DNA Mismatch Repair Proteins in *Tetrahymena thermophila*

by

Kate Kudynska
B.Sc., University of Victoria, 2004

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

University of Victoria

Supervisory Committee

DNA Mismatch Repair Proteins in *Tetrahymena thermophila*

by

Kate Kudynska
B.S., University of Victoria, 2004

Dr. Claire G.Cupples (Department of Biochemistry and Microbiology)
Supervisor

Dr. Francis E.Nano (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Perry Howard (Department of Biochemistry and Microbiology Department of Biology)
Outside Member

Dr. Patrick von Aderkas (Department of Biology)
External Examiner

Dr. Claire G.Cupples (Department of Biochemistry and Microbiology)
Supervisor

Dr. Francis E.Nano (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Perry Howard (Department of Biochemistry and Microbiology, Department of Biology)
Outside Member

Dr. Patrick von Aderkas (Department of Biology)
External Examiner

ABSTRACT

DNA mismatch repair (MMR) is an essential part of genomic stability, guarding the integrity of the genome in virtually all cells. MMR corrects mismatched bases in DNA and is one of several DNA repair pathways conserved from bacteria to humans. In *Escherichia coli,* MutS and MutL are the key proteins in MMR. In prokaryotes, the MMR proteins function as homodimers whereas in eukaryotes the MMR proteins come together as heterodimers. In the absence of MMR there is a great increase in mutation frequency and a higher susceptibility to cancer in mammals. Previous studies in our laboratory have identified five MutS homologs and two MutL homologs in the single-celled ciliated protozoan, *Tetrahymena thermophila.* MMR repair has been extensively studied in *E. coli* but less is known in eukaryotes.

T. thermophila's biology and the recent sequencing of its genome make it an attractive eukaryotic research model. In this study, poly-histidine tagged MMR genes from *T. thermophila* and human thymine DNA glycosylase (TDG) were cloned into two different types of *T. thermophila* expression plasmid. The integrated homologously recombinational *T. thermophila* vector approach and the espisomal rDNA vector approach which utilized Gateway® technology a cloning method based on the bacteriophage lambda site-specific recombination system. The homologous vector approach relies on mutant strains of *T. thermophila* harboring a negatively selectable allele of a β-tubulin gene producing sensitivity to paclitaxol. Upon knocking out the mutant β-tubulin with the gene of interest the resistance of the *Tetrahymena* strain is then resorted and selected for. The various *T. thermophila* expression plasmids and DNA transformations techniques to follow such as biolistic bombardment and conjugated electroporation were carried out in order to optimize the technical aspect of working with *T. thermophila.* and finally lead to the purification of histidine tagged *T. thermophila* MMR proteins for antibody production and further studies. This study will lead to insight into the inner workings of DNA mismatch repair and technical aspect of working with the organism and will permit functional and structural studies of the MMR homolog proteins in *T. thermophila.*

TABLE OF CONTENTS

List of Figures

List of Tables

LIST OF ABBREVIATIONS

°C	degrees Celsius
A	absorbance at x nm
aa	amino acid
BSA	bovine serum albumin
BTU	β-tubulin
bp	base pair
CFU	colony forming unit
dHO	deionized water
DMSO	dimethylsuphoxide
DNA	deoxyribonucleic acid
EDTA	(ethylene diamine) tetraacetic acid
EtBr	ethidium bromide
EtOH	ethanol
g	gram
h	hour
IgA	immunoglobulin A
IgG/IgM	immunoglobulin G/ immunoglobulinM
IPTG	isopropyl-β-D-thiogalactoside
kDa	kilodalton
kV	kilovolt
l	liter

LB	Luria-Bertani media
M	molar
mAb	monoclonal antibody
mg	milligram
min	minute
mL	milliliter
mM	millimolar
MMR	mismatch repair
M W	molecular weight
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PBS	phosphate buffered saline
SDS	sodium dodecylsulphate
TDG	thymine-DNA glycosylase
UVic	University of Victoria
V	Volt
μg	microgram
μl	microliter
μM	micromolar

Acknowledgements

There are many people whom I would like to thank for making my graduate studies possible. My deepest thanks goes out to my supervisor Dr. Claire Cupples. I would like to thank Dr. Claire Cupples for her ongoing support and continual encouragement, especially when I need that extra push or advice. I would like to thank Dr. Cupples for sharing her enthusiasm and sincere interest in science with me which has helped me deepen my appreciation for research and ignited my scientific curiosity. Thank you for giving me a chance to prove my abilities and for giving me an opportunity to explore my interest in research.

I want to thank my committee for their suggestions and guidance. Thank you for taking the time for the committee meetings and your advice.

And finally I must thank my family. My parents and brother, even though living on the other side of the continent were joined with me in sprit every step of the way. I want to thank my children Michelle and Andrew despite their young age for being very understating. Also, I would like to thank my friends and lab mates for exhilarating scientific conversations and camaraderie. Thank you to all my friends for lending me their shoulders to lean on without which I would not be where I am today.

"We totally missed the possible role of...repair. I came later to realize that DNA is so precious that probably many distinct repair mechanisms would exist. Nowadays one could hardly discuss mutation without considering repair at the same time."

Francis Crick, ("The Double Helix: A Personal View." Nature,1974.)

Chapter 1: Literature Review

1.1 DNA REPAIR

1.1.1 Maintenance of Genomic Integrity

Maintaining the integrity of genetic information lies at the heart of DNA repair. Damage to DNA can have dire consequences; this damage may permanently affect or eliminate the function of a gene and even make this change heritable. DNA repair mechanisms are present in all organisms examined to date and many repair enzymes are conserved from archaea to eukaryotes (Jaroudi *et al.,* 2006). In the absence of DNA repair, apoptotic pathways may fail to be activated and lead to cell survival, or persistent DNA damage may lead to loss of genetic material, genomic and chromosomal instability, mutagenesis, and ultimately malignant transformation or cell death (Reese *et al.,* 2003). The number, diversity, and complexity of DNA repair mechanisms reflect the wide range of DNA threats present (Gorbunova *et al.,* 2007).

1.1.2 Types of DNA Damage and the Corresponding DNA Repair System

DNA damage is caused by a variety of agents and may be divided into two major categories, endogenous and environmental (Loft *et al.,* 1996; Skinner *et al.,* 2005). The endogenous category includes spontaneous alteration mainly due to hydrolytic and oxidative reactions during DNA metabolism. Environmental damages are due to physical or chemical agents such as UV light exposure, ionizing radiation and chemical exposure (Figure 1). The insults to DNA generate reactive oxygen species (ROS) including free radicals (e.g., superoxide and hydroxyl radicals), nonradical oxygen species (e.g., hydrogen peroxide and peroxynitrite) and reactive lipids and carbohydrates (e.g., ketoaldehydes, hydroxynonenal) (Gracy et al., 1999). The cellular response to DNA damage may involve activation of cell cycle checkpoints, execution of DNA repair, or when the damage is severe, initiation of apoptosis (Peltomaki, 2001).

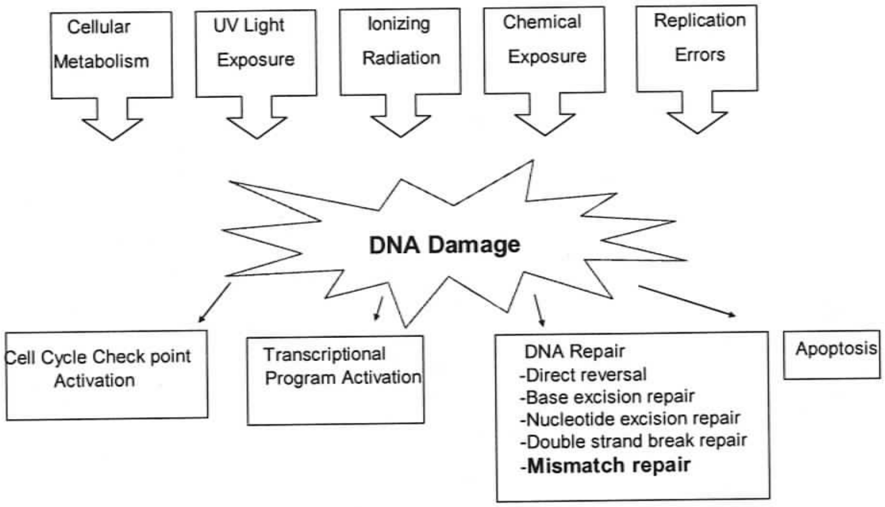


Figure 1: DNA Damage Response (Adapted from rndsystems.com)

The damaged DNA may present itself in a variety of alterations in the normal DNA structure including: single- and double-strand breaks, chemically modified bases, abasic sites, bulky adduct, inter- and intra-strand cross-links, and base-paring mismatches (Friedberg *et al,* 2003). These alterations can have many direct and indirect effects on cells and organism including: mutations, genetic recombination, chromosomal aberrations, tumorigenesis and cell death (Van Houten et al., 2002). Cells have evolved a multitude of repair pathways and are equipped to repair just about any type of DNA abnormality (Table 2) (Friedberg, 2003).

Table 1: Biological responses to DNA damage

DNA Damage	**Biological Response**
Reversal of base damage Excision of damaged, mispaired, or incorrect bases	Base excision repair (BER) Nucleotide excision repair (NER) Transcription-coupled nucleotide excision repair (TC-NER) Alternative excision repair (AER) Mismatch repair (MMR)
Strand break repair	Single-strand break repair (SSBR) Double-strand break repair (DSBR)
Tolerance of base damage	Translesion DNA synthesis (TLS) Postreplicative gap filling Replication fork progression
Cell cycle checkpoint activation	Apoptosis

1.1.3 DNA Mismatch Repair

DNA mismatch repair (MMR) guards the integrity of the genome in virtually all cells by targeting mismatched bases and insertion-deletion loops (IDLs) in duplex DNA for correction. MMR is one of several DNA repair pathways conserved from bacteria to humans (Iyer *et al.,* 2006). Mispaired bases may arise in variety of ways through replication errors, heteroduplex formation during homologous recombination, or as a result of DNA damage (Kunkel *et al,* 2005). During replication, DNA polymerase containing a proofreading function confers an error rate of ~10^{-7} per bp per replication; MMR elevates this fidelity further by 50-1000 fold (Schofield *et al.,* 2003). An array of base pair anomalies resulting from DNA damage are subject to processing by mismatch repair. These include base pairs containing O^6-methylguanine, 8-oxoguanine, carcinogen adducts, UV photo products, and cisplatin adducts (Iyer et al., 2006).

Elevated rates of spontaneous mutations are a hallmark of defects in genes associated with postreplicative mismatch repair and led to their original designation as mutator (*mut*) genes in bacteria. Many of these mutator genes have been found to be conserved throughout evolution (Acharya *et al*., 1996). Furthermore, there is frequent microsatellite instability at mono- and dinucleotide repeats in eukaryotic organisms deficient in MMR (Peltomaki, 2003).

MMR plays an important role in preserving the overall fidelity of replication, and various other functions have been assigned to the MMR system. Among them MMR proteins have been shown to be involved in diverse genetic pathways by interactions with other systems. MMR is found to police homologous recombination events and acts as a barrier to genetic exchange between species (Schofield *et al.* 2003). Furthermore, MMR is involved in promotion of meiotic cross over, in check-point and apoptotic responses to several classes of DNA damage (Iyer *et al,* 2006), and even in diversification of immunoglobulins (Figure 2) (Jun *et al,* 2006).

In order to be successful, MMR like other DNA repair systems has to complete a two step process as follows: 1) the DNA lesion needs to be detected and 2) it needs to be repaired. MMR is unique. Unlike other repair systems it mainly targets nucleotides which are normal rather than damaged, and the removal of the mismatched nucleotide is specific to the newly synthesized daughter strand (Bjornson *et al.,* 2003).

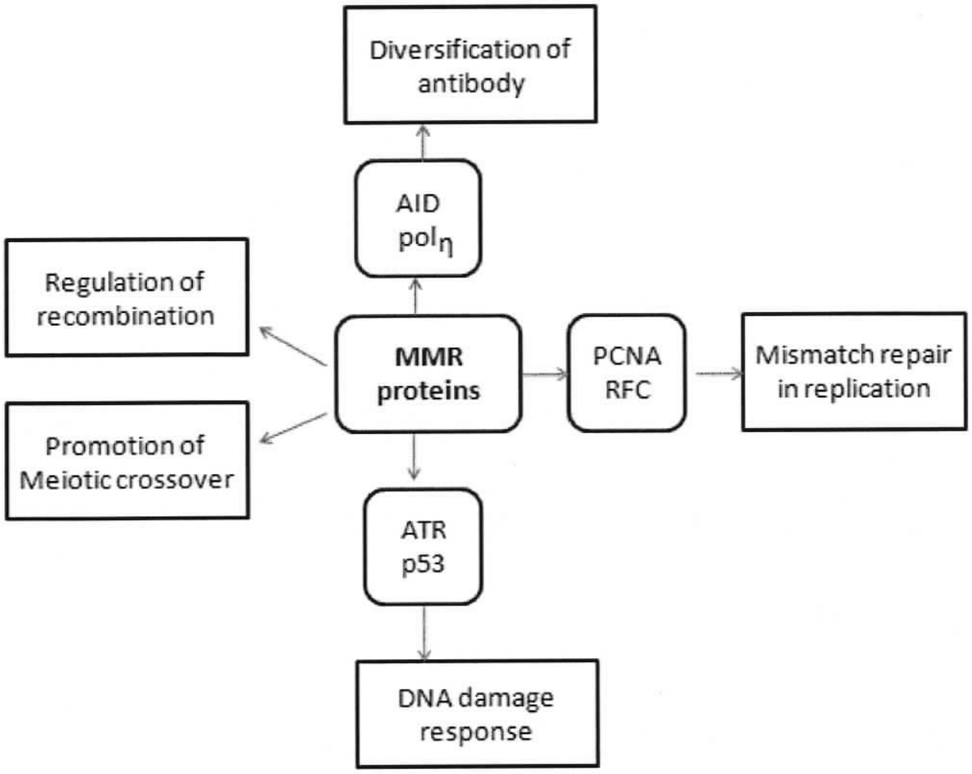


Figure 2: Various Functions of Mismatch Repair (MMR) Proteins. Proliferating cell nuclear antigen (PCNA) and replication factor C (RFC) work with MMR proteins during mismatch repair. DNA damage responses are activated through interactions with ATR and p53. MMR proteins introduce mutations in immunoglobulin genes in conjunction with activation-induced cytidine deaminase (AID) and DNA polymerase η, resulting in antibody diversification. In addition, MMR proteins regulate recombination and promote meiotic crossover (Adapted from Jun *et al.,* 2006).

In this review of the DNA mismatch repair system, the main components of MMR system in *E. coli* will be discussed, followed by an overview of eukaryotic MMR system.

1.1.4. Mismatch Repair Mechanism in *E. coli*

The mechanism of mismatch repair has been most extensively studied in *E. coli* (Iyer *et al.,* 2006) and the tasks of MMR machinery were first elucidated in *E. coli*. A vast amount of information about MMR comes from reconstitution in *vitro* of individual purified components in *E. coli* (Lahue *et al.,* 1989; Modrich 2006).

Reconstitution studies have indicated that a DNA lesion is recognized and bound by a MutS homodimer (Modrich *et al.,* 1996). MutS protein in the presence of ATP and a mismatch recruits MutL, and together they activate a weak, single stranded endonuclease MutH, whose activity is stimulated by association with MutL, MutS and a mismatch (Modrich *et al.,* 1996). The strand specificity is provided by the hemimethylated GATC site, which serves as a recognition sequence for Dam adenine methylase. In *E. coli*, MMR takes advantage of the fact that in the newly synthesized DNA strand adenines remain unmethylated for several minutes after the passage of the replication fork (Wagner *et al.,* 1979; Lu *et al.,* 1983). The hemi-methylated sequence is then cleaved on the unmethylated (new) strand by MutH. The GATC can reside on either side of the mismatch at distances up to 2kb away (Grilley et al., 1993). Following this incision by MutH, SSB and UvrD helicase are loaded at the nick via MutL protein-protein interactions and the UvrD helices unwinds the DNA in the direction of the mismatch Thus, once the mismatch is detected, the displaced DNA strand is degraded by one of the single-strand exonucleases: 3'→ 5' exonucleases ExoI or ExoX, or 5'→ 3' exonucleases RecJ or ExoVII (Burdett *et al.,* 2001). The repair synthesis is subsequently completed by DNA polIII and finally by re-sealing of the DNA backbone by DNA ligase (Jiricny 2006). Even though a lot is known about *E. coli* MMR, much remains to be understood.

Methyl-directed DNA Mismatch Repair

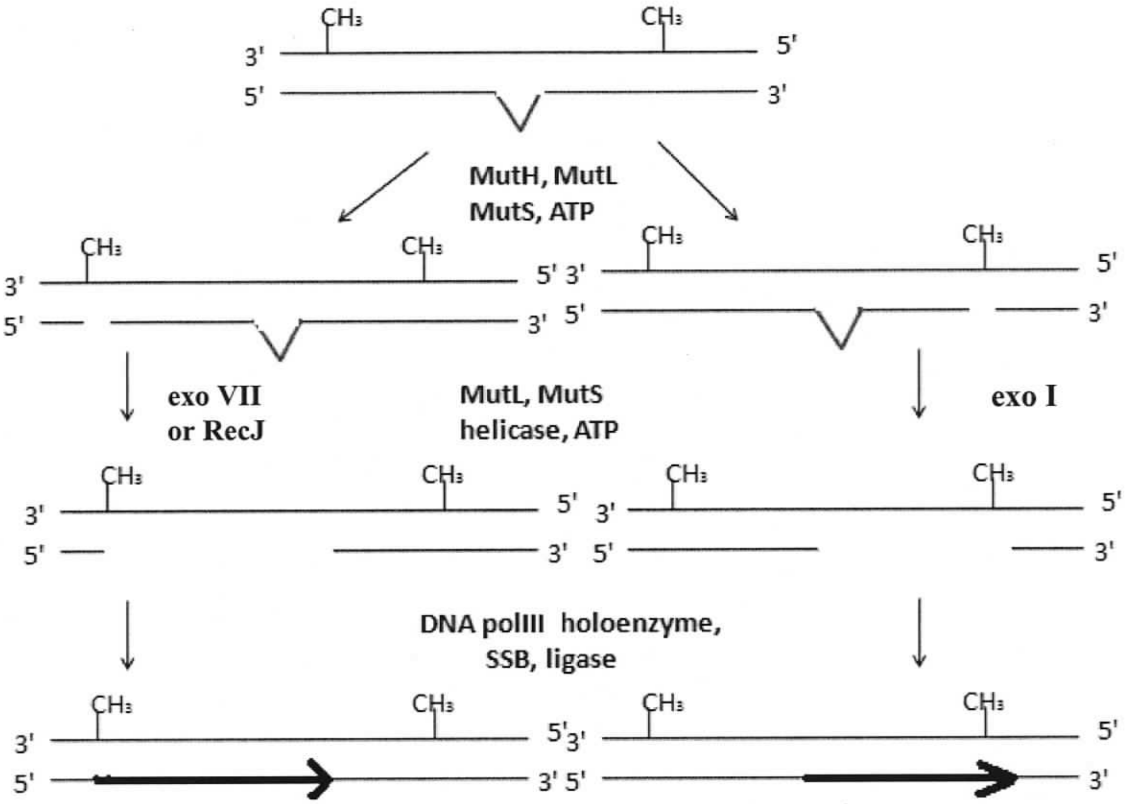


Figure 3: A Model of Methyl-directed DNA MMR in *E. coli*. MMR is initiated by mismatch recognition and strand discrimination. Following excision of the newly synthesized strand containing the mismatched base or insertion/deletion loop (IDL), strand resynthesis is performed to restore the fidelity of the DNA.

