, and annealing reaction. Gel was run at 100V/cm for 25 minutes and homoduplexes and heteroduplexes were gel excised using a Qiagen gel extraction kit.

Preparation of *E. coli* cell extracts:

E. coli DH5α cells were grown in LB broth supplemented with 0.1% glucose until OD_{600} 1.0-1.2. Cultures were centrifuged at 8000 x g for 3 minutes at 4°C and pellets were resuspended in 2mL 0.05M Tris pH 7.6 and 10% sucrose. Pellets were freeze dried in a dry ice/ethanol bath. Cells were lysed using the previously described method [180]. Briefly, cells are lysed by adding 0.23mg lysozyme, 1.2mM dTT, 0.15M KCl, and 4M NaCl and freeze/thawing. After lysis, cells are centrifuged and the supernatant is treated

with 0.42g/mL $(NH_4)_2SO_4$. Cell extract was precipitated at 8000 x g for 15 minutes at 4°C, resuspended in 0.025M Hepes pH 7.6, 0.1mM EDTA, 2mM dTT, and100mM KCl, and dialyzed against the same buffer for 90 minutes. Protein concentration was quantified using the Bradford assay [181]. Samples were flash frozen and stored at -80°C in small aliquots.

Preparation of *Halobacterium* cell extracts:

Halobacterium cultures *ΔmutS1, ΔmutS2, ΔmutS1ΔmutS2, ΔmutL,* and *Δura3* were started from a single colony in 5 ml GN101 media supplemented with 50mg/L uracil. Cultures were diluted to an OD_{600} 0.05 and grown until OD_{600} 0.60 for 10mL of culture. Cells were centrifuged at 8000 x g for 15 minutes at 4°C and pellets resuspended in either 1M, 2M, or 4M salt buffer. The 1M and 2M salt buffer consisted of 50mM potassium phosphate buffer pH 7.0, 1M or 2M NaCl, and 1% 2-mercaptoethanol. The 4M salt buffer contained 20mM potassium phosphate pH 7.4, 4M NaCl, and 1% 2-mercaptoethanol. Resulting solutions were sonicated for 30 seconds ON/OFF for 3 minutes on output setting 5 (VirSonic soncator, Virtis Corporation). Sonicated cells were centrifuged at 8000 x g for 20 minutes at 4°C. Supernatant was recovered and protein concentration was quantified using a Bradford assay [181]. Extracts were frozen in small aliquots at -80°C.

Mismatch repair assay:

Mismatch repair reactions were run in 25μL amounts and contained 30mM Hepes pH 7.8, 7mM $MgCl_2$, 4mM ATP, 100μM each dNTP, 15mM sodium phosphate pH 7.5, 5ng purified heteroduplex, and 50μg cell extract. Reactions were incubated at 37°C for 15

minutes and stopped with 2mg/mL proteinase K, 2% SDS, and 50mM EDTA at 37°C for 30 minutes. Resultant reactions were precipitated with 0.71mg/mL *E. coli* tRNA, 1.7M ammonium acetate, and an equal volume of isopropanol. Pellets were resuspended in 50μL TE pH 7.0 and extracted with phenol and chloroform. Final pellets were resuspended in DEPC treated water.

Electroporation and plating:

MMR reactions (1μL) were diluted in 50μL DEPC treated water and added to 50μL competent cells. Transfection was accomplished with an electroporator (Bio-Rad E. coli Pulser G-560) at 1800V in a 0.2cm cuvette. Immediately after electroporation, 1mL of SOC medium was added to cells. Following electroporation, 50-100μL of transfected cells were added to 2.5mL soft agar (heated to 50°C), treated with 500μg IPTG and 2.5μg X-gal, and mixed with 400μL CSH50 cells. Mixture was plated on minimal plates and incubated overnight at 37°C. Phage plaques were scored for color and repair efficiencies calculated using the formula: (1 – (treated/untreated) x mixed bursts) x 100.

B.3 Results: challenges in the development of the *in vitro* assay

We developed an *in vitro* assay to test the mismatch repair efficiency of cell extracts from *Δura3* (background), and the *ΔmutL, ΔmutS1,* and *ΔmutS1ΔmutS2* deletion strains of *Halobacterium*. Multiple problems were encountered during the course of this experiment. Cell extracts of *E. coli* (DH5α) and *Halobacterium* strains *Δura3, ΔmutL, ΔmutS1, ΔmutS2,* and *ΔmutS1ΔmutS2* were made successfully. Competent cells were made from the MC1061 (MutS+) strain of *E. coli* with ease but the construction of

competent cells in the NR9162 (MutS$^-$) strain of *E. coli* was challenging. Eventually, a new aliquot of NR9162 cells from the Kunkel lab was delivered and the construction of competent cells achieved. In the meanwhile, homoduplexes from the TGA phage derivative along with heteroduplexes containing a GT or a CA mismatch were constructed and purified. As a positive control, we transformed the homoduplex into a NR9162 (MutS$^-$) strain of *E. coli* and transfected into the *lacZα*-complementation strain, CSH50. Plaque formation was not seen after transfection with the complementation strain. Further analyses showed that the phage stock used to extract viral DNA for the duplex construction was contaminated, likely with *E. coli* strain NR9099 due to incomplete phage lysis of the bacteria and recovery. Problems were encountered during the production of new phage stock. Phage stock from the Kunkel lab was transfected into *E. coli* strain NR9099 and individual plaques picked into fresh media containing NR9099. Viral stranded DNA was extracted from these plaques and a phage stock made. Transfection of new phage stock into NR9099 resulted in no plaque formation. Troubleshooting included changing the media used in transfection, transfecting both exponential phase and stationary phase NR9099, incubating the transfection reaction for different time periods, and using different dilutions of phage stock for transfection.