Table 2: Summary of *E. coli* proteins required for DNA-mismatch repair

Initiation	**MutS**	Binds to mismatches and small IDLs.
	MutL	Molecular matchmaker. Interacts with MutS in a mismatch-specific manner to activate MutH cleavage activity.
	MutH	Nicks unmethylated strand at a hemimethylated GATC site, initiating repair.
	β-clamp	Processivity clamp for Pol III. Interacts with MutS in vitro. May recruit MutS to the replication fork.
Excision	Helicase II (UvrD)	Delivered to the nick by MutS and MutL. Unwinds DNA prior to its excision.
	RecJ ExoVII	Required for 3'-5' excision between nick and mismatch.
	ExoI ExoX	Required for 5'-3' excision between nick and mismatch.
Resynthesis	Pol III	DNA polymerase required for repair resythesis.
	SSB	Single stranded DNA binding protein, protects DNA from degradation.
	Ligase	Seals nicks.

(Adapted from Schofield *et al,* 2003)

1.1.5 Eukaryotic Mismatch Repair

Due to the evolutionary conservation of MMR between organisms, much of our understating about the eukaryotic MMR comes from the initial studies performed in *E. coli* (Kunkel *et al.,* 2005). Even though MMR has been best characterized in *E. coli*, much information is now available on yeast, mouse, plant and human systems, among others. Eukaryotic cells contain multiple homologs of MutS (MSH) and MutL (MLH/PMS) proteins which function as heterodimers as opposed to homodimers in bacteria. The initial steps of MMR in prokaryotes and eukaryotes appear to be the same (Modrich 2006). A mismatch in eukaryotic DNA can be recognized by either of the two heterodimeric complexes of MutS related proteins – MSH2/MSH6 (MutSα) or MSH2/MSH3 (MutSβ) depending on the kind of a mismatch present (Table 3) (Acharya et al., 1996).

Acharya and coworkers (1996) have demonstrated the physical and functional interactions between the human MutS homologues hMSH2, hMSH3, and hMSH6. Immunoprecipitation and cross-linking experiments have shown that hMSH2 forms a stable complex with both hMSH3 and hMSH6 *in vitro* also hMSH2 is capable of associating with itself (homodimer) *in vitro* (Acharya *et al.*, 1996).

The other main player in MMR is the MutL protein, which is known as a "matchmaker". In eukaryotes, the family of MutL homologs consist of MLH1 and either PMS1, PMS2 or MLH3 (Table 3) (Kosinski *et al.,* 2005). Furthermore, MutL is a member of the GHKL superfamily of ATPases, which includes gyrase, a type II topoisomerase, Hsp90, histidine kinase, and MutL (Table 3) (Dutta *et al.,* 2000).

Table 3: Eukaryotic MMR Components and their Function.

Complex	**Components**	**Function**
MutSα	MSH2/MSH6	Recognition of base-base mismatches and small IDLs
MutSβ	MSH2/MSH3	Recognition of IDLs
MutLα	MLH1/PMS2	Forms a ternary complex with mismatch DNA and MutSα; endonucleic activity, increases discrimination between heteroduplexes and homoduplexes; also functions in meiotic recombination
MutLβ	MLH1/PMS1	Unknown
MutLγ	MLH1/MLH3	Primary function in meiotic recombination; backup for MutLα in the repair of base-base mismatches and small IDLs

(Adapted from Jiricny 2006; Kadyrov *et al,* 2006)

A strand-directed endonuclease that nicks the daughter strand, like MutH in *E. coli*, is missing in all eukaryotes and in many bacteria. In 2006, Paul Modrich's group reported the discovery of an intrinsic endonuclease activity in human MutLα (MLH1-PMS2 heterodimer) which is dependent on the integrity of a PMS2 DQHA(X) 2E(X) 4E motif (Kadyrov et al., 2006). This breakthrough provides a satisfactory answer to the longstanding puzzle of a missing endonuclease activity in eukaryotic mismatch repair (Yang, 2006). The endonucleolytic activity of MutLα is activated by MutSα bound to a mismatch and the sliding clamp PCNA loaded by RFC at a free 3′ end, and it cleaves the daughter strand on either side of the mismatched base. The authors further showed that the endonucleolytic activity of MutLα requires Mn^{2+} and ATP (Kadyrov *et al.,* 2006). Thus, the endonuclease activity of MutLα and the 5′ → 3′ exonuclease of ExoI together can remove a mismatch from a 5′ or 3′ strand break (Yang, 2006).

Table 4: Comparison of Bacterial and Eukaryotic MMR Components.

E. coli	**Human**	**Function**
(MutS)2	hMutSα(MSH2-MSH6) hMutSβ(MSH2-MSH3)	DNA mismatch/damage recognition
(MutL)2	hMutLα (MLH1-PMS2) hMutLβ(MLH1-PMS1) hMutLγ (MLH1-MLH3)	Molecular matchmaker/chaperone endonucleic activity in eukaryotes
MutH	MutLα	endonuclease
UvrD	?	Unwinding DNA helix
ExoI, ExoVII, ExoX, RecJ	ExoI	Removing mispaired base
Pol III holoenzyme	Pol d, PCNA	DNA re-synthesis
SSB	RPA	Protecting template DNA from degradation
DNA Ligase	?	Nick ligation

(Adapted from Guo-Min Li, 2003)

1.1.6 MutS Protein Structural and Biochemical Features

MutS is a mismatch recognition protein that binds to 7 of 8 possible mismatches; it does not bind to C:C. The specificity of repair is the following, from highest affinities to lowest: GT, AC, small loops > GG, AA >>> TC, TT, CC, large loops (Carraway *et al.,* 1993). High-resolution crystal structures have been solved for the mismatch binding protein MutS of *Escherichia coli* (Lamers *et al.,* 2000) and its *Thermus aquaticus* (Obmolova, *et al.,* 2000) homologue. The MutS structure can be visualized as a pair of praying hands with the thumbs folded inwards, and the DNA passes between the fingertips and the thumbs (Jiricny 2006). MutS protein is shaped as an oval disk with two openings, of diameters approximately 30 Å and 40 Å, with the DNA passing through the latter (Obmolova *et al.,* 2000). Each subunit consists of five distinct domains. The amino-terminal domain I, which forms the top segment of the thumb, contains the conserved amino-acid motif that is required for mismatch recognition. Domains II and III form the second and third thumb segments; domain IV forms the fingers; and the carboxyl-terminal domain V the heels of the palms. This last domain contains the protein's ATP-binding site, which consists of the highly conserved Walker-type motifs (Jircny 2000).

MutS binds a mismatch in duplex DNA, in the absence of DNA the finger domains (domainIV) are unstructured and open, and the ATP-binding sites are dimerized. In the presence of a mismatch, the ADP-bound form of MutS wraps around the DNA like a pair of praying hands (praying hands of fidelity), this is anchored by a Phe-X-Glu wedge (thumb of one of the hands) into the minor groove of the duplex(Jiricny, 2006). Once bound to the DNA, the ADP/ATP exchange brings about a conformational change that releases the Phe-X-Glu thumb from the mismatch site allowing the clamp to freely translocate along the DNA in either direction. Although no eukaryotic MutSα structures is solved, biochemical experiments indicate that these heterodimers function analogously to the bacterial proteins (Lamers *et al.,* 2000).

1.1.7 MutL Biochemical and Structural Features

MutL has been assigned the function of a primary coordinator and regulator of MMR (Yang, 2006). Structurally, MutL can be divided into a highly conserved N-terminal ATPase domain (NTD), which dimerizes upon binding ATP, and a less conserved C-terminal domain (CTD) involved in homo- and heterodimerization, as well as a flexible, poorly conserved linker connecting the two domains (Guarne *et al.,* 2004) . Based on prokaryotic crystal structures of the N- and C-terminal fragments models of the entire MutL have been constructed (Guarne *et al,* 2004; Kosinski *et al,* 2005). Both the N- and C-terminal domains of *E. coli* and yeast MutL (PMS2-MLH1) have been shown to interact with DNA (Yang, 2006). The N-terminal ATPase region is highly conserved among all MutL homologs and shares four sequence motifs with other GHKL (for Gyrase, Hsp90, histidine kinase and MutL) ATPase/kinase superfamily, this unique structural ATP-binding motif is known as the Bergerat fold (Dutta *et al.,* 2000). Unlike, the N-terminal domain the C-terminal region shares little sequence similarity among MutL homologs but is essential for homo- or hetero-dimerization of MutL and its homologs (Wu *et al.,* 2003).

MutL activates *E. coli* UvrD helicase to unwind duplex DNA with a free end, the activation of UvrD on a nicked circular DNA substrate requires a MutS–MutL-mismatch complex (Yamaguchi *et al.,* 1998). Recently, Modrich group has shown the endonucleolytic activity of the human MutL complex.

Both MutS and MutL are ATPases, and mutations that impair either ATP binding or ATP hydrolysis by MutS or MutL abolish the mismatch repair process (Baitinger *et al,* 2003). Conformational changes of MutL have been suggested to enable the MutS–MutL-mismatch DNA complex to recruit different downstream effectors as outlined in Figure 2 (Junop et al, 2003).

1.1.8 Defects in Mismatch DNA Repair and Drug Resistance

Genetic inactivation of the mismatch repair system elevates spontaneous mutability 50-1000 fold (Schofield *et al,* 2003). Mismatch repair defects lead to highly elevated rates of base substitution and frameshift mutations, permit illegitimate recombination between quasi-homologous sequences (Harfe *et al,* 2000), and render mammalian cells resistant to the cytotoxic effects of several classes of DNA damaging agents (Stojic *et al,* 2004).

In the absence of MMR, cells are resistant (tolerant) to methylating agents and chemicals, such as S_N1-type methylating agents (ex. MMNG, TMZ), antimetabolites (6-thioguanine), intrastrand crosslinking agents (cisplatin a powerful anticancer chemotherapeutic agent) and intrastrand crosslinking agents (MMC, CCNU) (Stojic *et al.,* 2004). For example, there is now evidence to suggest that MMR acts as an upstream component of a cisplatin-induced apoptotic pathway mediated by the tumor suppressor p73 (Shimodaira et al., 2003). The "tolerant" cells have increased their survival despite an accumulation of mutations. These mutations may be point mutations, insertions, or deletions in simple repeat sequences throughout the genome (Shimodaira et al., 2003).

One of the hallmarks of mismatch repair deficiency is microsatellite instability (MSI) (Boyer *et al.,* 1995). Microsatellite instability is defined as alterations in the length of short repetitive sequences, such as $[A]_n$ or $[CA]_n$, where the changes in the DNA can be detected by the appearance of new bands after PCR amplification of microsatellite DNA from MMR-deficient cells. The reason the new bands can be detected is due to the fact that during DNA synthesis, the primer and the template strands in a microsatellite can dissociate and re-anneal incorrectly and these heterogeneities are known as insertion and /deletion loops (IDLs) (Jiricny, 2006).

The connection of mutator genes to frequently occurring human cancers was first recognized when inherited mutations in the hMSH2 gene were found to underlie a significant proportion (>90%) of hereditary nonpolyposis colorectal cancers (HNPCC) (Fishel *et al.,* 1993). Since 1993, HNPCC has been confirmed to be associated with germline mutations in several DNA MMR genes, including hMSH2, hMLH1, hMSH6, hPMS2 with mutations in hMSH2 and MLH1 being most common (Wang et al., 2007). Moreover, a significant proportion of human cancers of the endometrium and other organs also display MSI (Jiricny., 2006).

It has recently been reported that cancer stem cells (CSCs) having MMR deficiencies are also insensitive ("tolerant") to the treatments with cytotoxic agents and as a result increase the risk of relapse and metastasis. Thus, defective MMR status in a damaged cell directly contributes to the resistant phenotype against the cytotoxic effects of chemotherapeutic drugs allowing CSCs to survive chemotherapy (Vaish., 2007).

The activation of DNA MMR system can inhibit the survival mechanisms by raising the sensitivity of such cells to therapeutic drugs and consequently reduce its tumor potential by arresting their growth. Owning to mass effect, progeny cancer cells lead to renewed symptoms in the patients and therefore, combinational therapy for the eradication of cancer stem cells and differentiated cells would be most effective. The importance of MMR in designing therapeutic strategies specifically targeting the tumor cells is being explored and pursued in clinical trials (Vaish., 2007).

MMR is a multifaceted system and the loss of MMR genes directly or indirectly alters signal pathways, loss of strictly regulated expression of cytokine receptors, transcription factors and cell cycle regulation (Reese *et al.,* 2003).

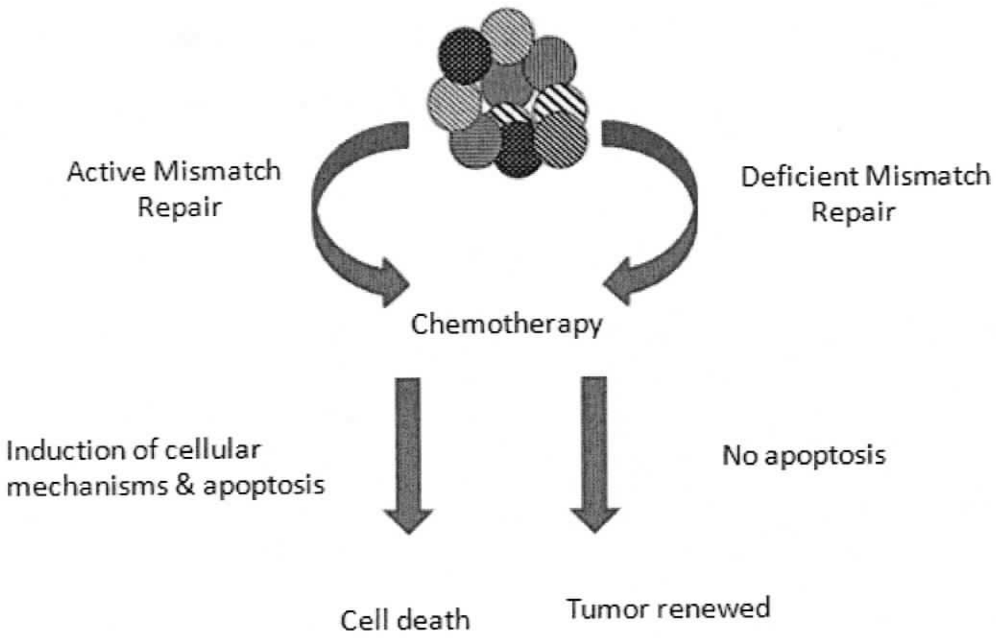


Figure 4: MMR Deficient Cancer cells and Chemotherapy. Primary tumors containing heterogeneous population of tumor cells contain a small percentage of cancer stem cells (CSCs) (solid gray colored oval). MMR deficient cells make tumor cells insensitive to chemotherapy and the cancer cells relapse (Adapted from Vaish., 2007).

1.2 TETRAHYMENA THERMOPHILA

1.2.1 *Tetrahymena thermophila*

Tetrahymena thermophila is a single-celled ciliate eukaryote. It has been used as a microbial animal model for more than 75 years -- ever since Nobel Laureate Andre Lwoff in 1923 succeeded in growing this cell in pure culture. This ciliate is a member of phylum Protozoa, subpylum Ciliopora, which includes the genera *Paramecium, Oxytricha,* and *Ichthyophthirius.* This freshwater non-parasitic ciliate is large in size measuring 40-50 µm along the anterior-posterior axis and commonly inhabits streams, lakes and ponds (Fjerdingstad., 2007). The ciliates are one of three major evolutionary lineages that make up the alveolates. The other two lineages are dinoflagellates and the exclusively parasitic apicomplexa, which includes the *Plasmodium* species, the causative agent of malaria. Although experimental tools are improving for the apicomplexa, they can still be challenging to work with, and in some situations *T. thermophila* can serve as a useful "distant cousin" model for this particular group (Eisen *et al.,* 2006).

The *Tetrahymena* genome has recently been sequenced and the organism is emerging as a model for unicellular eukaryotic biology (Turkewitz *et al.,* 2002). In the last decades studies in *T. thermophila* have contributed to fundamental biological discoveries such as the first cell whose division was synchronized, discovery of catalytic RNA (Zaug and Cech., 1986), RNAi-guided genome rearrangement (Mochizuki et al., 2004b), telomeres and telomerase (Greider and Blackburn, 1985), the function of chromatin acetylation (Brownell *et al.,* 1996) and the first example of a noncanonical nuclear code (Weide *et al.,* 2007). Although *T. thermophila* is unicellular, it possesses many core processes conserved across a wide diversity of eukaryotes (including humans) that are not found in other single-celled model systems (e.g., the yeasts

Saccharomyces cerevisiae and *Schizosaccharomyces pombe*) (Eisen *et al.,* 2006) and thus this allows *T. thermophila* to be an active model organism for many studies with direct relevance to human health, including telomerase function, regulated secretion, ciliary motility, post-translational modification of tubulins, chromatin modification, phagocytosis, RNAi, chromosome condensation and segregation, signal transduction, apoptosis, mitotic checkpoints, immunoglobulin-like DNA rearrangement and DNA replication, amplification, and copy number control (White Paper., 2007).

In addition, the attractiveness of this single celled eukaryote stems from the fact that *Tetrahymena* cells grow rapidly to high density in a variety of inexpensive media conditions over a wide temperature range: 12°C to 41°C, with a generation time of approximately 2.5 hours at 30°C (Wheatley *et al.,* 1994). Furthermore, the sexual stage in its life cycle (conjugation) can be induced at will, with high efficiency and synchrony (Collins *et al.,* 2005).

1.2.2 Nuclear Dimorphism in *Tetrahymena thermophila*

Like other ciliates *T. thermophila* possesses two types of nucleus; this is know as nuclear dimorphism and it separates the germline and the soma functions into two functionally and structurally distinct nuclei within a single cell (Figure 5)(Karrer., 2000). *Tetrahymena* has a micronucleus (MIC) and a macronucleus (MAC); even though the nuclei are distinct they do contain closely related genomes (Yao *et al.,* 2005). The MIC is the transcriptionally silent germline nucleus and serves as the store of genetic information for the progeny produced by conjugation in the sexual stage of *T. thermophila* life cycle. The MIC nucleus is diploid and contains five pairs of chromosomes. The MAC is transcriptionally active polypoid somatic nuclei, derived from the MIC (Mochizuki *et al.,* 2004b).

A shotgun sequencing analysis of the MAC DNA revealed that *T. thermophila* MAC genome is 104 Mb in length and has approximately 225 MAC chromosomes that contain more than 27,000 protein coding genes. These MAC chromosomes are autonomously replicating species derived from the 5 MIC chromosomes by site –specific fragmentation. So far, about 15,000 genes match genes of other organisms. In addition the genome analysis elucidated that a large number of genes are based on gene duplication mechanisms, this being yet another illustration of the complexity of this single celled eukaryotic microorganism (Eisen *et al.,* 2006).