B.4 Discussion

We know from genomic mutation rate analysis that *Halobacterium* has a low incidence of mutation. Results from phenotypic characterization of *mutS1*, *mutS2*, and *mutL* deletion strains in *Halobacterium* demonstrate that these bacterial-like MMR homologs are not essential. It is surprising that the MutS and MutL homologs found in

Halobacterium do not show a mutator phenotype when deleted. A study looking at the bacterial-like UvrA/B/C proteins in *Halobacterium* showed that they are essential for the nucleotide excision repair of UV damage in the absence of light [88]. This leads into the question of what is responsible for the low incidence of mutation. General hypotheses described in Chapter 3 are (1) high fidelity of the archaeal polymerase, and (2) recruitment of an archaeal-specific pathway. Benefits of a high fidelity polymerase would be a decreased incidence of replication error resulting in a low requirement for MMR. Alternatively, if there is a requirement for MMR, recruitment of an archaeal-specific system is likely. In both of these hypotheses, the bacterial-like MutS1, MutS2, and MutL proteins found in *Halobacterium* would not be essential.

To distinguish between these two hypotheses, we used an *in vitro* assay to test the capability of *Halobacterium* wildtype and MMR deletion strains to repair mismatches. The assay we are using is modified from an *in vitro* assay developed to look at heteroduplex repair in human HeLa extracts [170]. This assay demonstrated the viability of using M13mp2 heteroduplexes to assay repair. In HeLa extracts, after transfection of the heteroduplex into MMR deficient *E. coli,* approximately 55% repair efficiency was calculated [179]. Transfection of the heteroduplexes into a MMR+ strain of *E. coli* resulted in an approximately 72% repair efficiency [179]. While this assay has not been completed using *E. coli* cell extracts, we expected similar findings to what has been seen in HeLa extracts because the MMR pathways, with the exception of the strand discrimination signal, are very similar [3, 4, 179]. A previous study utilized cell extracts of *E. coli* to assay repair of a mismatch within a restriction site [13]. While this assay

differs from ours in several ways, it provides evidence that cell extracts of *E. coli* can stimulate MMR.

We were unable to complete this assay using *E. coli* cell extracts. Problems arose with the phage stock we used to create the homoduplexes and heteroduplexes. We anticipate that this assay could also be successful in *Halobacterium,* however difficulties arise in determining the concentration of salts needed in the *Halobacterium* cell extract for proper function of the proteins. Holmes *et al* showed that enzymes in *Halobacterium* cell extracts are unstable at a concentration of less than 3.4M NaCl [182]. Further optimization will be needed to successfully use the above *in vitro* assay.

Bibliography

1. Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell, Fourth Edition.* 2002, New York, NY: Garland Science.
2. Zharkov, D.O., *Base excision DNA repair.* Cell Mol Life Sci, 2008. **65**(10): p. 1544-65.
3. Harfe, B.D. and S. Jinks-Robertson, *DNA mismatch repair and genetic instability.* Annu Rev Genet, 2000. **34**: p. 359-399.
4. Schofield, M.J. and P. Hsieh, *DNA mismatch repair: molecular mechanisms and biological function.* Annu Rev Microbiol, 2003. **57**: p. 579-608.
5. Iyer, R.R., et al., *DNA mismatch repair: functions and mechanisms.* Chem Rev, 2006. **106**(2): p. 302-23.
6. Lin, Z., M. Nei, and H. Ma, *The origins and early evolution of DNA mismatch repair genes multiple horizontal gene transfers and co-evolution.* Nucleic Acids Res, 2007.
7. Madigan, A., and Parker, Eds. , *Brock Biology of Microorganisms 9th Edition.*
8. Barbeyron, T., K. Kean, and P. Forterre, *DNA adenine methylation of GATC sequences appeared recently in the Escherichia coli lineage.* J Bacteriol, 1984. **160**(2): p. 586-90.
9. Pukkila, P.J., et al., *Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in Escherichia coli.* Genetics, 1983. **104**(4): p. 571-82.
10. Herman, G.E. and P. Modrich, *Escherichia coli K-12 clones that overproduce dam methylase are hypermutable.* J Bacteriol, 1981. **145**(1): p. 644-6.
11. Glickman, B.W. and M. Radman, *Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction.* Proc Natl Acad Sci U S A, 1980. **77**(2): p. 1063-7.
12. Cupples, C.G. and J.H. Miller, *A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions.* Proc Natl Acad Sci U S A, 1989. **86**(14): p. 5345-9.
13. Lu, A.L., S. Clark, and P. Modrich, *Methyl-directed repair of DNA base-pair mismatches in vitro.* Proc Natl Acad Sci U S A, 1983. **80**(15): p. 4639-43.
14. Wang, H. and J.B. Hays, *Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions.* Mol Biotechnol, 2001. **19**(2): p. 133-40.
15. Lamers, M.H., et al., *The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch.* Nature, 2000. **407**(6805): p. 711-7.
16. Obmolova, G., et al., *Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA.* Nature, 2000. **407**(6805): p. 703-10.
17. Mendillo, M.L., C.D. Putnam, and R.D. Kolodner, *Escherichia coli MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo.* J Biol Chem, 2007. **282**(22): p. 16345-54.
18. Acharya, S., et al., *The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair.* Mol Cell, 2003. **12**(1): p. 233-46.

19. Bergerat, A., et al., *An atypical topoisomerase II from Archaea with implications for meiotic recombination.* Nature, 1997. **386**(6623): p. 414-7.
20. Hall, M.C., J.R. Jordan, and S.W. Matson, *Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD.* Embo J, 1998. **17**(5): p. 1535-41.
21. Hall, M.C. and S.W. Matson, *The Escherichia coli MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity.* J Biol Chem, 1999. **274**(3): p. 1306-12.
22. Welsh, K.M., et al., *Isolation and characterization of the Escherichia coli mutH gene product.* J Biol Chem, 1987. **262**(32): p. 15624-9.
23. Au, K.G., K. Welsh, and P. Modrich, *Initiation of methyl-directed mismatch repair.* J Biol Chem, 1992. **267**(17): p. 12142-8.
24. Lahue, R.S., K.G. Au, and P. Modrich, *DNA mismatch correction in a defined system.* Science, 1989. **245**(4914): p. 160-4.
25. Bateman, A., et al., *The Pfam protein families database.* Nucleic Acids Res, 2004. **32**(Database issue): p. D138-41.
26. Moreira, D. and H. Philippe, *Smr: a bacterial and eukaryotic homologue of the C-terminal region of the MutS2 family.* Trends Biochem Sci, 1999. **24**(8): p. 298-300.
27. Pinto, A.V., et al., *Suppression of homologous and homeologous recombination by the bacterial MutS2 protein.* Mol Cell, 2005. **17**(1): p. 113-20.
28. Kang, J., S. Huang, and M.J. Blaser, *Structural and functional divergence of MutS2 from bacterial MutS1 and eukaryotic MSH4-MSH5 homologs.* J Bacteriol, 2005. **187**(10): p. 3528-37.
29. Yamaguchi, M., V. Dao, and P. Modrich, *MutS and MutL activate DNA helicase II in a mismatch-dependent manner.* J Biol Chem, 1998. **273**(15): p. 9197-201.
30. Burdett, V., et al., *In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair.* Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6765-70.
31. Viswanathan, M. and S.T. Lovett, *Exonuclease X of Escherichia coli. A novel 3'-5' DNase and Dnaq superfamily member involved in DNA repair.* J Biol Chem, 1999. **274**(42): p. 30094-100.
32. Matson, S.W. and A.B. Robertson, *The UvrD helicase and its modulation by the mismatch repair protein MutL.* Nucleic Acids Res, 2006. **34**(15): p. 4089-97.
33. *Mismatch Repair Pathway in E. coli.* . [cited; Available from: www.sinauer.com/cooper4e/sample/Figures/Chapter%2006/lowres/CELL4e-Fig-06-0.jpg.
34. Mennecier, S., et al., *Mismatch repair ensures fidelity of replication and recombination in the radioresistant organism Deinococcus radiodurans.* Mol Genet Genomics, 2004. **272**(4): p. 460-9.
35. Alani, E., et al., *Genetic and biochemical analysis of Msh2p-Msh6p: role of ATP hydrolysis and Msh2p-Msh6p subunit interactions in mismatch base pair recognition.* Mol Cell Biol, 1997. **17**(5): p. 2436-47.
36. Harfe, B.D., B.K. Minesinger, and S. Jinks-Robertson, *Discrete in vivo roles for the MutL homologs Mlh2p and Mlh3p in the removal of frameshift intermediates in budding yeast.* Curr Biol, 2000. **10**(3): p. 145-8.