Nuclear Dimorphism in *Tetrahymena thermophila*

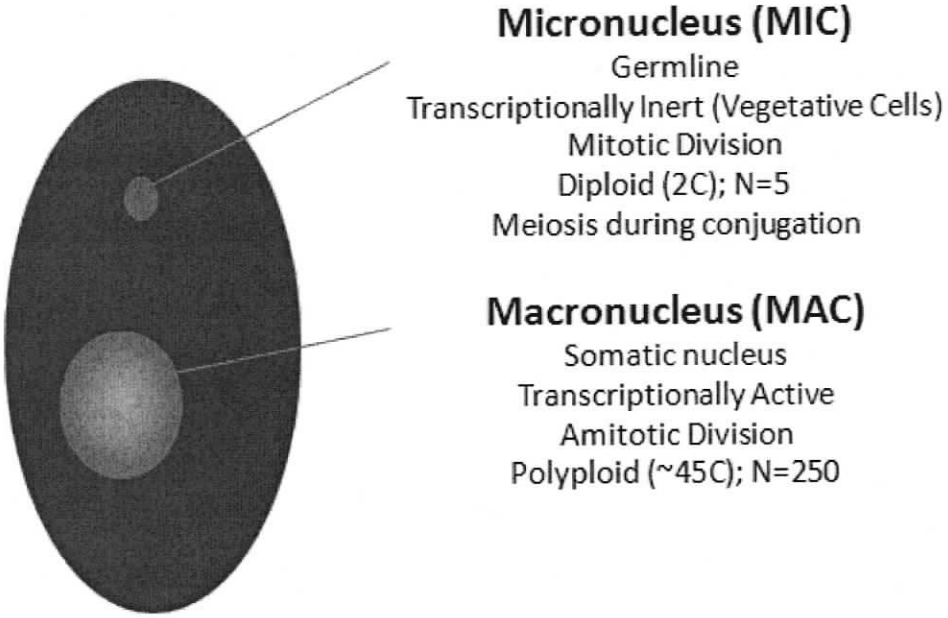


Figure 5: Nuclear Dimorphism in *Tetrahymena thermophila.*

1.2.3 The Life Cycle of *Tetrahymena thermophila*

1.2.3.1 Conjugation and Mating Type

In the absence of plentiful nutrient resources, *Tetrahymena* undergo sexual reproduction, pairing with cells of a complementary mating type. An individual cell expresses one of seven mating types, designated by the Roman numerals I-VII. Mating can take place between any of the two different mating types. The mating type is determined by a single locus called mat. The locus has been mapped but the molecular bases of the mating-type determination and mating type interactions in *Tetrahymena* are unknown. However, once the mating-type is determined in the cell, it is extremely stable in its asexual progeny (Simon et al., 1987). Thus, pure cultures for a single mating-type are easily maintained.

1.2.3.2 MAC differentiation/programmed / DNA rearrangement

Conjugation involves meiosis, fusion of haploid MIC gametes to produce a new zygotic nucleus, and differentiation of new MACs from mitotic copies of the zygotic nucleus. After formation of the MAC, cells reproduce asexually until the next sexual conjugation. During this asexual growth, all gene expression occurs in the MAC, which is thus considered the somatic nucleus (Eisen *et al.,* 2006). The mating process can take somewhere between 12 to 24 hours; an outline is found in Figure 6.

During MAC differentiation the number of copies of the MAC chromosomes is amplified. The rDNA chromosome, which encodes the 5.8S, 17S, and 26S rRNAs, is maintained at an average of 9,000 copies per MAC. Six other sub-chromosomal fragments that have been examined are each maintained at an average of 45 copies per MAC (Mochizuki et al., 2004b). During asexual reproduction, the MAC divides amitotically, with apparently random distribution of copies that behave as if acentromeric. In contrast, MIC chromosomes are metacentric and are distributed mitotically. Parental MAC DNA is not transmitted to sexual progeny, although it does have an epigenetic influence on post-zygotic MAC genome rearrangement, mediated by RNA interference (Eisen *et al.,* 2006).

1.2.4 *Tetrahymena* and Biotechnological Applications

Now this remarkable single celled organism has made the leap from a simple model organism in basic biology to a new role as potentially useful expression system in biotechnology (Turkewitz et al., 2002). For example, it is used for determining the protein nutritional value of human foods, the toxicity of compounds of chemical or biological origin (Sauvant *et al.,* 1999), and for monitoring water quality (Power *et al.,* 2006; Pinheiro et al., 2007). It has potential for industrial synthesis of pharmaceuticals. Its capacity for massive, synchronous regulated secretion confers the potential for industrial synthesis of useful proteins (Wheatley *et al.,* 1994). Since mosquitoes, snails and fish are parasitized by other *Tetrahymena* species and closely related genera; *T. thermophila* has the potential to develop tools for the biological control of major world-wide human diseases for which those organisms are intermediate vectors. Examples of these diseases are schistosomiasis, malaria, and yellow fever (Peterson et al., 2002). *Tetrahymena*-based biotech strategies to fight fish hatchery and pet store epidemics of the closely related ciliate Ichthyophthirius ("Ick") are currently being explored. It is an ideal candidate system for production of pharmaceuticals and vaccines (Geartig et al., 1999).

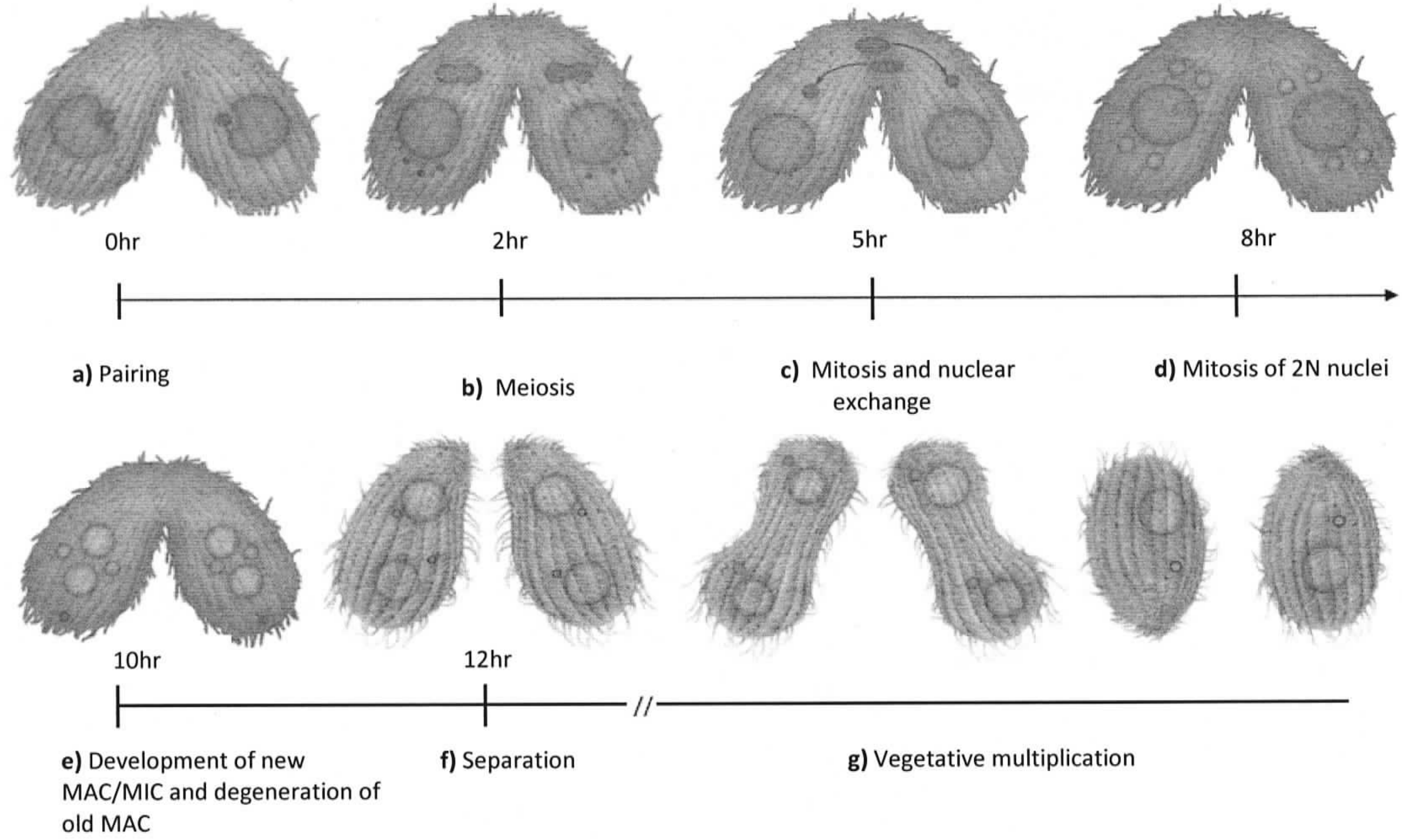


Figure 6: Starvation Induced Conjugation in *Tetrahymena thermophila*. (a) Pair formation of different mating types. (b)The micronucleus undergoes meiosis to generate haploid meiotic products. (c)Three of the meiotic products are destroyed, and the remaining one divides mitotically to generate two gamete pronuclei of identical genetic composition. Genetic exchange between conjugating partners is reciprocal and generates zygote nuclei of identical genotype in the two conjugants. (d) The zygote nucleus undergoes two mitotic divisions. (e) The zygotic nuclei differentiate into MICs and MACs, and the old MAC is destroyed. (f) The conjugants then separate, becoming exconjugants, and resume vegetative multiplication. (K'ari Fisher - Tetrahymean drawings, 2007)

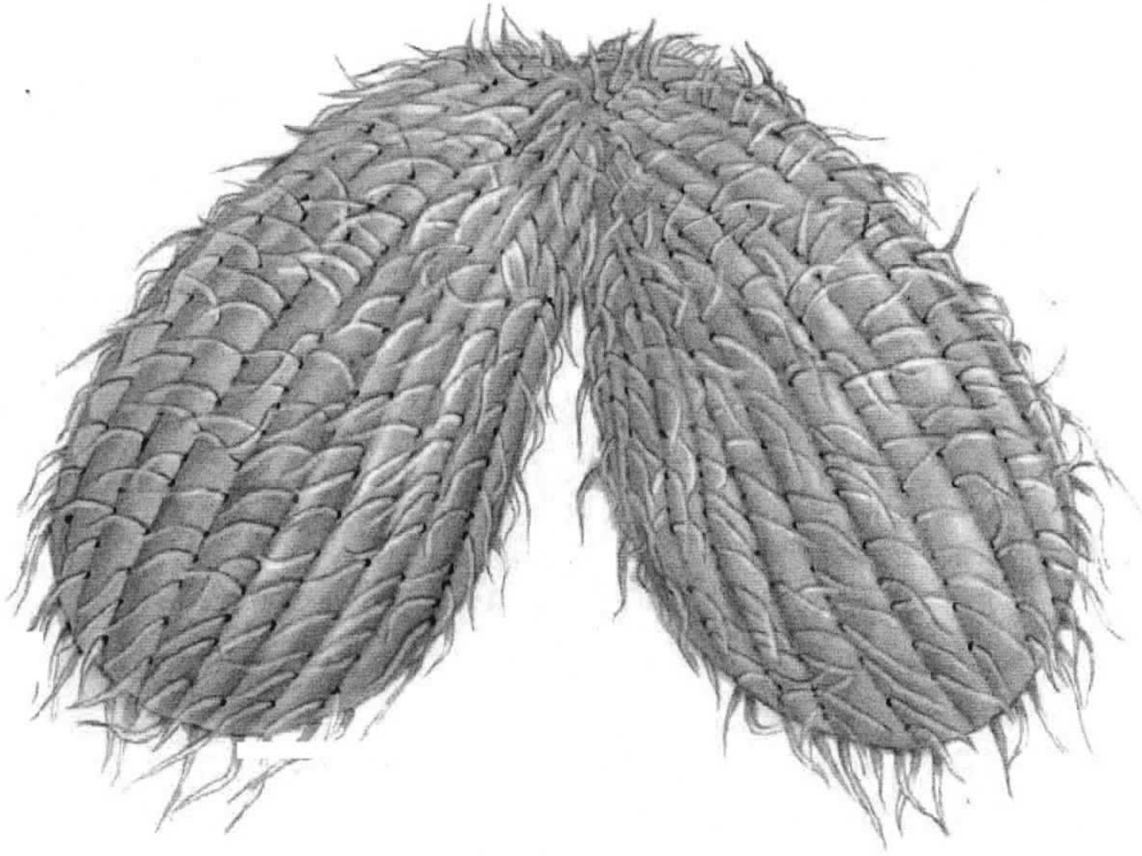


Figure 7: Pictorial Representation of *T. thermophila* Pairing during Conjugation. (K'ari Fisher - Tetrahymena drawing, 2007)

1.2 MISMATCH REPAIR IN *TETRAHYMENA THERMOPHILA*

A members of Cupples' laboratory (Derek Bell) has identified two *Tetrahymena* MutL homologs, related to the mammalian PMS2 and MLH1 genes, one MSH2 and four MutS homologs related to the mammalian MSH6 (Table 5). The Cupples' lab cDNA sequences (Bell, Mui, Roberts) revealed several differences from the gene structures predicted by TIGR. The corrected genes are shown in Table 5; among the corrections are exon placements.

Table 5: Summary of MutL and MutS homologs identified in *Tetrahymena thermophila.* Molecular weight of each of the MMR proteins was deduced from the corresponding amino acid sequence. Also indicated are the corresponding eukaryotic MMR homologs which the putative *T. thermophila* homologs most closely resemble.

TDG name	**Official name**	**cDNA size**	**Genomic size**	**# amino acids**	**kD**	**TIGR correct**	**Eukaryotic homologs**
27.m00362	msh2	2442	2623	813	93	No	Human MSH2 Yeast MSH2 Paramecium MSH2
16.m00421	msh6_1	3699	4046	1232	142	Yes	Human MSH6 Rat MSH6 Mouse MSH6
47.m00253	msh6_2	3405	4642	1134	131	No	Human MSH6 Mouse MSH6 ZebrafishMSH6
12.m00434	msh6_4	4170	4579	1389	161	No	Human MSH6
11.m00326	msh6_3	3774	4046	1257	147	Yes	Human MSH6 ZebrafishMSH6 Mouse MSH6
9.m00384	mlh1	2271	4553	756	87	No	Human MLH1 Rat MLH1 Mouse MLH1
239.m00043	pms2	2841	3018	946	110	No	Human PMS2 S.pombe PMS1 Arabidopsis PMS1

Phylogenetic analysis comparing the first 180 amino acids of the two MutL homologs (PMS1 , PMS2) to the same residues of bacterial MutL and eukaryotic MLH1, MLH3 and PMS2 protein families confirms that they belong to the MLH1 and PMS2 families; analysis of full length sequences from a bigger data set gave very similar results.

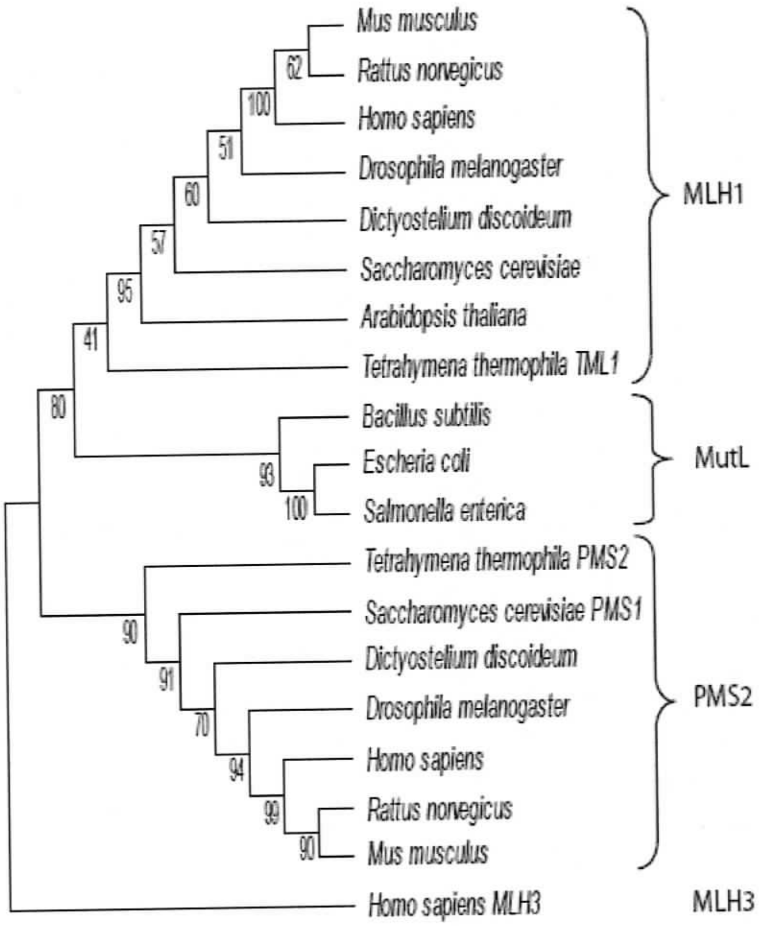


Figure 8: Phylogenetic analysis of *T. thermophila* MutL homologs. (Erin Annendale)

The Cupples' laboratory has chosen to compare the mitotic and meiotic functions of eukaryotic MutS and MutL homologs using the unicellular, ciliate protozoa *T. thermophila.* The partitioning of vegetative and meiotic functions in the same cell makes Tetrahymena ideal for studying the diverse functions of the MMR proteins. Microscopic studies in the lab have shown that Pms2 protein localizes solely to the micronucleus, suggesting that it does not play a role in vegetative MMR in the macronucleus. (Localization is detected by GFP fusion to *T. thermophila* MMR proteins under MMT1 promoter.) Pms2 localizes early in the nuclear differentiation to the new micronuclei. This is when transcription of genes increases during meiotic prophase being in sync with meiotic recombination (unpublished). Further localization studies indicate that Tetrahymena MLH1 localizes to the macronucleus, MSH-2 localizes to mic, and MSH6-7 shows localization on the periphery of the cell- possibly the basal bodies of Tetrahymena. Through quantitative PCR studies there is a meiosis-specific increase in PMS2 and in the other six MMR genes. These findings suggest that *Tetrahymena* will provide an attractive model system for defining the ròle of PMS2 and other MMR genes in meiotic recombination and overall cell function.

1.4 OBJECTIVES OF THE STUDY

The objectives of this study were to:

(1) Amplify and clone six MMR genes from *Tetrahymena*

(2) Construct an expression cassette for expressing his-tagged proteins for a *Tetrahymena* homologously recombinational vector pKK101

(3) Clone *T. thermophila* MMR genes and human TDG gene into pKK101

(4) Clone *T. thermophila* MMR genes via Gateway® into rDNA expression vector

(5) Transform Tetrahymena biolistically and via conjugant electrotransformation (CET) followed by selection of transformants.

(6) Analyze *Tetrahymena* protein extracts for the presence of recombinant his-tagged proteins by anti-his antibody Immunoblotting.

CHAPTER 2: MATERIALS AND METHODS

2.1 Bacterial strains, growth conditions and culture media

E. coli strain DH5α (F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)*U169, hsdR17(r_K^- m_K^+), λ–) was used for the routine propagation of plasmid DNA. SURE® *E. coli* strain (*e14– (McrA–) (mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5* (Kan^r) *uvrC* [F' *proAB* *lac*I^q*ZM15 Tn10* (Tet^r)] (Stratagene) was used for difficult to clone *T.thermophila* DNA constructs. These strains were routinely grown at 37°C for 24h with aeration from a single colony of each strain. Rich medium was made with Difco™ LB Broth, (Lennox) at 20g per liter. Antibiotics and supplements were added as required. Both the SURE® and DH5α *E. coli* strains were made electrocompetent according to the recommended protocol (Sambrook *et al.,* 2000)

2.2 Tetrahymena strains, growth conditions and culture media

The wild-type *Tetrahymena thermophila* strains CU427, CU428 and paclitaxel-sensitive strain CU522 were generously provided by Ron Pearlman. All cell cultures were grown in super peptone protease (SPP) medium (2% proteose peptone, 0.1% yeast extract, 0.2% dextrose, 10uM $FeCl_3$, 250ng/ml Penicillin and Streptomycin) supplemented with 100 U/ml Penicillin, 100 μg/ml Streptomycin, and 0.25 μg/ml amphotericin B (SPPA).