37. Habraken, Y., et al., *Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3.* Curr Biol, 1996. **6**(9): p. 1185-7.
38. Prolla, T.A., D.M. Christie, and R.M. Liskay, *Dual requirement in yeast DNA mismatch repair for MLH1 and PMS1, two homologs of the bacterial mutL gene.* Mol Cell Biol, 1994. **14**(1): p. 407-15.
39. Marsischky, G.T., et al., *Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair.* Genes Dev, 1996. **10**(4): p. 407-20.
40. Alani, E., *The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs.* Mol Cell Biol, 1996. **16**(10): p. 5604-15.
41. Williamson, M.S., J.C. Game, and S. Fogel, *Meiotic gene conversion mutants in Saccharomyces cerevisiae. I. Isolation and characterization of pms1-1 and pms1-2.* Genetics, 1985. **110**(4): p. 609-46.
42. Fang, W.H. and P. Modrich, *Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction.* J Biol Chem, 1993. **268**(16): p. 11838-44.
43. Wang, H. and J.B. Hays, *Mismatch repair in human nuclear extracts. Time courses and ATP requirements for kinetically distinguishable steps leading to tightly controlled 5' to 3' and aphidicolin-sensitive 3' to 5' mispair-provoked excision.* J Biol Chem, 2002. **277**(29): p. 26143-8.
44. Wang, H. and J.B. Hays, *Mismatch repair in human nuclear extracts. Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides.* J Biol Chem, 2002. **277**(29): p. 26136-42.
45. Johnson, R.E., et al., *Requirement of the yeast RTH1 5' to 3' exonuclease for the stability of simple repetitive DNA.* Science, 1995. **269**(5221): p. 238-40.
46. Tran, H.T., D.A. Gordenin, and M.A. Resnick, *The 3'-->5' exonucleases of DNA polymerases delta and epsilon and the 5'-->3' exonuclease Exo1 have major roles in postreplication mutation avoidance in Saccharomyces cerevisiae.* Mol Cell Biol, 1999. **19**(3): p. 2000-7.
47. Dzantiev, L., et al., *A defined human system that supports bidirectional mismatch-provoked excision.* Mol Cell, 2004. **15**(1): p. 31-41.
48. Kunkel, T.A. and D.A. Erie, *DNA mismatch repair.* Annu Rev Biochem, 2005. **74**: p. 681-710.
49. Umar, A., et al., *Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis.* Cell, 1996. **87**(1): p. 65-73.
50. Modrich, P. and R. Lahue, *Mismatch repair in replication fidelity, genetic recombination, and cancer biology.* Annu Rev Biochem, 1996. **65**: p. 101-33.
51. Pavlov, Y.I., C.S. Newlon, and T.A. Kunkel, *Yeast origins establish a strand bias for replicational mutagenesis.* Mol Cell, 2002. **10**(1): p. 207-13.
52. Pavlov, Y.I., I.M. Mian, and T.A. Kunkel, *Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast.* Curr Biol, 2003. **13**(9): p. 744-8.

53. Clark, A.B., et al., *Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes.* J Biol Chem, 2000. **275**(47): p. 36498-501.
54. Bertrand, P., et al., *Physical interaction between components of DNA mismatch repair and nucleotide excision repair.* Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14278-83.
55. Zhao, J. and M.E. Winkler, *Reduction of GC --> TA transversion mutation by overexpression of MutS in Escherichia coli K-12.* J Bacteriol, 2000. **182**(17): p. 5025-8.
56. Ni, T.T., G.T. Marsischky, and R.D. Kolodner, *MSH2 and MSH6 are required for removal of adenine misincorporated opposite 8-oxo-guanine in S. cerevisiae.* Mol Cell, 1999. **4**(3): p. 439-44.
57. Karran, P. and M.G. Marinus, *Mismatch correction at O6-methylguanine residues in E. coli DNA.* Nature, 1982. **296**(5860): p. 868-9.
58. Kat, A., et al., *An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair.* Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6424-8.
59. Branch, P., et al., *Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage.* Nature, 1993. **362**(6421): p. 652-4.
60. Bawa, S. and W. Xiao, *A mutation in the MSH5 gene results in alkylation tolerance.* Cancer Res, 1997. **57**(13): p. 2715-20.
61. Wang, T.F., N. Kleckner, and N. Hunter, *Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13914-9.
62. Borts, R.H., S.R. Chambers, and M.F. Abdullah, *The many faces of mismatch repair in meiosis.* Mutat Res, 2000. **451**(1-2): p. 129-50.
63. Saparbaev, M., L. Prakash, and S. Prakash, *Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in Saccharomyces cerevisiae.* Genetics, 1996. **142**(3): p. 727-36.
64. Sugawara, N., et al., *Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination.* Proc Natl Acad Sci U S A, 1997. **94**(17): p. 9214-9.
65. Drake, J.W., et al., *Rates of spontaneous mutation.* Genetics, 1998. **148**(4): p. 1667-86.
66. Drake, J.W., *A constant rate of spontaneous mutation in DNA-based microbes.* Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7160-4.
67. Grogan, D.W., G.T. Carver, and J.W. Drake, *Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius.* Proc Natl Acad Sci U S A, 2001. **98**(14): p. 7928-33.
68. Mackwan, R.R., et al., *An Unusual Pattern of Spontaneous Mutations Recovered in the Halophilic Archaeon Haloferax volcanii.* Genetics, 2006.
69. Barry, E.R. and S.D. Bell, *DNA replication in the archaea.* Microbiol Mol Biol Rev, 2006. **70**(4): p. 876-87.

70. Ishino, Y. and I.K. Cann, *The euryarchaeotes, a subdomain of Archaea, survive on a single DNA polymerase: fact or farce?* Genes Genet Syst, 1998. **73**(6): p. 323-36.
71. Ishino, Y., et al., *A novel DNA polymerase family found in Archaea.* J Bacteriol, 1998. **180**(8): p. 2232-6.
72. Lundberg, K.S., et al., *High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus.* Gene, 1991. **108**(1): p. 1-6.
73. Mattila, P., et al., *Fidelity of DNA synthesis by the Thermococcus litoralis DNA polymerase--an extremely heat stable enzyme with proofreading activity.* Nucleic Acids Res, 1991. **19**(18): p. 4967-73.
74. Cline, J., J.C. Braman, and H.H. Hogrefe, *PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases.* Nucleic Acids Res, 1996. **24**(18): p. 3546-51.
75. Kunkel, T.A., *DNA replication fidelity.* J Biol Chem, 1992. **267**(26): p. 18251-4.
76. Bloom, L.B., et al., *Fidelity of Escherichia coli DNA polymerase III holoenzyme. The effects of beta, gamma complex processivity proteins and epsilon proofreading exonuclease on nucleotide misincorporation efficiencies.* J Biol Chem, 1997. **272**(44): p. 27919-30.
77. Grogan, D.W., *Stability and repair of DNA in hyperthermophilic Archaea.* Curr Issues Mol Biol, 2004. **6**(2): p. 137-44.
78. Vijayvargia, R. and I. Biswas, *MutS2 family protein from Pyrococcus furiosus.* Curr Microbiol, 2002. **44**(3): p. 224-8.
79. Ng, W.V., et al., *Genome sequence of Halobacterium species NRC-1.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12176-81.
80. Breuert, S., et al., *Regulated polyploidy in halophilic archaea.* PLoS ONE, 2006. **1**: p. e92.
81. Dassarma, S., *Saline Systems: A research journal bridging gene systems and ecosystems.* Saline Systems, 2005. **1**: p. 1.
82. Martin, E.L., et al., *The effects of ultraviolet radiation on the moderate halophile Halomonas elongata and the extreme halophile Halobacterium salinarum.* Can J Microbiol, 2000. **46**(2): p. 180-7.
83. Potts, M., *Desiccation tolerance of prokaryotes.* Microbiol Rev, 1994. **58**(4): p. 755-805.
84. Baliga, N.S., et al., *Systems level insights into the stress response to UV radiation in the halophilic archaeon Halobacterium NRC-1.* Genome Res, 2004. **14**(6): p. 1025-35.
85. Whitehead, K., et al., *An integrated systems approach for understanding cellular responses to gamma radiation.* Mol Syst Biol, 2006. **2**: p. 47.
86. Kottemann, M., et al., *Physiological responses of the halophilic archaeon Halobacterium sp. strain NRC1 to desiccation and gamma irradiation.* Extremophiles, 2005. **9**(3): p. 219-27.
87. Kish, A. and J. Diruggiero, *Rad50 is Not Essential for the Mre11-Dependant Repair of DNA Double Strand Breaks in Halobacterium sp. str. NRC-1.* J Bacteriol, 2008. **190**(15): p. 5210-6.

88. Crowley, D.J., et al., *The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1.* Saline Systems, 2006. **2**: p. 11.
89. Woods, W.G. and M.L. Dyall-Smith, *Construction and analysis of a recombination-deficient (radA) mutant of Haloferax volcanii.* Mol Microbiol, 1997. **23**(4): p. 791-7.
90. Salerno, V., et al., *Transcriptional response to DNA damage in the archaeon Sulfolobus solfataricus.* Nucleic Acids Res, 2003. **31**(21): p. 6127-38.
91. Lee Bi, B.I., et al., *Molecular interactions of human Exo1 with DNA.* Nucleic Acids Res, 2002. **30**(4): p. 942-9.
92. Biswas, I., et al., *Disruption of the helix-u-turn-helix motif of MutS protein: loss of subunit dimerization, mismatch binding and ATP hydrolysis.* J Mol Biol, 2001. **305**(4): p. 805-16.
93. Schofield, M.J., et al., *The Phe-X-Glu DNA binding motif of MutS. The role of hydrogen bonding in mismatch recognition.* J Biol Chem, 2001. **276**(49): p. 45505-8.
94. Kelman, L.M. and Z. Kelman, *Archaea: an archetype for replication initiation studies?* Mol Microbiol, 2003. **48**(3): p. 605-15.
95. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.
96. Kelman, L.M. and Z. Kelman, *Multiple origins of replication in archaea.* Trends Microbiol, 2004. **12**(9): p. 399-401.
97. Riley, P.A., *Free radicals in biology: oxidative stress and the effects of ionizing radiation.* Int J Radiat Biol, 1994. **65**(1): p. 27-33.
98. Hutchinson, F., *Chemical changes induced in DNA by ionizing radiation.* Prog Nucleic Acid Res Mol Biol, 1985. **32**: p. 115-54.
99. Schaaper, R.M. and R.L. Dunn, *Spontaneous mutation in the Escherichia coli lacI gene.* Genetics, 1991. **129**(2): p. 317-26.
100. Schaaper, R.M., B.N. Danforth, and B.W. Glickman, *Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the Escherichia coli lacI gene.* J Mol Biol, 1986. **189**(2): p. 273-84.
101. Hartman, P.E., Z. Hartman, and R.C. Stahl, *Classification and mapping of spontaneous and induced mutations in the histidine operon of Salmonella.* Adv Genet, 1971. **16**: p. 1-34.
102. Lee, G.S., et al., *The base-alteration spectrum of spontaneous and ultraviolet radiation-induced forward mutations in the URA3 locus of Saccharomyces cerevisiae.* Mol Gen Genet, 1988. **214**(3): p. 396-404.
103. Whelan, W.L., E. Gocke, and T.R. Manney, *The CAN1 locus of Saccharomyces cerevisiae: fine-structure analysis and forward mutation rates.* Genetics, 1979. **91**(1): p. 35-51.
104. Halliday, J.A. and B.W. Glickman, *Mechanisms of spontaneous mutation in DNA repair-proficient Escherichia coli.* Mutat Res, 1991. **250**(1-2): p. 55-71.
105. Farabaugh, P.J., et al., *Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots in the lacI gene of Escherichia coli.* J Mol Biol, 1978. **126**(4): p. 847-57.