2.2.1 Measuring Concentration of *T.thermophila*

For measuring concentration of *Tetrahymena* cells either for growing or for mating preparations, 100 μl of cells were killed and stained with 1 μl of iodine for 1 minute. Then 10 μl aliquots of the prepared cells were analyzed under a microscope with a hemocytometer to determine cell concentration.

2.3 Construction of *Tetrahymena* homologous recombination vectors

2.3.1 Vector - pBICH3

pBICH3 (Gaertig *et al,* 1999) was kindly provided by Ron Pearlman (Figure 9). pBICH3 contains the coding region of the surface antigen gene of the ciliate *Ichthyophthirius multifiliis* (1340bp) inserted between the flanking sequences of the *BTU1* gene of *T. thermophila,* made up of 5'*BTU1* (990bp) and 3'*BTU1* (~1400bp). This cassette is in the context of Bluescript SK plasmid containing an ampicillin resistance marker. The *BTU1* regions of the plasmid targets the gene of interest to the endogenous *BTU1* locus of *T. thermophila.*

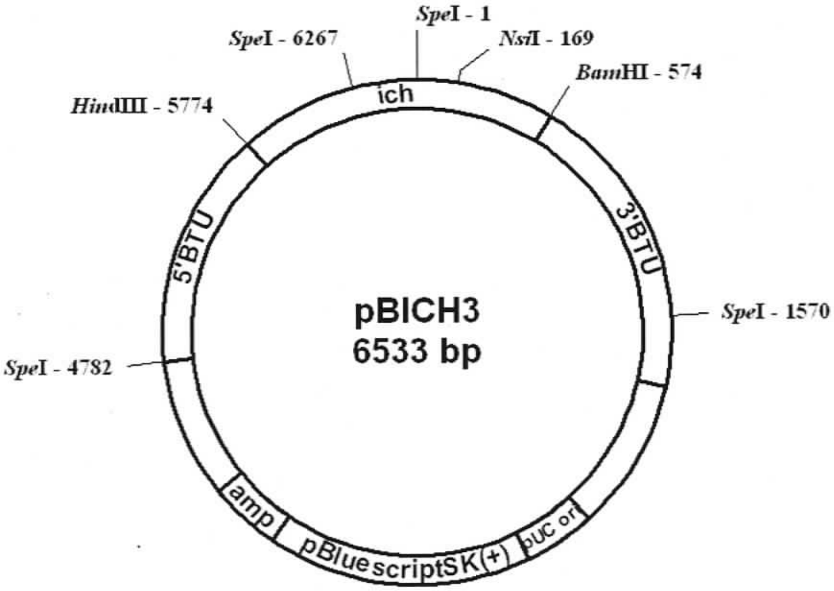


Figure 9: Vector map of pBICH3. The pBICH3 vector homologously recombines into the *BTU1* locus of the *T.thermophila* genome. (The restriction sites indicated on the map are the ones used during this study)

The identity of the pBICH3 plasmid was confirmed by restriction enzyme digests. Approximately 300-500ng of pBICH3 plasmid was digested with NdeI and a buffer recommended by the supplier (NEB) for 2h at 37°C. The reaction was analyzed on a 1% agarose gel in 1x TAE buffer at 100V and the pattern of restriction fragments was compared to the expected fragment pattern.

2.3.2 Construction of pKK101

pBICH3 was used to construct another derivative, pKK101, which contains a 6xHistadine Tag followed by a thrombin cut site and a multiple cloning site cassette. The 6xHis Tag, Thrombin cut site and MCS site cassette were assembled from two DNA oligonucleotides purchased from Alpha DNA (Table 6). Restriction sites were added at 5'(HindIII) and 3' (NsiI) ends of the synthetic oligonucleotides to facilitate cloning.

Table 6: Oligonucleotides used in the assembly of the 6xHis Tag, thrombin cut site and MCS site cassette for making pKK101.

Name	Length	Sequence
Oligo#1	81bp	5'AGCTTAGGCATGCACCATCAT CATCA TCAC AG CAGCGGC CTGGTGCCGCGCGGCAGCACCATGGCACGTGGGCGCCATGCA
Oligo#2	73bp	3' ATCCGTACGTGGTAGTAGTAGT AGTGT CGTC GCCGG ACCACGGCGCGCCGTCGTGGTACCGTGCACCCGCGGT'5

Annealing of the two oligonucleotides for the HisTag:Thr:MCS cassette was done by mixing equamolar amount (50pmol) of each oligo in 1X SSC in 50 µl total reaction volume. The annealing reaction was carried out in a Biometra® TGradient thermal cycler with the following program: (i) heat to 95°C and remain at 95°C for 2 minutes, (ii) ramp cool to 25°C over a period of 50 minutes. The annealed oligonucleotides were run on a polyacrylimide gel for verification and then stored on ice or at 4°C until ready to use.

In order to prepare the pBICH3 for cloning in the HisTag:Thr:MCS cassette a double restriction endonuclease digest was preformed on pBICH3 with HindIII and NsiI to excise part of the gene (928bp) coding sequence for protein ICH. This created sticky ends for cloning in the 81 bp cassette containing the HisTag:Thr:MC. Following the double digest pBICH3 vector was gel purified and the remaining 5605bp were ligated by T4 ligase with the HisTag:Thr:MCS cassette to produce the *Tetrahymena* expression vector pKK101. The ligation reaction was carried out as recommended by the manufacturer. The ligation reaction was then transformed into DH5α cells. Following the transformation, colonies were picked and screened for appropriate vector, and then the pKK101 vector was verified by restriction digest analysis.

2.4 *Tetrahymena* DNA and RNA isolations

2.4.1 Extraction of Total DNA from *Tetrahymena* (based on the method of Gaertig *et al* 1994)

To extract total *Tetrahymena* DNA, 25 ml of stationary-phase culture was concentrated to 0.5ml and combined with 3.5ml of 42% (wt/vol) urea/0.35 M NaCl/ 10 mM Tris-HCl, pH 7.4/ 10 mM EDTA/ 1% SDS. The solution was gently shaken until homogeneous and was then extracted twice with an equal volume of phenol/chloroform/isoamyl alcohol (24:1). Subsequently 1 ml of 5M NaCl was added to 3 ml of aqueous phase and DNA was precipitated with an equal volume of isopropyl alcohol. The DNA was spooled on a glass rod, washed with 70% (vol/vol) ethanol, and dried. The sample was suspended in 600µl of 10 mM Tris-HCl/ 1 mM EDTA, pH 8, plus

6 µl of RNase A (10mg/ml) at 55°C overnight and then centrifuged at 31, 000 x g for 45 minutes.

2.4.2 RNA Isolation and cDNA preparation from *Tetrahymena*

To extract mRNA, 5ml (2.5 x 10^6 cells) of *Tetrahymena* cells were centrifuged at 3200g for 5 min and resuspended in 1ml TRI Reagent (Ambion). Samples were taken from log phase cells. Total RNA isolation was performed according to the manufacturer's protocol (Ambion). Following isolation, RNA was resuspended in DEPC-treated RNase-free water and stored at -80°C. Traces of contaminating genomic DNA were removed from all RNA samples using the TURBO DNA-free kit (Ambion). 1µl of total RNA was annealed with 50ng random hexamers and cDNA was synthesized using the Thermoscript RT-PCR System (Invitrogen). cDNA samples were treated with RNaseH to remove complementary RNA.

2.5 Cloning *T. thermophila* mismatch repair genes into pKK101

2.5.1 Amplification of the six putative *T. thermophila* MMR genes

Nested temperature gradient PCRs were preformed on all of the genes with Phusion™ (NEB) polymerase. 20 µl PCR reactions contained 4 µl of 5x Phusion™ HF buffer, 0.4 µl of 10 mM dNTPs, 1.25 µl of each specific forward and reverse primer (Table 7), 1.5 µl of template DNA and 0.2 µl of Phusion polymerase. The reactions were amplified in a Biometra® TGradient thermal cycler with the following conditions: an initial denaturation at 98°C for 3 minutes followed by 35 cycles of a denaturation 98°C for 30 seconds, gradient annealing temperatures ranging from 48.5 °C to 61.5 °C for 45 seconds, and an initial extension at 72 °C for 4 minutes, with a final extension at 72 °C for 10 minutes. Amplicon sizes were confirmed by gel electrophoresis run on a 1% agarose gel in 1x TAE buffer at 100V. The correct sized PCR fragments were then gel purified using the QIAquick gel extraction kit (QIAGEN) according to the manufactures recommendations in preparation for cloning.

Table 7: Primers used in PCR amplification of *T.thermophila* MMR genes for the purpose of cloning into pKK101

Name	**Size in bp**	**Restriction site**	**Sequence form 5'to 3'**
3238-F	35-mer	NcoI	ccagcatcc atg gaa atgtcgccaaacaaaaaaac
3238-R	35-mer	BamHI	cgatgca ggatcc tcttgagcttatttgtgaattc
4645-F	37-mer	NcoI	cgtcacc atg gat atg aatcattcctaacccttagaa
4645-R	36-mer	alternative	acgac ggatcc tgaataac tcaacatctttcgaaga
504-F	32-mer	NcoI	cgtgctcc atg gAA atg ataaagaagcttgac
504-R	35-mer	BamHI	acggcatggatcctcagtctcaaagatcatatgtc
3288-F	46-mer	NcoI	cgtacccatggaagaggaaattgaattagttaatgataagaacttg
3288-R	36-mer	BamHI	ggcatggatccacaaagaacttaatatgttttcaag
3297-F	34-mer	NcoI	acgcat cc atg gca atgcaaatcaacttgaaagc
3297-R	35-mer	BamHI	acggat ggatcc tgaactaattattatggaagctt
174-F	29-mer	alternative	tgcatCCATGGAAatgaagcaaagttcac
174-R	32-mer	alternative	catgc ggatcc ttgattttctgcaagaaatgc

2.5.2 Cloning pKKMSH2

The PCR purified (Qiagen) MSH2 PCR product was cloned into pDrive (Qiagen) PCR cloning vector (Figure 10). The MSH2 PCR product was obtained with the primers listed in Table 7 in the 3288 row, with the addition of 0.2 µl of Platinum Taq DNA polymerase (Invitrogen) at cycle 30 of the PCR reaction outlined in the above section 2.6.1 in order to generate poly A tails for ligation into pDrive. The ligation reaction between MSH2 MMR gene and pDrive was carried out according to the procedure outline in the pDrive manufacture protocol (Qiagen).

A double endonuclease digestion of MSH2 from pDriveMSH2 with NcoI and BamHI was carried out and the appropriate size fragment was gel extracted using QIAquick gel purification kit (Qiagen) in preparation for pKK101 cloning.

pKK101 was prepared by a double digest with NcoI and BamHI. Next, a ligation reaction with T4 ligase (Invitrogen) of insert (MSH2 derived from pDrive) and vector pKK101 was carried out at 16°C overnight according to the manufacturer's protocol. Next SURE® *E. coli* cells were electrotransformed with the ligation reaction. The transformants were plated on LB ampicillin plates and colonies where then screened for the correct clone. This was done by isolating the plasmids using the QIAprep Spin Miniprep Kit (Qiagen). The correct clone pKKMSH2 was verified by restriction digest with NcoI and BamHI.

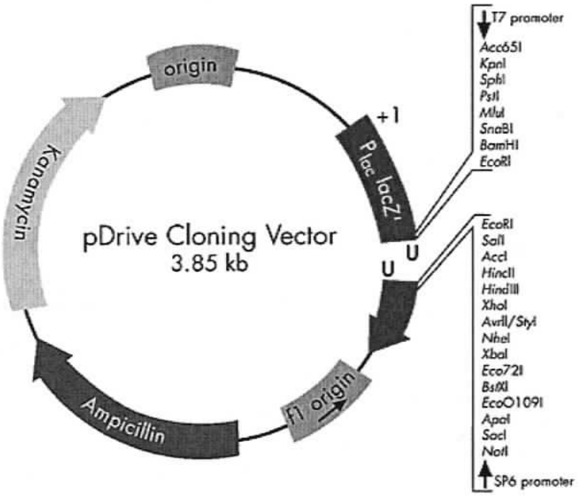


Figure 10: pDrive cloning vector. This vector was used as an intermediate in cloning the putative MMR gene from *T.thermophila* into an expression vector pKK101.

2.6 Cloning Human Thymine-DNA Glycosylase into pKK101

The Thymine-DNA glycosylase (TDG) gene was first amplified using the primers in Table 8 from pET28a vector. A double digest of the amplicon with NcoI and BamHI was carried out for ligation into pKK101. pKK101 was prepared as outlined in section 2.6.2 Then a ligation with T4 DNA ligase (Invitrogen) at 16°C overnight according to the manufacturer's protocol was carried out. The final vector was verified by restriction endonuclease digest.

Table 8: TDG Primers for amplification from pET28a.

Name	**Sequence from 5' to 3'**	**Features**
TDG-F	acgatgcgtccggcgtagagga	This primer amplifies TDG upstream NcoI and start codon
TDG-R2	Tgcatggatcctcaagcatggctttc	With a stop codon specific for *T. thermophila* and a BamHI site

2.7 rDNA Vectors and Gateway® Cloning

2.7.1 Tetrahymena episomal rDNA vector - pICSH

The Gateway® (Invitrogen) compatible rDNA vector pBICSH (Figure 11) was generously provided by Douglas Chalker. The Gateway® compatible destination vector pICSH is built on the rDNA and pIC19H backbone containing the MTT1 cadmium inducible promoter and a C-terminal histadine/streptavidin tag. pICSH contains a Gateway® recombinase cassette consisting of att sites (attR1 and attR2), chloramphenicol resistance gene (Cm^r) and *ccd*B gene (Figure 11).

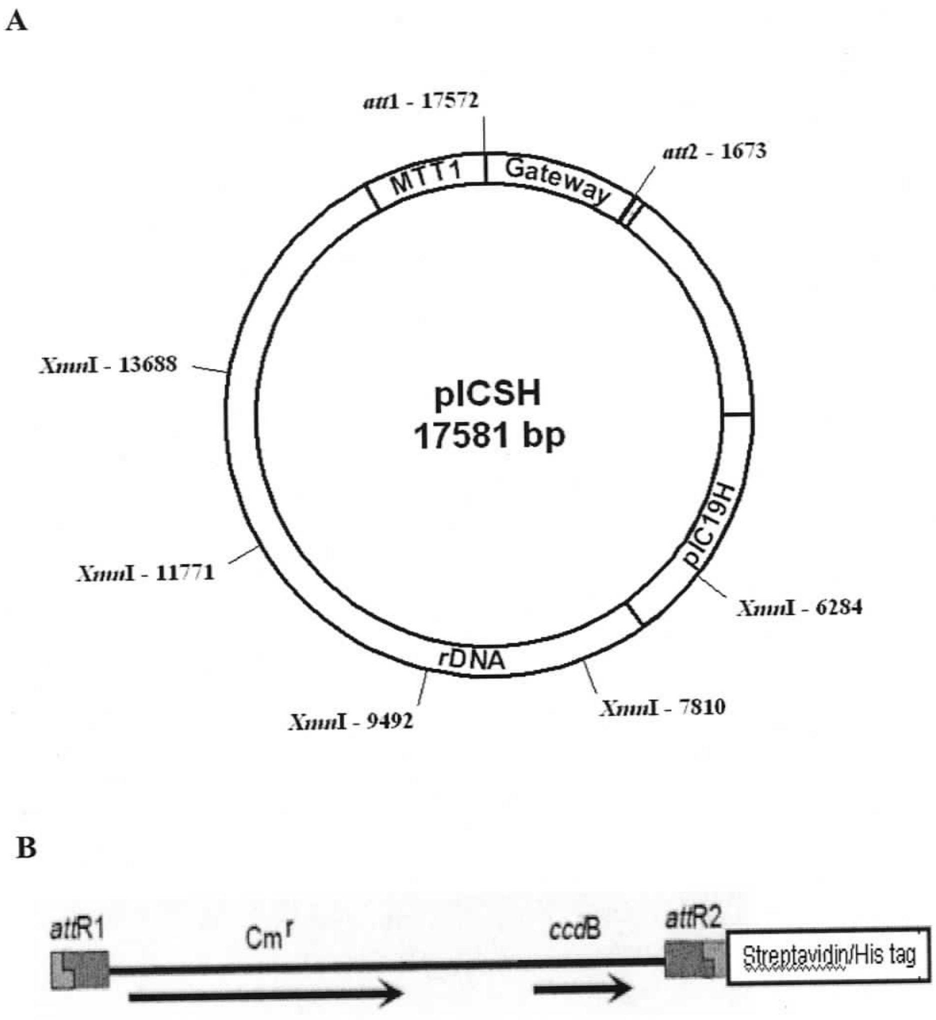


Figure 11: pICSH Vector Map a) vector map of pICSH, **b)**The schematic of the GATEWAY® cassette, comprised of att sites (attR1 and attR2), chloramphenicol resistance gene (Cm^r) and *ccd*B gene.

In order to create Strep/His fusion proteins with the MMR genes the Gateway® cloning system was utilized. The identity of pICSH plasmids was confirmed by restriction enzyme digests as outlined in section 2.3.1 (NsiI restriction endonuclease enzyme was used in this case).

2.7.2 Gateway entry vectors – Gateway® BP reaction

T. thermophila MMR genes were amplified with appropriate primers containing the specific att sites and subsequently cloned (recombined) into pENTR-D vectors. The entry vectors used and created via BP Gateway reaction were: N-504, N-3288, N-174, N-c4645, and N-3238. The entry vectors were kindly provided by Erin Annandale.

2.7.3 Gateway expression vectors – Gateway® LR reaction

The final expression vectors pICSH-PMS2, pICSH-MSH2, pICSH-MLH1, pICSH-MSH6-3 and pICSH-MSH6-1 were recombined by the Gateway® LR reaction. All of the reactions were carried out in accordance to the manufacturer's protocol. The identity of the final expression vectors was confirmed by restriction digests (Table 9). Due to the large size of the plasmids (17kb and up) Qiagen Plasmid Midi Kit was used to purify the rDNA plasmids.

Purified rDNA plasmid	**Restriction enzymes used for single digest**
pICSH (control)	NsiI, XmnI, NdeI
pICSH-MSH6-1	XmnI
pICSH-MSH2	NsiI
pICSH-PMS2	NdeI
pICSH-MSH6-3	XmnI
pICSH-MLH1	XmnI

Table 9: Restriction endonuclease enzymes for screening successful LR Gateway® reactions. pICSH is the control rDNA vector; the rest of the rDNA vectors contain MMR gene inserts.

2.8 DNA Transformation of *Tetrahymena thermophila*

2.8.1 Biolistic Transformation of *Tetrahymena thermophila.*

T. thermophila strain CU522 was used as a transformation host. This strain carries a single substitution (Lys350Met) in the β-tubulin BTU1 gene, which confers increased resistance to several microtubule depolymerizing drugs, such as oryzalin and colchicines, and increased sensitivity to microtubule stabilizing agent, paclitaxol (Geartig *et al,* 1999). Prior to the transformation, CU522 cells were grown in 50 ml of SPPA medium, in 250 ml Erlenmeyer flasks with shaking at 150 rpm at 30°C to a density of about 6×10^5 cells/ml. Twenty-four hours before transformation, 50 ml of growing cells were washed and suspended in 10 mM Tris-HCl (pH 7.5) buffer in the original volume. After 4–5 h starving, cells were counted again, cell concentration was adjusted to 3×10^5 cells/ml, and cells were left at 30°C without shaking for another 18–20 h .