106. Lanyi, J.K., *Salt-dependent properties of proteins from extremely halophilic bacteria.* Bacteriol Rev, 1974. **38**(3): p. 272-90.
107. http://www.genome.jp/dbget-bin/get_pathway?org_name=hal&mapno=00240. *KEGG Pyrimidine Metabolic Pathway: Halobacterium sp. strain NRC-1*. [cited.
108. Bitan-Banin, G., R. Ortenberg, and M. Mevarech, *Development of a gene knockout system for the halophilic archaeon Haloferax volcanii by use of the pyrE gene.* J Bacteriol, 2003. **185**(3): p. 772-8.
109. Rosche, W.A. and P.L. Foster, *Determining mutation rates in bacterial populations.* Methods, 2000. **20**(1): p. 4-17.
110. Luria, S.E. and M. Delbruck, *Mutations of Bacteria from Virus Sensitivity to Virus Resistance.* Genetics, 1943. **28**(6): p. 491-511.
111. Jacobs, K.L. and D.W. Grogan, *Rates of spontaneous mutation in an archaeon from geothermal environments.* J Bacteriol, 1997. **179**(10): p. 3298-303.
112. Chen, L., et al., *The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota.* J Bacteriol, 2005. **187**(14): p. 4992-9.
113. Lang, G.I. and A.W. Murray, *Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae.* Genetics, 2008. **178**(1): p. 67-82.
114. LeClerc, J.E., et al., *High mutation frequencies among Escherichia coli and Salmonella pathogens.* Science, 1996. **274**(5290): p. 1208-11.
115. Schaaper, R.M. and R.L. Dunn, *Spectra of spontaneous mutations in Escherichia coli strains defective in mismatch correction: the nature of in vivo DNA replication errors.* Proc Natl Acad Sci U S A, 1987. **84**(17): p. 6220-4.
116. Schaaper, R.M., *Base selection, proofreading, and mismatch repair during DNA replication in Escherichia coli.* J Biol Chem, 1993. **268**(32): p. 23762-5.
117. Xiao, W., et al., *DNA mismatch repair mutants do not increase N-methyl-N'-nitro-N-nitrosoguanidine tolerance in O6-methylguanine DNA methyltransferase-deficient yeast cells.* Carcinogenesis, 1995. **16**(8): p. 1933-9.
118. Mellon, I. and G.N. Champe, *Products of DNA mismatch repair genes mutS and mutL are required for transcription-coupled nucleotide-excision repair of the lactose operon in Escherichia coli.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1292-7.
119. Mellon, I., et al., *Transcription-coupled repair deficiency and mutations in human mismatch repair genes.* Science, 1996. **272**(5261): p. 557-60.
120. Harfe, B.D. and S. Jinks-Robertson, *Mismatch repair proteins and mitotic genome stability.* Mutat Res, 2000. **451**(1-2): p. 151-67.
121. Peck, R.F., S. Dassarma, and M.P. Krebs, *Homologous gene knockout in the archaeon Halobacterium salinarum with ura3 as a counterselectable marker.* Mol Microbiol, 2000. **35**(3): p. 667-76.
122. Wang, G., et al., *Arsenic resistance in Halobacterium sp. strain NRC-1 examined by using an improved gene knockout system.* J Bacteriol, 2004. **186**(10): p. 3187-94.
123. Sambrook, J.a.R., D, *Molecular Cloning: A Laboratory Manual*. 2001.
124. Holmes, M.L. and M.L. Dyall-Smith, *Sequence and expression of a halobacterial beta-galactosidase gene.* Mol Microbiol, 2000. **36**(1): p. 114-22.

125. Hidaka, M., et al., *Trimeric crystal structure of the glycoside hydrolase family 42 beta-galactosidase from Thermus thermophilus A4 and the structure of its complex with galactose.* J Mol Biol, 2002. **322**(1): p. 79-91.
126. Nolling, J. and W.M. de Vos, *Identification of the CTAG-recognizing restriction-modification systems MthZI and MthFI from Methanobacterium thermoformicicum and characterization of the plasmid-encoded mthZIM gene.* Nucleic Acids Res, 1992. **20**(19): p. 5047-52.
127. Sweder, K.S., et al., *Mismatch repair mutants in yeast are not defective in transcription-coupled DNA repair of UV-induced DNA damage.* Genetics, 1996. **143**(3): p. 1127-35.
128. Lasken, R.S., D.M. Schuster, and A. Rashtchian, *Archaebacterial DNA polymerases tightly bind uracil-containing DNA.* J Biol Chem, 1996. **271**(30): p. 17692-6.
129. Yang, H., et al., *Direct interaction between uracil-DNA glycosylase and a proliferating cell nuclear antigen homolog in the crenarchaeon Pyrobaculum aerophilum.* J Biol Chem, 2002. **277**(25): p. 22271-8.
130. Connolly, B.A., et al., *Uracil recognition by archaeal family B DNA polymerases.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 699-702.
131. Fogg, M.J., L.H. Pearl, and B.A. Connolly, *Structural basis for uracil recognition by archaeal family B DNA polymerases.* Nat Struct Biol, 2002. **9**(12): p. 922-7.
132. Haber, J.E., et al., *Repairing a double-strand chromosome break by homologous recombination: revisiting Robin Holliday's model.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1441): p. 79-86.
133. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.
134. Dianov, G.L., P. O'Neill, and D.T. Goodhead, *Securing genome stability by orchestrating DNA repair: removal of radiation-induced clustered lesions in DNA.* Bioessays, 2001. **23**(8): p. 745-9.
135. Imlay, J.A. and S. Linn, *DNA damage and oxygen radical toxicity.* Science, 1988. **240**(4857): p. 1302-9.
136. Imlay, J.A., *Pathways of oxidative damage.* Annu Rev Microbiol, 2003. **57**: p. 395-418.
137. Aguirre, J., et al., *Reactive oxygen species and development in microbial eukaryotes.* Trends Microbiol, 2005. **13**(3): p. 111-8.
138. Keyer, K., A.S. Gort, and J.A. Imlay, *Superoxide and the production of oxidative DNA damage.* J Bacteriol, 1995. **177**(23): p. 6782-90.
139. Imlay, J.A., S.M. Chin, and S. Linn, *Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro.* Science, 1988. **240**(4852): p. 640-2.
140. Joshi, P. and P.P. Dennis, *Characterization of paralogous and orthologous members of the superoxide dismutase gene family from genera of the halophilic archaebacteria.* J Bacteriol, 1993. **175**(6): p. 1561-71.
141. May, B.P., P. Tam, and P.P. Dennis, *The expression of the superoxide dismutase gene in Halobacterium cutirubrum and Halobacterium volcanii.* Can J Microbiol, 1989. **35**(1): p. 171-5.

142. Schmid, A.K., et al., *The anatomy of microbial cell state transitions in response to oxygen.* Genome Res, 2007. **17**(10): p. 1399-413.
143. Dassarma, S., et al., *Genomic perspective on the photobiology of Halobacterium species NRC-1, a phototrophic, phototactic, and UV-tolerant haloarchaeon.* Photosynth Res, 2001. **70**(1): p. 3-17.
144. Carbonneau, M.A., et al., *The action of free radicals on Deinococcus radiodurans carotenoids.* Arch Biochem Biophys, 1989. **275**(1): p. 244-51.
145. Cannio, R., et al., *Oxygen: friend or foe? Archaeal superoxide dismutases in the protection of intra- and extracellular oxidative stress.* Front Biosci, 2000. **5**: p. D768-79.
146. Kish, A., *Unpublished*. 2008.
147. Spudich, E.N. and J.L. Spudich, *Control of transmembrane ion fluxes to select halorhodopsin-deficient and other energy-transduction mutants of Halobacterium halobium.* Proc Natl Acad Sci U S A, 1982. **79**(14): p. 4308-12.
148. Baliga, N.S., et al., *Coordinate regulation of energy transduction modules in Halobacterium sp. analyzed by a global systems approach.* Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14913-8.
149. Ideker, T., et al., *Testing for differentially-expressed genes by maximum-likelihood analysis of microarray data.* J Comput Biol, 2000. **7**(6): p. 805-17.
150. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.
151. Shannon, P.T., et al., *The Gaggle: an open-source software system for integrating bioinformatics software and data sources.* BMC Bioinformatics, 2006. **7**: p. 176.
152. Storz, G. and J.A. Imlay, *Oxidative stress.* Curr Opin Microbiol, 1999. **2**(2): p. 188-94.
153. Brioukhanov, A.L., A.I. Netrusov, and R.I. Eggen, *The catalase and superoxide dismutase genes are transcriptionally up-regulated upon oxidative stress in the strictly anaerobic archaeon Methanosarcina barkeri.* Microbiology, 2006. **152**(Pt 6): p. 1671-7.
154. Zheng, M., et al., *DNA microarray-mediated transcriptional profiling of the Escherichia coli response to hydrogen peroxide.* J Bacteriol, 2001. **183**(15): p. 4562-70.
155. Shahmohammadi, H.R., et al., *Protective roles of bacterioruberin and intracellular KCl in the resistance of Halobacterium salinarium against DNA-damaging agents.* J Radiat Res (Tokyo), 1998. **39**(4): p. 251-62.
156. Spudich, E.N. and J.L. Spudich, *The photochemical reactions of sensory rhodopsin I are altered by its transducer.* J Biol Chem, 1993. **268**(22): p. 16095-7.
157. Zhu, J., et al., *Effects of substitutions D73E, D73N, D103N and V106M on signaling and pH titration of sensory rhodopsin II.* Photochem Photobiol, 1997. **66**(6): p. 788-91.
158. Holmgren, A., *Thioredoxin and glutaredoxin systems.* J Biol Chem, 1989. **264**(24): p. 13963-6.
159. Russel, M., P. Model, and A. Holmgren, *Thioredoxin or glutaredoxin in Escherichia coli is essential for sulfate reduction but not for deoxyribonucleotide synthesis.* J Bacteriol, 1990. **172**(4): p. 1923-9.