To target the *BTU1* gene derivatives to the endogenous *BTU1* locus of *T. thermophila*, the transformation plasmids pKK3288, pKKTDG and pCH504 were linearized.

pKK3288 was digested with *XhoI* and *EagI*, restriction endonucleases to separate the insert (*btu1-2::3288*) from the plasmid. pKKTDG was digested with *KpnI* and *SacI* restriction endonucleases to separate the insert (either *btu1-2::TDG*). The particle bombardment of Tetrahymena was carried out with the procedure outlined by Hanley-Cassidy *et al.*, (1997). Fifty micrograms of digested DNA were purified by a single phenol/chloroform/isoamyl alcohol (25:24:1) extraction followed by chloroform/isoamyl alcohol (24:1) extraction, precipitated with an equal volume of isopropyl alcohol in the presence of 0.15 M sodium acetate, dried and resuspended. Ten micrograms of linearized plasmid DNA were used to coat 60 mg of 0.6 µm gold particles (Bio-Rad). Next, 2.4 × 10^7 of starved *Tetrahymena* cells were spun down at 600 G for 4 min, washed with 30 ml of 10 mM Tris (pH 7.5), and resuspended in 1 ml of Tris buffer. One milliliter of cells was bombarded using 1-5 µg of DNA-coated gold particles at 900 psi using DuPont Biolistic PDS-1000/He particle delivery system (Bio-Rad).

2.8.2 Conjugant Electrotransformation (CET) of *T. thermophila*

Tetrahymena strain CU427 and CU428 (all strains are Pm^{r+}/ Pm^{r+} [paromomycin-sensitive (pm^s)] and contain B-type rDNA) were grown to 5 x 10^5 cells/ml and washed once with Tris HCl (pH 7.5) prior to overnight starvation at 30°C without shaking. Cells were then recounted and equal volumes of cells were mixed and incubated at 30°C without shaking. Ten to 11 hours after the initiation of mating the cells were pelleted, washed and resuspended in 10mM HEPES pH 7.5. Cells were transferred to 0.4µm cuvettes with 1-5µg of vector DNA in HEPES and pulsed at 0.44kV, 200Ω, 25uF using the Bio-Rad Gene Pulser. A 1:100 dilution in SSP media was then plated onto three 96 well plates and incubated overnight at 30°C. The conjugant electrotransformation was carried out as developed by Gaertig and Gorovsky (1992).

2.9 Selection of *T. thermophila* Transformants - Phenotypic Assortment

2.9.1 Paclitaxol selection

Following biolistic transformation with linearized plasmid the bombarded *Tetrahymena* cells were resuspended in 50 ml of pre-warmed to 30 °C SPPA, left for 2–3 h at 30°C with shaking to recover. Paclitaxol was added to a final concentration of 20µM. The cells were then transferred into 96 well microtiter plates, 200ul cells/ well, incubated for 3-4 days at 30°C, in a humidified, dark (paclitaxol is light sensitive), chamber. Also, included were mock transformants to monitor the efficiency of the drug for killing the untransformed *Tetrahymena* cells. Visual checks were made under the microscope. Wells containing paclitaxel-resistant transformants (blooms) were apparent after 6-7 days of selection. Complete replacement of all endogenous *BTU1* gene copies by transformed *BTU1* derivative was achieved using paclitaxel selection for phenotypic assortment by the method of Gaertig *et al* (1994). The cells were passaged into fresh medium containing increasing amount of paclitaxel up to 40μM concentration, by transfering 20 ul to 180 ul of fresh SSPA medium plus paclitaxel. Single cell isolations were made by picking a single cell into a drop of SPPA media. To verify complete assortment the cells were lysed and diagnostic PCR was carried out.

2.9.2 Paromomycin selection of CET transformed strains

100 μg/ml of paromomycin (Sigma) was added to *T. thermophila* cells transformed with rDNA plasmid and incubated for 3-4 days until significant growth was evident in some wells (blooms). Resistant cells were matured with continual passage of media for 7-10 days. Single cells were then isolated, cultured and tested.

2.9.3 Diagnostic PCR for *Tetrahymena* transformants

Once the cell bloom was established; 300-400 cells were lysed in 30 μl of lysis buffer : 0.5% Tween-20, 0.5% NP-40, 10 mm Tris-HCl (pH 8) and 300 μg/ml proteinase K, for 20min at 65°, followed by incubation at 95° for 10 min. One-tenth of each lysate was used as a template in a consecutive PCR reaction.

Table 10: Primers used for the verification of *T. thermophila* transforments. Transfroments checked, pKKMSH2, pKKTDG and CH-504.

Name	**Sequence**	**Checking Transforments**
pGTF	5' GGTAAATAAAGGGCGTTGCAC	pKKMSH2 pKKTDG cH-504
pGTR	5' TAGCTGACCGATTCAGTTCG	pKKMSH2
p3288R	5'GGCATGGATCCACAAAGAACTTAATATGTTTTCAAG	pKKMSH2
pTDG-R2	5'TGCATGGATCCTCAAGCATGGCTTTC	pKKTDG
p504-R1	5' TTTAAAGGGATTCACTGGGCGA	CH-504

2.10 Protein Extracts from *Tetrahymena* Cells for SDS-PAGE and Immunoblot analysis

2.10.1 *Tetrahymena* SDS-PAGE

Tetrahymena protein lysates were prepared from exponentially growing cells $2x10^5$ cells/ml. 10ml of cells @ approx 1-2 x 10^6 cells were spun down for 5 min at 2000 rpm @ 4°C. The pellet was washed with ice cold 10mM Tris-HCl at pH 7.5 and spun one more time. The cells were then resuspend in 125µl ice-cold Tris buffer supplemented with a mixture of protease inhibitors. Next, the sample was sonicated on ice. When sonication was considered completed the cells visually checked for rupture under the microscope. Then 125 µl of boiling 2x SDS-PAGE sample buffer was added and boiled for 7 minutes, after which the cells are centrifuged for 5 minutes at maximum speed in a bench top microfuge. 10-12 µl of the cell lysates was loaded per lane and run on 10% SDS-PAGE gel. In order to visualize proteins, gels were incubated for 60 minutes at room temperature in a Coomassie staining solution (0.2% w/v Brilliant Blue G

(Sigma), 45% methanol, 45% water, 10% acetic acid) and then destained overnight in destining solution (40% methanol, 10% acetic acid, and 40% water).

2.10.2 Western Blot Analyis

After separation by 10% SDS-PAGE, the proteins in the gel were electrophoretically transferred onto a PVDF membrane at 28 volts for 1 hour, then at 84 volts for 14-16 hours at 4°C in transfer buffer (22 mM Tris base, 19 2mM glycine, and 20% v/v methanol) using a Bio-Rad Mini Trans-Blot Cell. The blotted membranes were then incubated in 40ml blocking/incubation buffer (10% w/v non-fat dry milk, 1xPBS, 0.05% Tween™-20) for 1-16 hours with gentle shaking. Next, the membrane was incubated with in a 1: 3 000 dilution of the primary anti-his antibody (Amersham) in blocking/incubation buffer for 1 hour at room temperature with gentle mixing. The membrane was then washed 3 x 5 minutes in a wash buffer (1xPBS, 0.05% Tween-20) and then incubated for 1 hour in 20 ml of 1:4 000 secondary antibody, IRDye800 conjugated affinity purified anti-mouse IgG (goat) (Rockland) in blocking/incubation buffer. After the membrane was washed 3 x 5 minutes with 20 ml of wash buffer the membrane was analyzed by the fluor base detection system (Licor Odyssey).

CHAPTER 3: RESULTS

3.1 Introduction and Overview

In previous studies in our laboratory *T. thermophila* MMR genes were identified and annotated including five MutS homologues and two MutL homologues (Table 5). My study began with the amplification of the *T. thermophila* MMR genes and cloning the genes into *T. thermophila* expression vectors. Following the amplification of the *T. thermophila* MMR genes, two different cloning approaches were investigated. The first approach focused on a homologously recombinational *T. thermophila* expression vector resulting in a MMR fusion protein with an N-terminal histadine tag under the endogenous β-tubulin promoter.

The second approach consisted of utilizing an episomal plasmid based on a *T. thermophila* rDNA replicon. The MMR genes were cloned into Gateway® compatible rDNA vector producing a C-terminal histadine/streptavidine tag this time under the control of a cadmium inducible metallothionein (MMT1) promoter.

The various vectors were transformed into the appropriate strains of *T. thermophila*. Again two different transformational approaches were taken depending on the type of vector used, whether homologous recombination or episomal rDNA. Vectors that homologously recombine into the *T. thermophila* genome (i.e. the pKK101) were introduced by biolistic transformation, while the rDNA vectors were introduced by the conjugant electrotransformation (CET) method.

Transformants due to homologous recombination were selected by phenotypic assortment (Gaertig et al., 1994a), by addition of paclitaxol. A mutant strain of *T. thermophila* CU522 was used for these transformations exploiting the drug sensitivity to paclitaxol, where knocking out the mutant β-tubulin gene restores resistance. For rDNA vectors paromomycin was used for the selection of transformants. The putative transformants were then confirmed by diagnostic PCR and finally *T. thermophila* protein extracts were obtained for visualization on SDS-PAGE and immunoblots.

3.2 Construction of a pKK101- Homologously Recombinational Vector

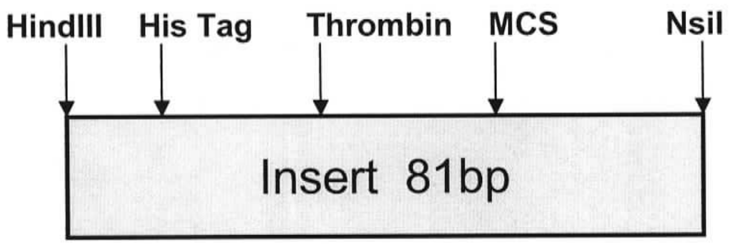
A derivative of pBICH3, pKK101 was constructed by cloning in the 6xHisTag:Thr:MCS cassette constructed from two oligonucleotides. The oligonucleotides were annealed and run on a 10 % polyacrylamide gel and stained with ethidum bromide for visualization (Figure 11).

The annealed oligonucleotides making up the 6xHisTag:Thr:MCS cassette were ligated into pBICH3 which was digested with HindIII and NsiI previously. The new clone pKK101 was obtained and tested by restriction digestion (Figure 12). As indicated in figure 12 A, lane 4 the double digest with BamHI and NcoI results in two bands 5259 bp and 427 bp corresponding with the expected bands for pKK101 (Figure 12).

3.3 Amplification of the Six Putative MMR genes from *T. thermophila*

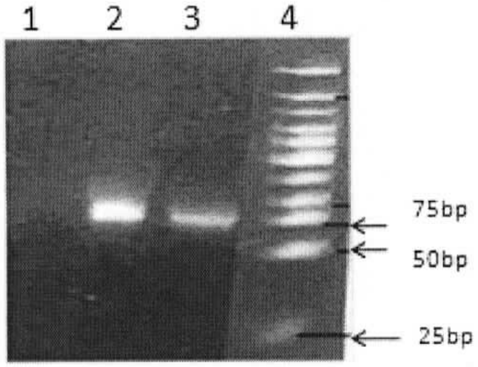
Nested temperature gradient PCRs were performed for all of the genes outlined in Table 11 from genomic Tetrahymena DNA substrate. The nested PCR was designed with two sets of primers for each gene. The "outer" set of primers amplified a larger section of the DNA containing the gene of interest and 5' and 3' flanks while the second set of primers the "inner" primers amplified the actual gene of interest with the primers containing specific restriction cloning sites on 3' and 5' to facilitate further cloning. All of the amplicons shown in Figure 13 are the products obtained with the "inner" set of primers (Table 7). The correct sized PCR

A



B

```
HindIII          SphI           6x His Tag                                     Thrombin  cut site                  NcoI      PmlI      NarI     NsiI
'AGC  TTA  GGC ATG'   CAC CAT CAT CAT CAT CAC  AGC AGC GGC  CTG GTG CCG CGC GGC AGC AC'CATGGCAC'GTGGG'CGCC ATGCA'
        '  AT  CCG TAC   GTG GTA GTA GTA GTA GTG  TCG TCG CCG  GAC CAC GGC GCG CCGTCG  TG GTACC GT G CA CCCGC'GGT'
```



C

Figure 11. Annealing two oligonucleotides for 6xHisTag:Thr:MCS cassette. A) the 6xHisTag:Thr:MCS cassette indicating the main components of the cassette. B) sequence of the two oligonucleotides, C) 10 % polyacrylamide gel run at 90V, stained in EtBr for 20 min, Lane 1: 5 µl of 15 ng/µl unannealed oligos, Lane 2: 10 µl of 15 ng/µl annealed oligos, Lane 3: 5 µl of 15 ng/µl annealed oligos, Lane 4: DNA MW ladder

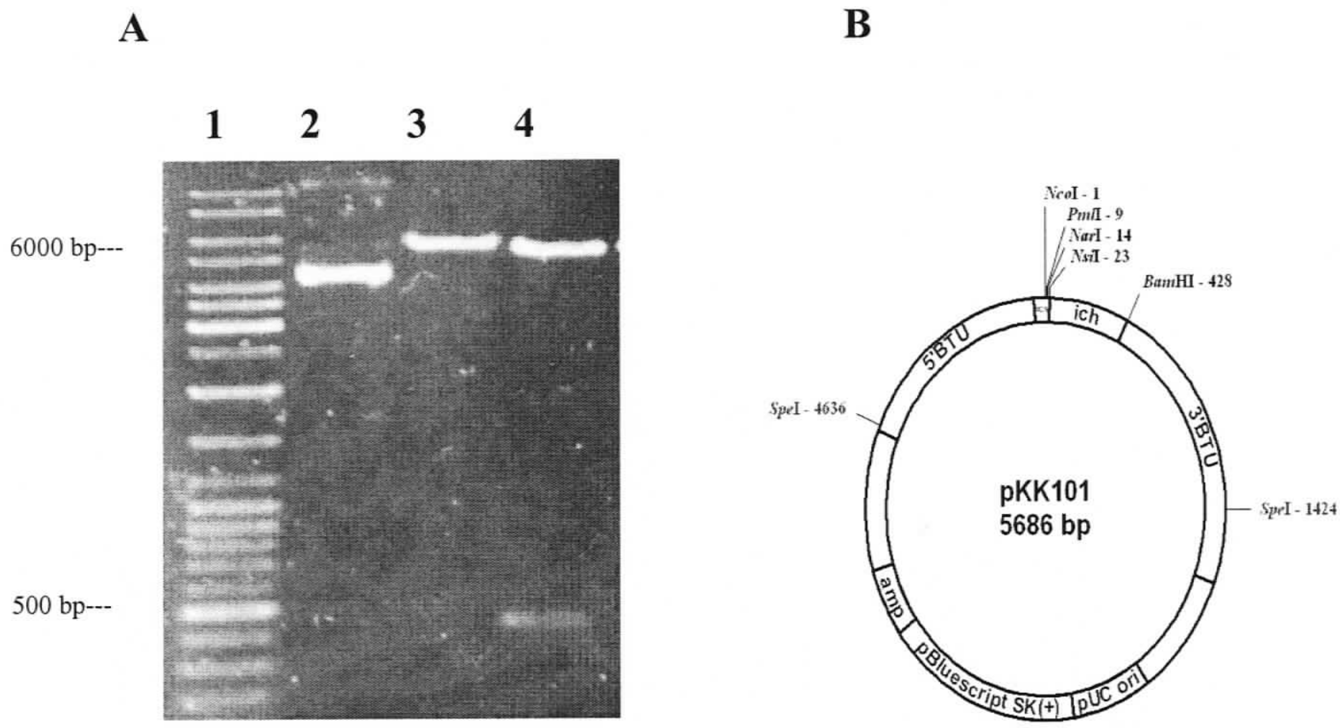
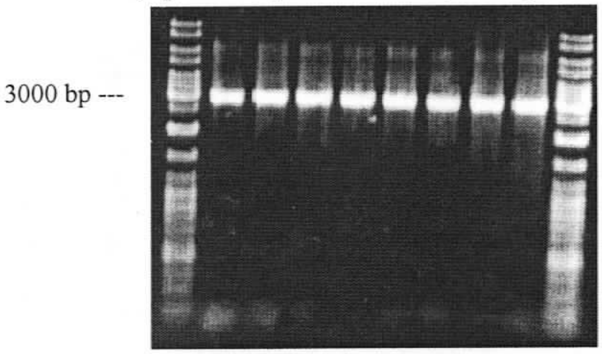
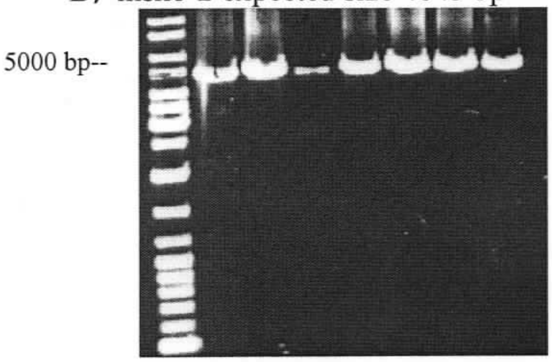


Figure 12: **Diagnostic restriction endonuclease digest of pKK101 with BamHI and NcoI.**

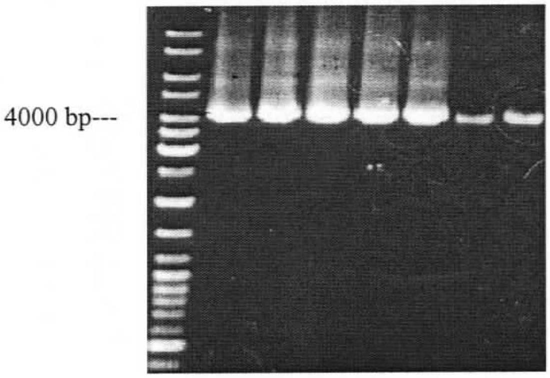
A) Lane 1: MW DNA ladder, Lane 2: undigested pkk101, Lane 3: pKK101 digested BamHI , lane 4: double digest with BamHI and NcoI. Gel run on 1% Agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min. **B)** Vector map of pKK101 Shown are all of the restriction endonuclease sites which were used in either cloning or identification of the plasmid.



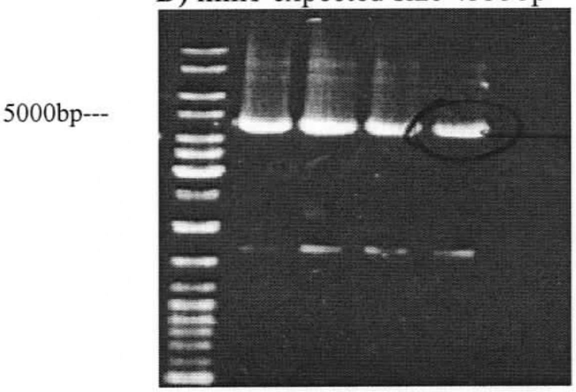
A) pms2 expected size 3,221 bp



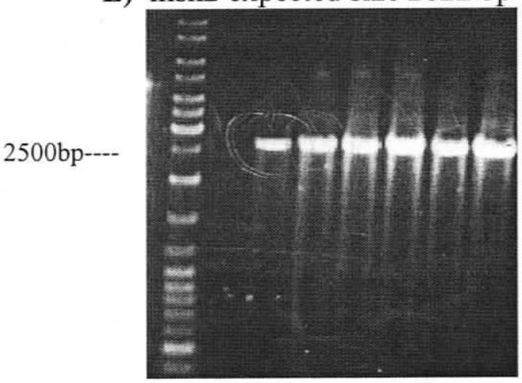
B) msh6-2 expected size 4645 bp



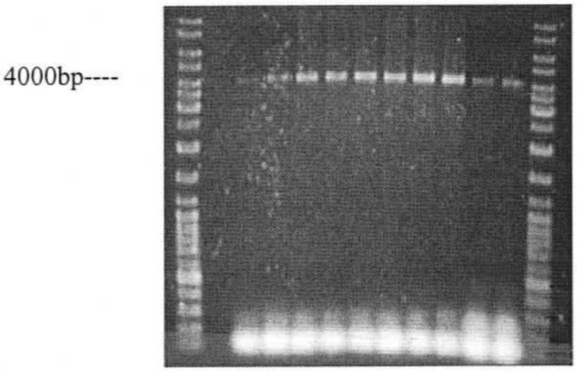
C) msh6-3 expected size 4045 bp.



D) mlh1 expected size 4555bp



E) msh2 expected size 2622 bp



F) msh6-1 expected size 4045bp

Figure 13: MMR PCR product for *T. thermophila.* All gels run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min.

Fragments were then gel purified in preparation for cloning into the pKK101 *T. thermophila* expression vector.

Table 11: *T. thermophila* MMR genes and their corresponding sizes.

MMR gene	genomic DNA bp	cDNA bp	# introns
pms2 (504)	3,221	3,045	2
msh6-2 (3297)	4,645	3,386	6
msh6-3 (3238)	4,045	3,774	2
mlh1 (4645)	4,555	2,281	8
msh2 (3288)	2,622	2,442	3
msh6-1 (174)	4,045	3,699	5

3.4 Construction of pKK3288 (MSH2)

The *T. thermophila* MSH2 was amplified (Figure 13 E), cleaned up and cloned into a PCR cloning vector pDrive. Then *E. coli* SURE® competent cells were transformed with the ligation reaction and plated on LB ampicillin plates with IPTG + X-gal for blue/white selection. The correct insert was digested out of pDrive with BamHI and NcoI (Figure 14), then gel purified and cloned into pKK101. PkkMSH2 was obtained; restricton digests with NcoI and BamHI confirmed the correct vector (Figure 15), expected bands match up with the obtained bands of 5359 bp plus 2645 bp.

3.5 Construction of pKKTDG

The TDG gene was amplified from pET28a, provided by Kevin So. The 3' primer contained one base difference in the last codon, changing the stop codon from UAA to UGA. *T. thermophila* uses only one stop codon (UGA) during protein synthesis, compared to the three

(UGA, UAG, UAA) which are standard in most eukaryotes; UAA and UAG encode glutamine instead (Gorovsky, et al., 1985).

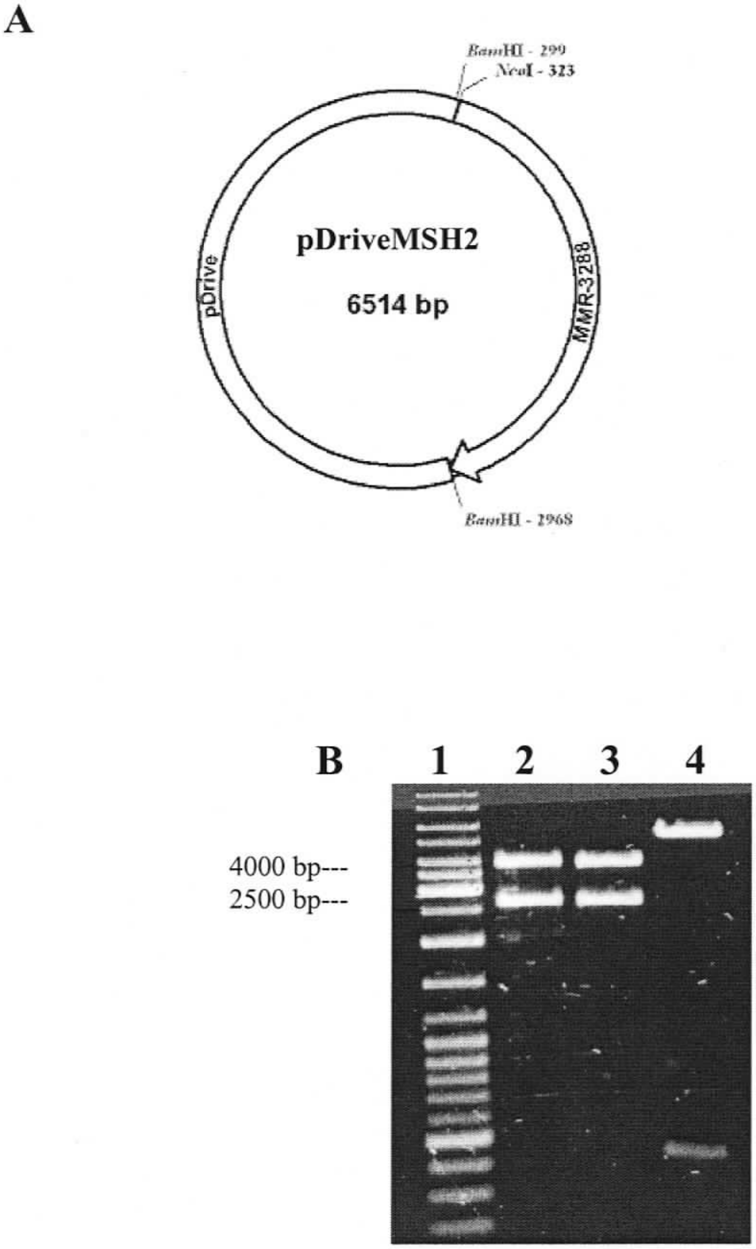


Figure 14: pDriveMSH2. **A** vector map of with MMR-3288 insert (MSH2) in pDrive. **B** pDrive with 3288 insert digested with NcoI and BamHI, Lane 1: MW DNA ladder, Lane 2 and 3 pD3288, Lane 4: not a correct clone. Run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min.

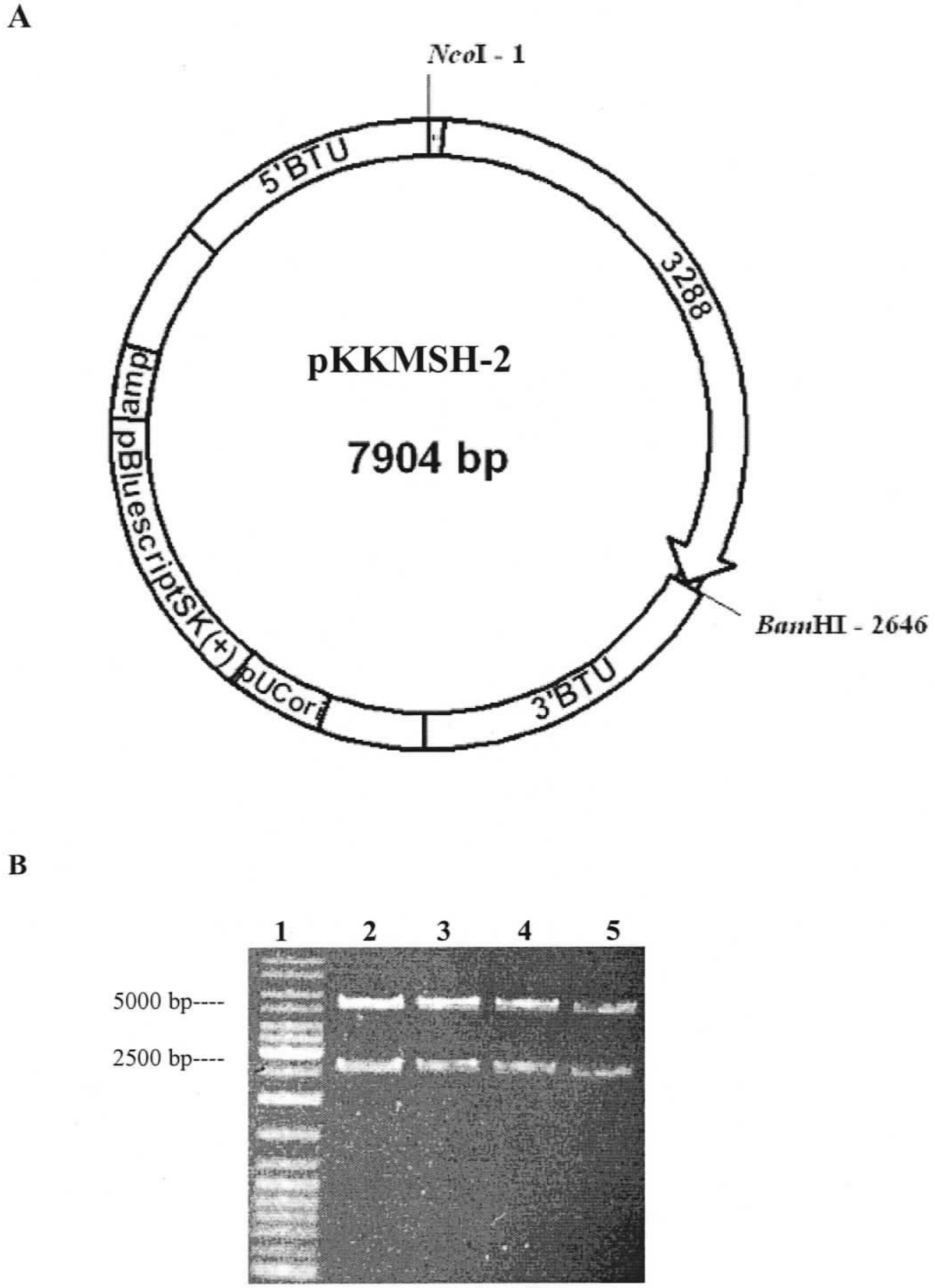


Figure 15: pKKMSH-2. **A** vector map of pKKMSH-2. **B** Digest of pKKMSH-2 with NcoI and BamHI. Lane 1: MW DNA ladder, Lane 2-5: pKKMSH-2 digested with NcoI and BamHI. Run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min.

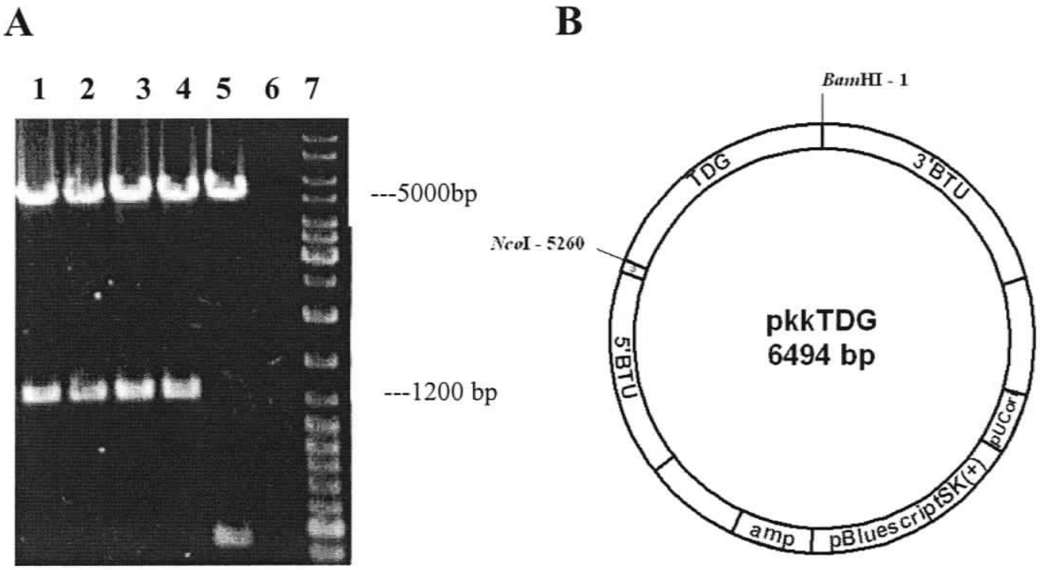


Figure 16: pKKTDG A) Diagnostic restriction endonuclease digest of pkkTDG with NcoI and BamHI, Lane 1-4: pKKTDG digested with NcoI and BamHI, Lane 5: (negative control-no insert) pkk101 digest with NcoI and BamHI. Run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min. B) vector map of pkkTDG.

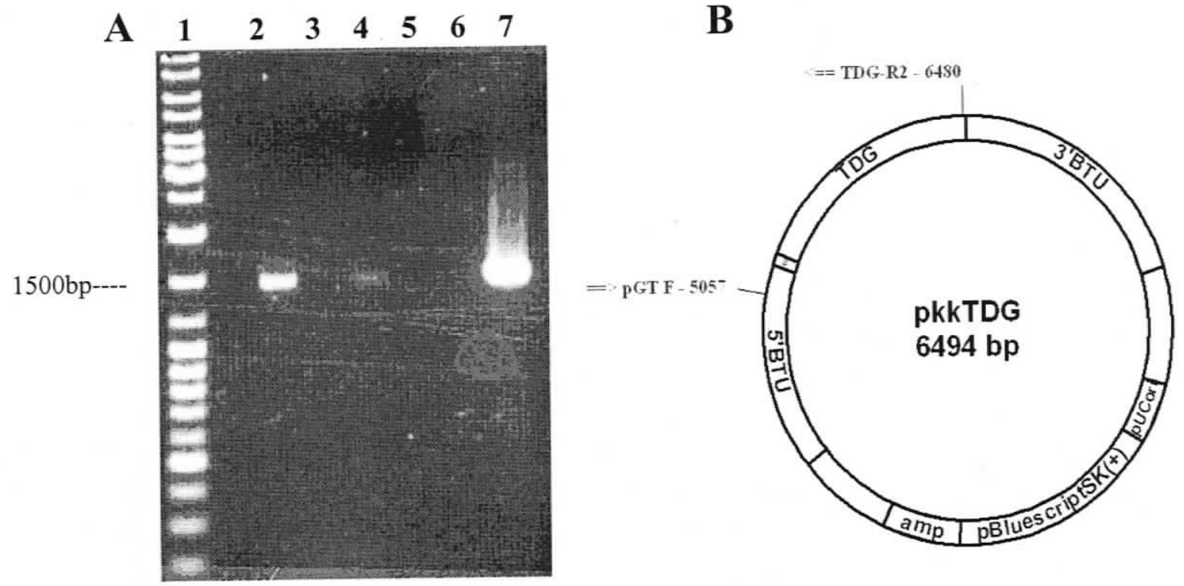


Figure 17: Diagnostic PCR for pKKTDG integration into *T. thermophila* via biolistic bombardment.

A) Ethidium bromide stained 1% agarose gel showing diagnostic PCR with primers TDG-R2 and pGTF. Lane 1: MW DNA ladder, Lane 2: *T. thermophila* sample 1, Lane 3: *T. thermophila* sample 2, Lane 4: *T. thermophila* sample 3, Lane 5: untransformed *T. thermophila* sample 4, Lane 6: no DNA control, Lane 7: (positive control) vector pKKTDG. B) pKKTDG vector map indicating the primers used and their location.

As in many other organisms, UGA itself is also used in some genes to encode the amino acid selenocysteine, making *T. thermophila* the only known organism to translate all 64 codons (Robinson 2006).

The TDG was cloned into pKK101 and confirmed by restriction endonuclease digestion with NcoI and BamHI (Figure 16).

3.6 Biolistic Transformations of *T. thermophila*

The constructs pKKTDG and pKK3288 were targeted into the BTU1 gene of the MAC genome of vegetative *T. thermophila* cells using biolistic bombardment DNA transformation. The CU522 *T. thermophila* mutant strain was used, it contains a single substitution (Lys350Met) in the β-tubulin *BTU1* protein, conferring increased resistance to several microtubule depolymerizing drugs such as oryzalin and colchicines, and increased sensitivity to a microtubule-stabilizing agent, paclitaxol (Gaertig *et al.,* 1999).

3.6.1 PCR Confirmation of pKKTDG *T. thermophila* transformant

Diagnostic PCR was carried out to check for TDG integration into *T. thermophila* β-tubulin region. Primers used TDG-R2 and pGTF (Table 10). The gel indicates (Figure 17) a band at approximately 1500 bp in lane 2, corresponding to *T. thermophila* transformant 1. There are no bands visible for *T. thermophila* sample 2 in lane 3 or the untransformed *T. thermophila* in lane a faint band is present in lane 4 corresponding to *T. thermophila* transformant 4.

3.6.2 PCR Confirmation of pKKMSH2 *T. thermophila* transformant

Diagnostic PCR of pKKMSH2 *T. thermophila* transformants was carried out to verify the integration of MMR gene MSH2 via biolistic transformation into β-tubulin region of *T. thermophila* genome. The pGTF/p3288R pair of primers was used (Figure 18) (Table 10 for

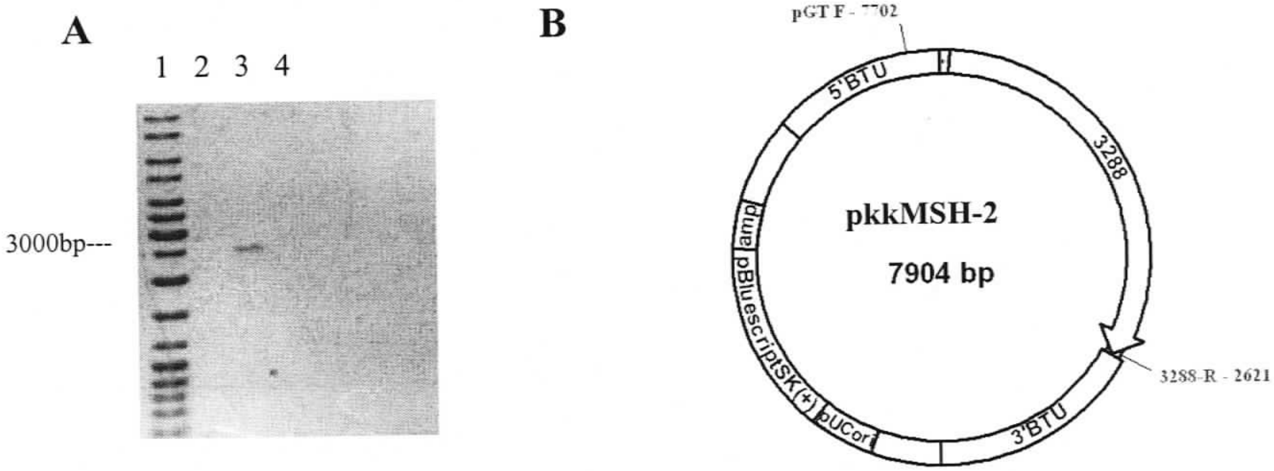


Figure 18: Diagnostic PCR for the integration of pKK3288 with pGTF and 3288-R primers. A) Lane 1: MW DNA ladder, Lane 2: untransformed *T. thermophila* sample , Lane 3 and 4: *T. thermophila* transformants samples 1 and 2 respectively. Run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min. B) pKK3288 vector map indicating the position of primers pGTF and 3288-R.