160. Ritz, D., et al., *Thioredoxin 2 is involved in the oxidative stress response in Escherichia coli.* J Biol Chem, 2000. **275**(4): p. 2505-12.
161. Takemoto, T., Q.M. Zhang, and S. Yonei, *Different mechanisms of thioredoxin in its reduced and oxidized forms in defense against hydrogen peroxide in Escherichia coli.* Free Radic Biol Med, 1998. **24**(4): p. 556-62.
162. Liu, Y., et al., *Transcriptome dynamics of Deinococcus radiodurans recovering from ionizing radiation.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 4191-6.
163. Robey-Bond, S.M., et al., *Clostridium acetobutylicum 8-oxoguanine DNA glycosylase (Ogg) differs from eukaryotic Oggs with respect to opposite base discrimination.* Biochemistry, 2008. **47**(29): p. 7626-36.
164. Gogos, A. and N.D. Clarke, *Characterization of an 8-oxoguanine DNA glycosylase from Methanococcus jannaschii.* J Biol Chem, 1999. **274**(43): p. 30447-50.
165. Chung, J.H., et al., *Repair activities of 8-oxoguanine DNA glycosylase from Archaeoglobus fulgidus, a hyperthermophilic archaeon.* Mutat Res, 2001. **486**(2): p. 99-111.
166. van der Kemp, P.A., et al., *Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine.* Proc Natl Acad Sci U S A, 1996. **93**(11): p. 5197-202.
167. Koulis, A., et al., *Uracil-DNA glycosylase activities in hyperthermophilic micro-organisms.* FEMS Microbiol Lett, 1996. **143**(2-3): p. 267-71.
168. Begley, T.J. and R.P. Cunningham, *Methanobacterium thermoformicicum thymine DNA mismatch glycosylase: conversion of an N-glycosylase to an AP lyase.* Protein Eng, 1999. **12**(4): p. 333-40.
169. Yang, H., et al., *Characterization of a thermostable DNA glycosylase specific for U/G and T/G mismatches from the hyperthermophilic archaeon Pyrobaculum aerophilum.* J Bacteriol, 2000. **182**(5): p. 1272-9.
170. Thomas, D.C., Asad Umar, and Thomas A. Kunkel, *Measurement of Heteroduplex Repair in Human Cell Extracts.* Methods, 1995. **7**: p. 187-197.
171. Su, S.S. and P. Modrich, *Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs.* Proc Natl Acad Sci U S A, 1986. **83**(14): p. 5057-61.
172. Su, S.S., et al., *Mispair specificity of methyl-directed DNA mismatch correction in vitro.* J Biol Chem, 1988. **263**(14): p. 6829-35.
173. Iaccarino, I., et al., *MSH6, a Saccharomyces cerevisiae protein that binds to mismatches as a heterodimer with MSH2.* Curr Biol, 1996. **6**(4): p. 484-6.
174. Invitrogen (Carlsbad, C. *pET100/D/lacZ map*. [cited; Available from: https://commerce.invitrogen.com/index.cfm?fuseaction=iProtocol.unitSectionTree&treeNodeId=29691752BF5DC5A35D85F9C3C6CEB6C6.
175. Constantinesco, F., et al., *A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea.* Nucleic Acids Res, 2004. **32**(4): p. 1439-47.
176. Shimono, K., et al., *Functional expression of pharaonis phoborhodopsin in Escherichia coli.* FEBS Lett, 1997. **420**(1): p. 54-6.

177. Hohenfeld, I.P., A.A. Wegener, and M. Engelhard, *Purification of histidine tagged bacteriorhodopsin, pharaonis halorhodopsin and pharaonis sensory rhodopsin II functionally expressed in Escherichia coli.* FEBS Lett, 1999. **442**(2-3): p. 198-202.
178. Kennedy, S.P., et al., *Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence.* Genome Res, 2001. **11**(10): p. 1641-50.
179. Thomas, D.C., J.D. Roberts, and T.A. Kunkel, *Heteroduplex repair in extracts of human HeLa cells.* J Biol Chem, 1991. **266**(6): p. 3744-51.
180. Wickner, W., et al., *RNA synthesis initiates in vitro conversion of M13 DNA to its replicative form.* Proc Natl Acad Sci U S A, 1972. **69**(4): p. 965-9.
181. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.
182. Holmes, P.K., I.E. Dundas, and H.O. Halvorson, *Halophilic enzymes in cell-free extracts of Halobacterium salinarium.* J Bacteriol, 1965. **90**(4): p. 1159-60.