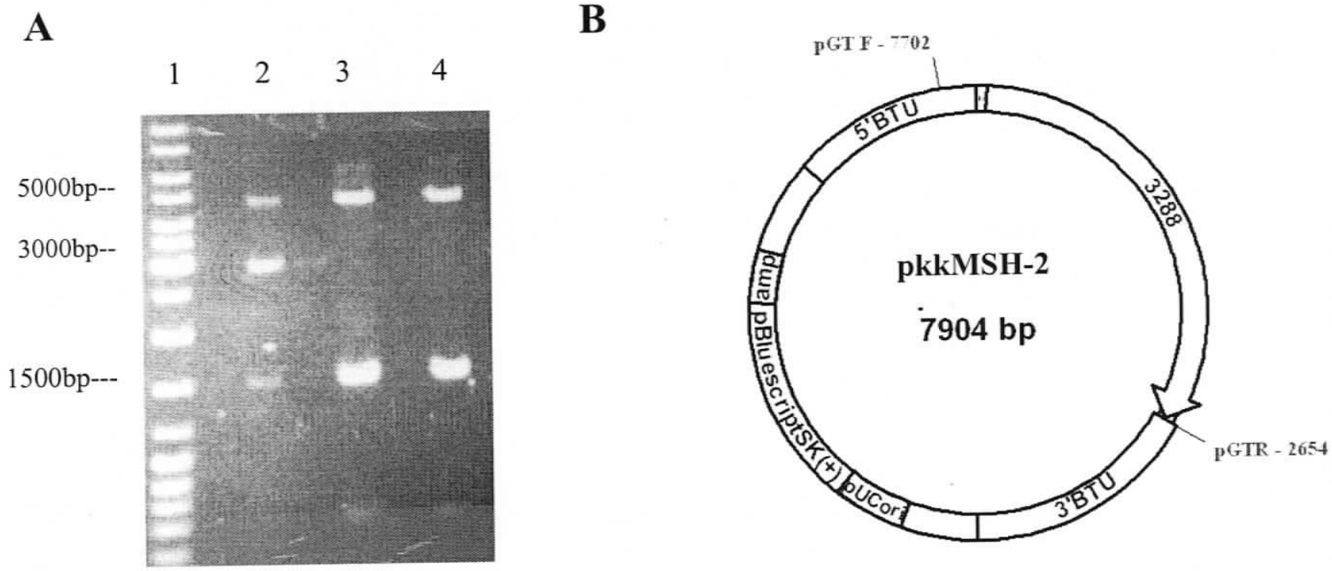


Figure 19: Diagnostic PCR for the integration of pkkMSH-2 with pGTF and pGTR primers. A) Lane 1: MW DNA ladder, Lane 2: *T. thermophila* sample 1 transformed with pkkMSH-2 , Lane 3 and 4 : *T. thermophila* transformant sample 2 and 3. Run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min. B) pkkMSH-2 vector map indicating the position of primers pGTF and pGTR.

sequence). A band between 2500bp and 3000bp was obtained (lane 3 Figure 18) corresponding with the predicted sizes for this transformant. Same transformant samples were analyzed but with a different set of primers (Figure 19): pGTF and pGTR both primers amplify outside the MSH2 gene in the β-tubulin *BTU1* non-coding region. Lane 2 has a band corresponding to MSH2 gene while lane 3 and 4 do not. There is an additional band visible at approximately 1500 bp and a non-specific band at 5000 bp in all of the samples analyzed in Figure 19.

3.7 SDS-PAGE and Immunoblots

T. thermophila were grown to 1-5 x 10^5 cells/ml. 15 ml of cells were spun and washed and resuspened in the final volume of 125 μl of 2x SDS-PAGE sample buffer. 10 μl of this preparation was loaded on the 10% polyacrylamide gel. Figure 20 shows a typical *T. thermophila* SDS-PAGE stained with Coomassie. Large molecular bands are visible, which is essential because all the *T. thermophila* MMR proteins are of high molecular weight.

Immunoblot analysis was carried out in order to visualize the recombinant hexa-histadine tagged proteins. Figure 21 shows a typical *T. thermophila* Western blot visualized by the fluor base detection system (Licor Odyssey). The positive control is visible in lane 2, while *T. thermophila* transformed cell with pKKTDG do not show a signal at expected molecular weight.

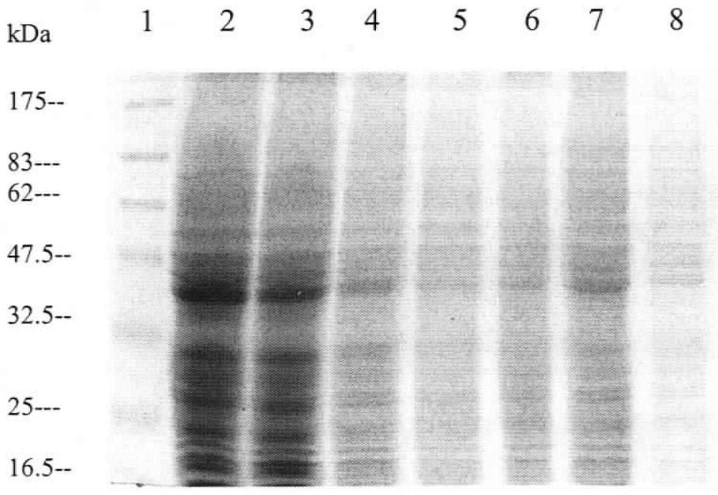


Figure 20: *T. thermophila* SDS-PAGE analysis. Lane 1: Prestained protein marker, broad range (NEB), Lane 2: 10 µl *T. thermophila* transformed with pkk3288, Lane 3 and 4: 10 µl of *T. thermophila* transformed with CH504, Lane 5,6 and 7: 10 µl of *T. thermophila* transformed with pkkTDG, Lane 8: untransformed *T. thermophila.* The proteins were separated on 10% acrylamide gel run at 15mA for 1 hour and visualized by Coomassie staining.

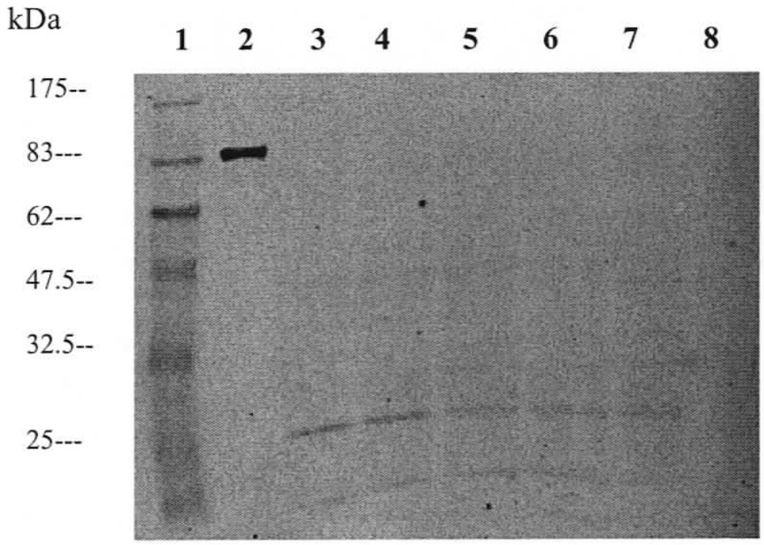


Figure 21: Western analysis of Tetrahymena transformed with pKKTDG. Lane 1: Prestained protein marker, broad range (NEB), Lane 2: positive control *E. coli* His-tagged MutS protein, Lane 3 to 7: *T. thermophila* transformed with pKKTDG, Lane 8: non-transformed *T. thermophila*. The proteins were separated on 10% acrylamide gel for 1 hour at 15mA and then electrophoretically transferred onto PVDF membrane at 28 volts for 1 hour, then at 84 volts for 14-16 hours at 4°C. The primary antibody used was 1: 3 000 dilution of the primary anti-his antibody. The secondary antibody on the membrane, IRDye800 conjugated affinity purified anti-mouse IgH was visualized by fluor base detection system (Licor Odyssey).

3.8 Gateway® Cloning and rDNA Vectors

Gateway$^{®}$ cloning technology provided a versatile system for transferring MMR genes between vectors. It greatly facilitated cloning of PCR products by replacing restriction endonucleases and ligases with site-specific recombination.

In more detail, two reactions constitute the Gateway$^{®}$ cloning system. The LR reaction is the main pathway of this system. The LR reaction is a recombination reaction between an entry clone and a destination vector, mediated by the LR clonase mix of recombination proteins. This reaction transfers DNA segments in the entry clone to the destination vector, to create an expression clone (Figure 22). The att sites are recognition sites for enzyme mediated recombination by bacteriophage lambda in *E. coli.* These sites are specifically recognized by the recombination proteins that constitute the clonase enzyme mix cocktails. The Gateway$^{®}$ cloning reactions are equivalent to concerted, highly specific, cutting and ligation reactions. The orientation of the gene is maintained throughout the subcloning, because attL1 reacts only with attR1, and attL2 reacts only with attR2. The desired transformants can be selected by kanamycin resistance. The unreacted destination vector does not produce kanamycin-resistant colonies, even though it carries the kanamycin-resistance gene, because it contains a gene lethal to *E. coli, ccd*B amd thus it never produces a colony.

The second major pathway of the Gateway$^{®}$ cloning system is the BP reaction. This is essentially the reverse of the LR reaction, the BP reaction transfers the expression clone into the donor clone to produce a new entry clone. The major use of the BP reaction is for cloning PCR products as entry clones. PCR products made with primers containing terminal attB sites (25nt + 4Gs) are efficient substrates for the BP reaction (Gateway Manual).

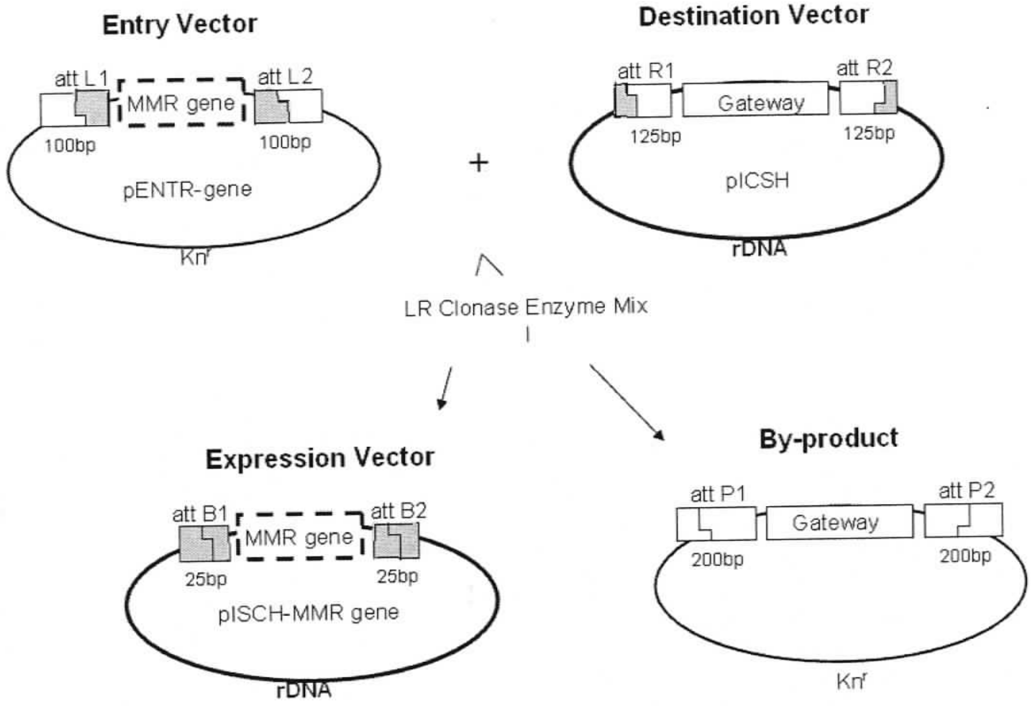


Figure 22: Gateway® cloning *T. thermophila* MMR genes. The LR reaction is a recombination reaction between and entry clone and a destination vector, mediated by the LR clonase mix of recombination proteins. The reaction transfers the *T. thermophila* MMR gene in the entry clone to the destination vector, to create an expression vector.

3.8.1 Gateway® cloning Tetrahymena MMR genes into rDNA vectors

T. thermophila MMR gene sequences were converted into entry clones N-MSH6-1, N-MSH6-2, N-MSH6-3, N-MLH-1, N-MSH2, N-PMS2 and then were subcloned into a pICSH destination vector (Figure 22), maintaining reading frame and orientation. This created the final six *T. thermophila* expression vectors: pICSH-MSH6-1, pICSH-MSH6-2, pICSH-MSH6-3, pICSH -MLH1, pICSH-MSH2 and pICSH-PMS2 (Figures 23,24 and 26). These final six expression vectors were digested with appropriate endonuclease enzymes (Table 9) (Figures 25 and 27) to verify a successful recombination of the insert (MMR gene) with the att sites of the destination vector pICSH (Figure 22). The endonuclease enzymes were picked for their ability to cut within the actual insert; for example XmnI sites are present within MSH6-1, MSH6-2, MSH6-3 and MLH1 genes (Figures 23 and 24). MSH2 gne has a NsiI recognition sequence and PMS2 has a NdeI recognition sequece (Figure 26). All of the six final expression vectors were positively identified by the diagnostic restriction endonuclease digestion.

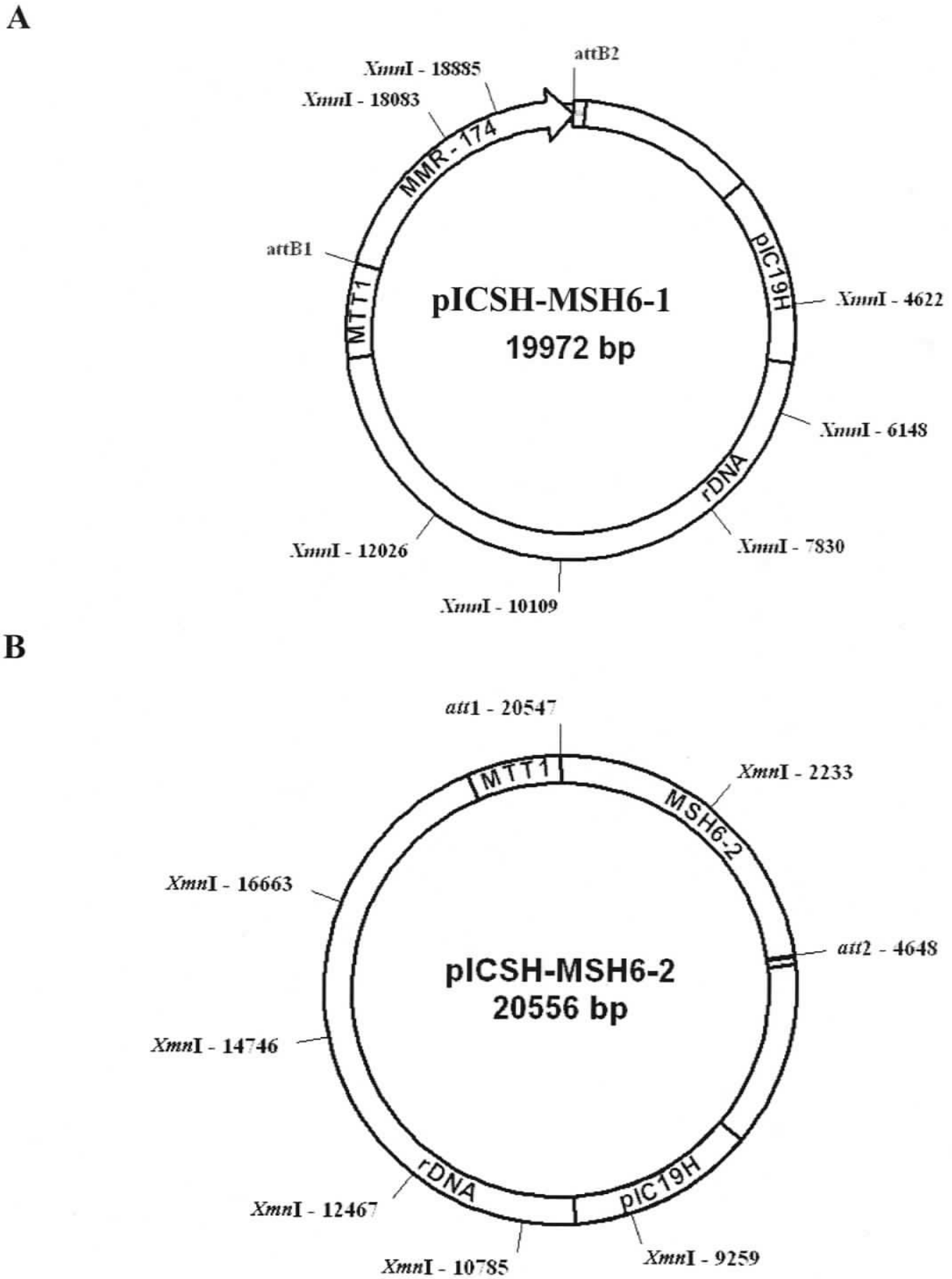


Figure 23: Vector maps of pICSH-MSH6-1 and pICSH-MSH6-2 with XmnI sites.

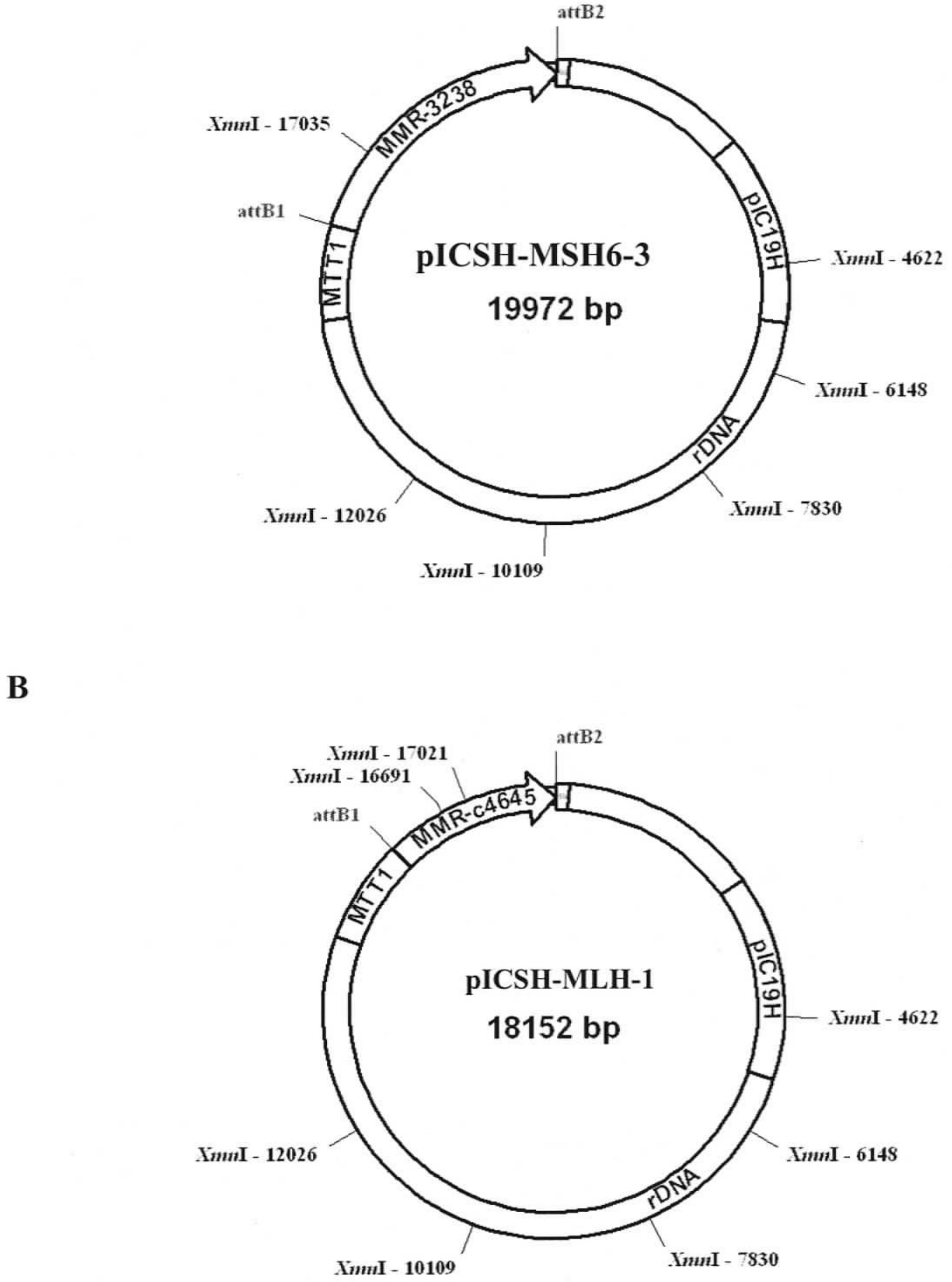
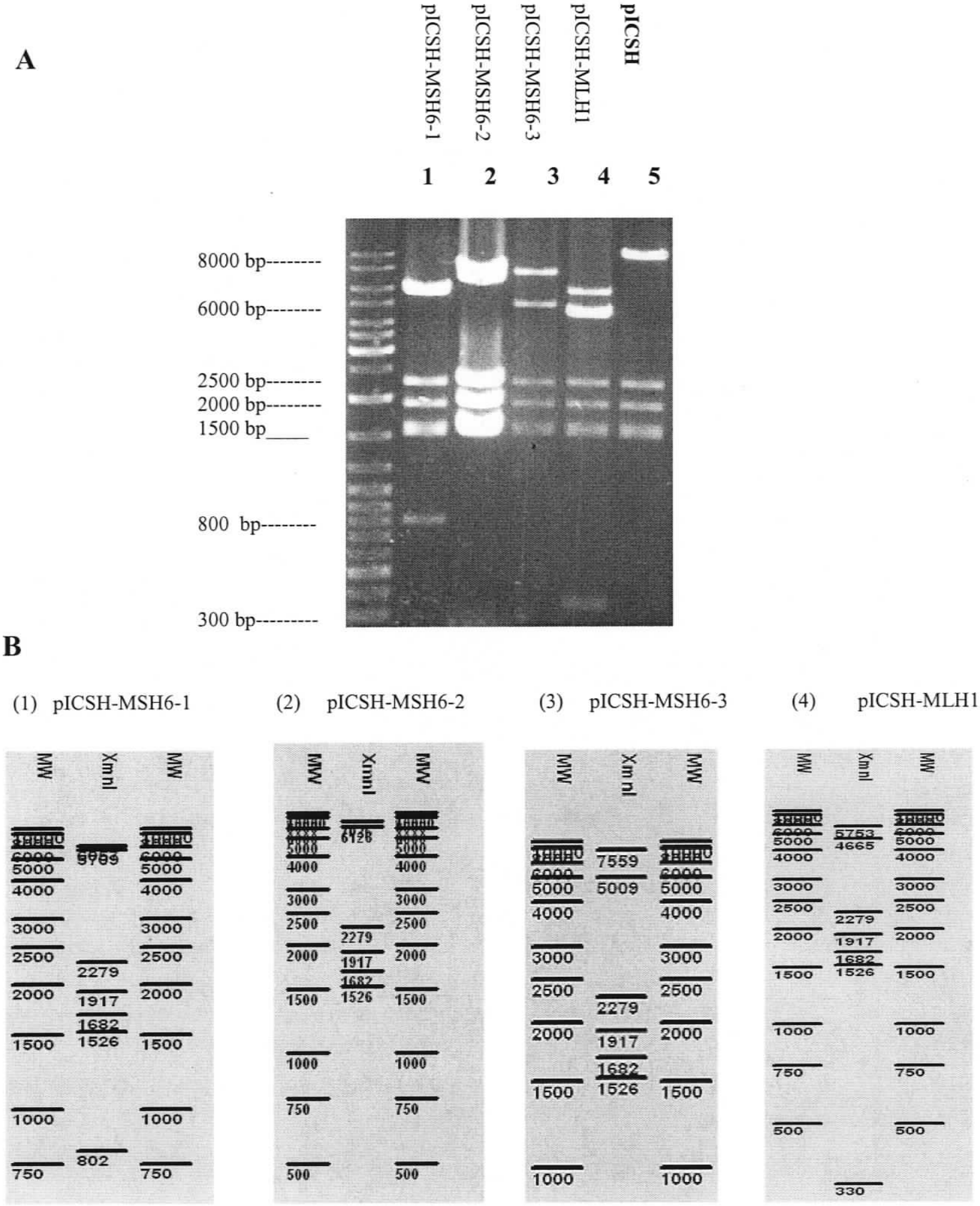


Figure 24: Vector maps of pICSH-MSH6-3 and pICSH-MLH1 with XmnI sites.



(1) pICSH-MSH6-1 (2) pICSH-MSH6-2 (3) pICSH-MSH6-3 (4) pICSH-MLH1

Figure 25: Diagnostic restriction endonuclease digest of pICSH-MSH6-1, pICSH-MSH6-2, pICSH-MSH6-3 , pICSH-MLH1 and pICSH with XmnI. A. Gel run on 1% agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min. **B**. Predicted fragments for pICSH-MSH6-1, pICSH-MSH6-2, pICSH-MSH6-3 and pICSH-MLH1 as digested with XmnI.

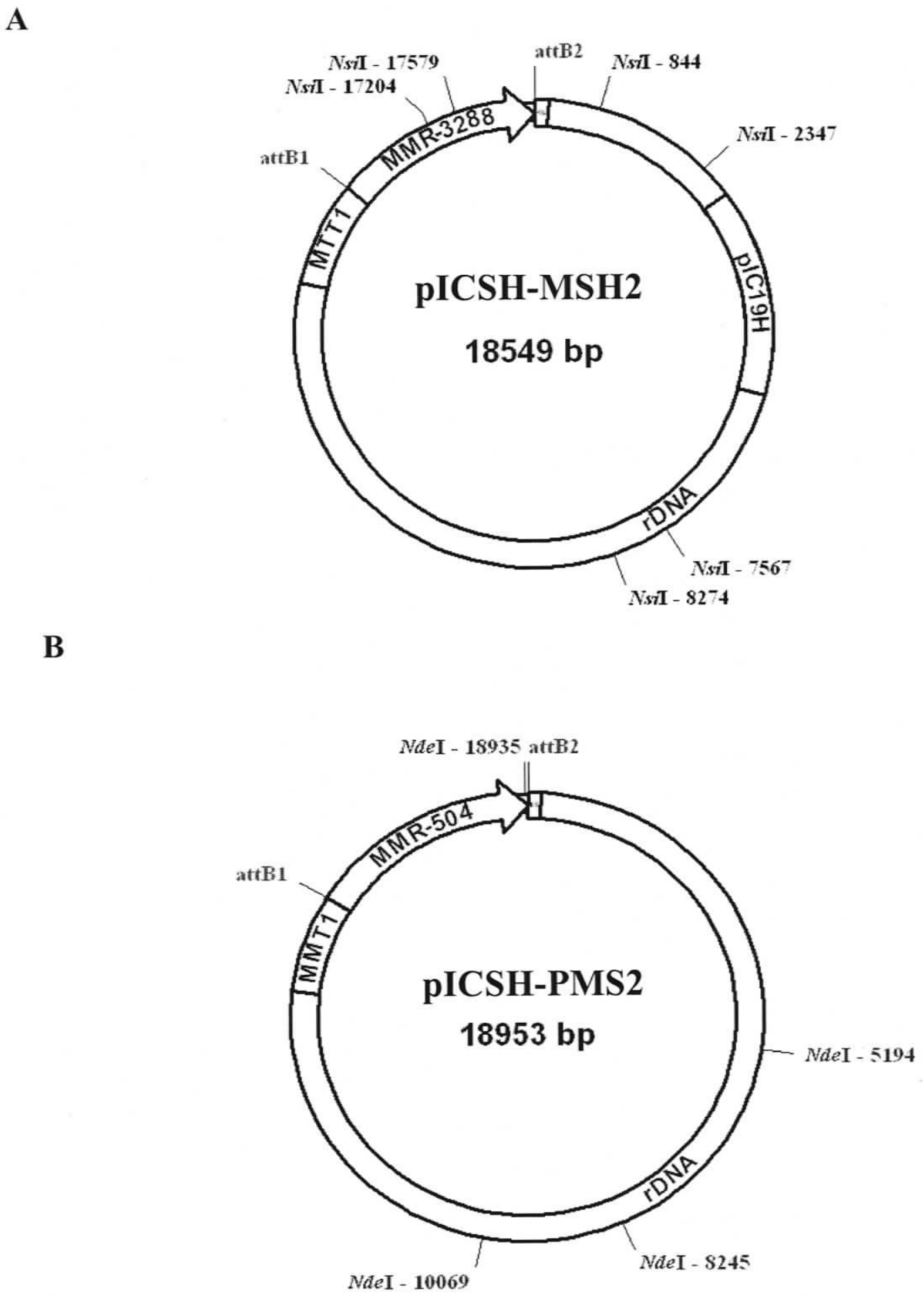


Figure 26: Vector maps of pICSH-MSH2 and pICSH-PMS2 **A.** vector map of pICSH-MSH2 with NsiI sites. **B.** Vector map of pICSH-PMS2 with NdeI sites.

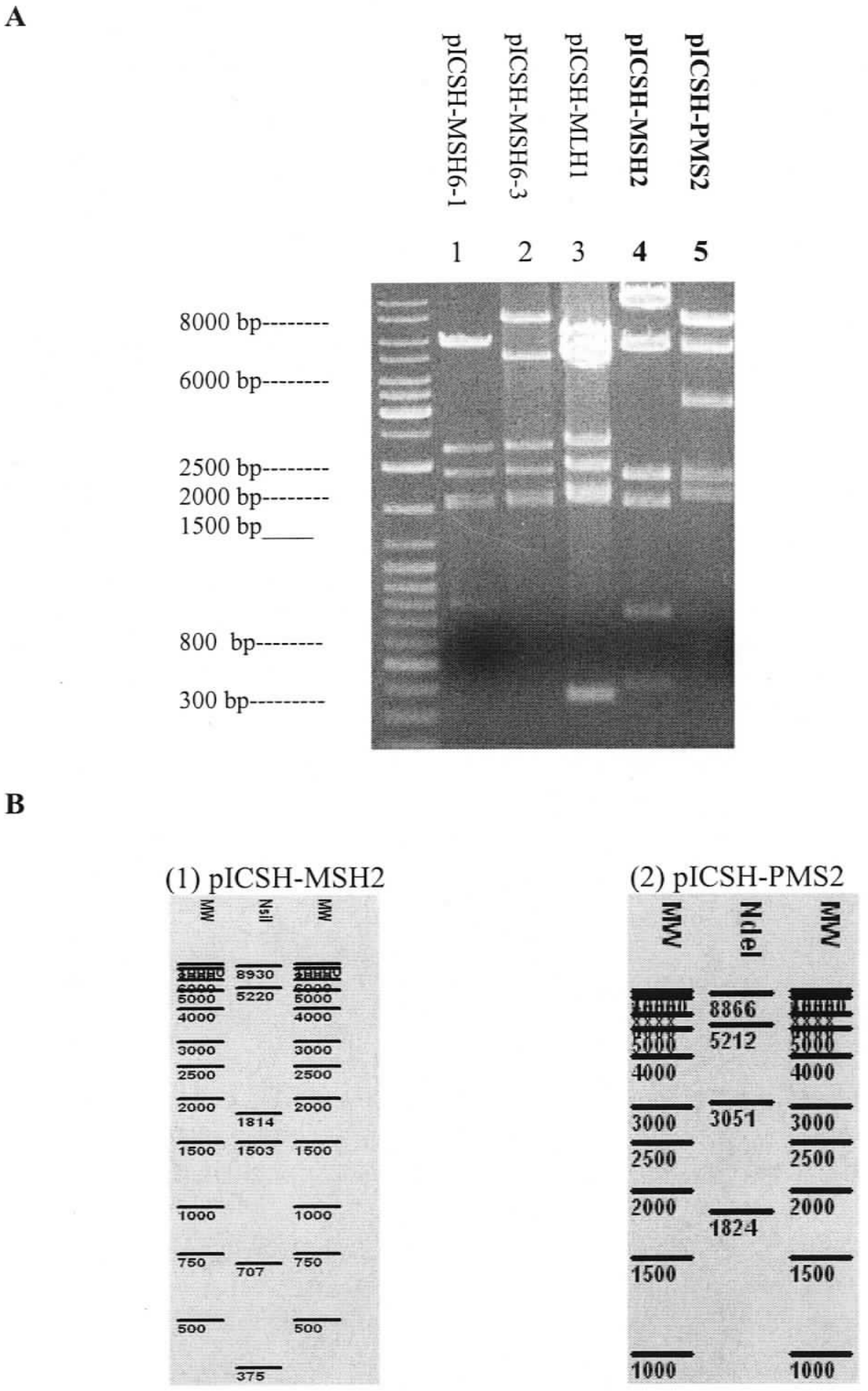


Figure 27: Diagnostic restriction endonuclease digest of pICSH-MSH2 and pICSH-PMS2. **A.**Lanes 1-3 are a repeat of the digests form figure 25, where lane 1: pICSH-MSH6-1 digested with XmnI, lane 2: pICSH-MSH6-3 digested with XmnI, lane 3: pICSH-MLH1 digested with XmnI, lane 4: pICSH-MSH2 digested with NsiI and lane 4: pICSH-PMS2 digested with NdeI. Gel run on1% Agarose gel at 100V in 1xTAE buffer and stained with EtBr for 20 min.

B. Predicted fragments for pICSH-MSH2 digest with NsiI and pICSH-PMS2 digested with NdeI.

Chapter 4: Discussion

4.1 Introduction and Objectives

Tetrahymena thermophila is one of the best characterized unicellular eukaryotes with its genome recently sequenced in its entirety (Eisen *et al.,* 2007). In addition, the nuclear dimorphism (MIC and MAC) of the ciliates offers the possibility of manipulating the organism's properties (Cassidy-Hanley *et al.,* 1997). The transcriptionally active MAC performs the somatic functions of the cell and the trascriptionally silent MIC serves as a germinal nucleus during mating. Due to the fact that the MAC undergoes an amitotic type of fission during cell division, it uses biochemical machinery which is distinct from the mitotically dividing MIC (Sweet *et al.,* 1997). Moreover, Tetrahymena cells are capable of surviving extreme damage to the MIC up to and including complete loss of both diploid copies of one or more micronuclear chromosomes (Burns *et al.,* 1981). Thus, one may generate viable cell lines which are homozygous for nonconditional mutations which interfere with meiotic and mitotic cellular functions, mutations which would prove lethal in almost any other eukaryotic system. In addition, the ability to induce and visualize the nuclear events associated with sexual behavior in these cells provides an excellent assay for screening and characterizing novel conjugation mutations (Cole *et al.,* 1997).

Within the last decade, molecular biological techniques have been developed to alter *T. thermophila's* genome and proteome, such as DNA transformation methods ranging from microinjection and electroporation into the MAC to biolistic bombardment of MIC and MAC. Episomal plasmids based on an rDNA-replicon have been developed, as well as knock out/in techniques based on homologous recombination. Furthermore, *T.*

thermophila's biotechnological potential has been demonstrated by its fast growth, high biomass, ability to ferment in ordinary bacteria/yeast equipment, up-scalability and existence of cheap, chemically defined media (Collins *et al.,* 2005). So far, ciliate expression vectors rely on large double rDNA origin stretches, to ensure stable propagation in *T. thermophila* cells, or on large flanking integrations sites of non-coding regions that are necessary for a proper and efficient homologous recombination into the gene loci of the MIC or MAC. In both cases the AT-richness and the unusual codon usage of these functional DNA sequences cause problems in handling and cloning *T. thermophila* in *E. coli* (Weide *et al.,* 2007).

Therefore, the objectives of this study were to: a) amplify MMR genes from *T. thermophila*, b) clone *Tetrahymena* MMR genes into suitable *T. thermophila* expression vectors, c) transform *T. thermophila* with the various constructs and d) detect recombinant MMR proteins from *T. thermophila.*

4.2 Homologously Recombinational Vector pKK101

The pKK101 vector was made as a derivative of the original pBICH3 vector. Upon transformation into *Tetrahymena,* pKK101 recombines homologously into the β-tubulin region due to the large flanking integration sites of non-coding regions of β-tubulin within the vector, thereby placing the gene of interest under the control of the endoginous β-tubulin promoter.

The HisTag:Thr:MCS cassette was cloned in order to facilitate cloning of *T. thermophila* MMR genes and to create recombinant MMR and histidine tagged fusion proteins for further expression and purification. *T. thermophila* MMR genes are relatively large is size ranging from 2.5 kb to 4.6 kb (Table 5). The original vector pBICH3 did not posses adequate restriction sites for cloning in the genes thus the multiple cloning site with unique cut sites was created. To sum up, pKK101 was designed to create fusion proteins and to streamline the cloning of *T. thermophila* MMR genes primarily using two enzymes NcoI and BamHI.

4.3 Amplification of the *T. thermophila* Mismatch Repair Genes

PCR products were not obtained from *T. thermophila* genomic DNA directly with the set of primers designed to amplify at the start and end of the genes. This was most likely due to the very high AT richness of the primers (an inherent property of the *T. thermophila* genome) as well as the size of the genes involved (2.5 kb to 4.6 kb). To overcome this problem nested PCR was designed involving two sets of primers, an "outer" and "inner" primer set. Once the larger PCR product was obtained containing the gene of interest and the flanking regions with the “outer” primers then the actual MMR gene was amplified with the specific set of "inner" primers containing restriction sites on the 5’ and 3’ ends to facilitate cloning.

4.4 Cloning *T. thermophila* MMR Genes into pKK101

Once the *T. thermophila* MMR genes were obtained, the products were gel purified in order to remove primer dimers and/or nonspecific bands, followed by digestion with appropriate restriction enzymes in order to create sticky ends for ligation into the receiving vector. To ensure that the sticky ends were truly created in the inserts to be cloned in, pDrive PCR cloning was utilized, thus making the cloning much more successful.

It was found that DNA rearrangement took place upon cloning *T. thermophila* genes and transforming them into competent *E. coli* DH5α cells or Top 10 cells . Thus, a new cell line of *E. coli* SURE ® was used. The *E. coli* SURE ® are designed for hard to clone genes. Upon utilizing the pDrive cloning intermediate step and transforming the ligation reactions into the *E. coli* SURE ® strain, pKKMSH2 was created.

Another approach in order to facilitate cloning is to amplify genes from cDNA in order to reduce their large size.

4.5 Gateway® Cloning *T. thermophila* MMR Genes

Gateway® cloning technology provided a versatile system for transferring DNA segments between vectors greatly faciliting cloning of PCR products by replacing restriction endonucleases and ligase with site-specific recombination. Once gene sequences are converted to entry clones, they can be subcloned into any type of destination vector, maintaining reading frame and orientation intact. This is a very useful approach considering the AT-richness of the *T. thermophila* genome which causes frequent problems in cloning and expression of the ciliate DNA. In order to further our insight into functional aspects of the *T. thermophila* genome, molecular biology tools are necessary that allow easy handling of *T. thermophila* genes to form the basis of the postgenomic age of this model organism. From the technical point of view this technique allows the flexible and fast transfer of DNA sequences from donor plasmids into multiple acceptor vectors thereby circumventing restriction and ligation reactions. Thus, once inserted into donor plasmid the selected DNA does not need to be subcloned. This is of high significance and relevance when cloning very large and AT-rich sequences such as the MMR genes of *T. thermophila*.

4.6 *T. thermophila* rDNA Vectors

The rDNA vectors created in this study pICSH-PMS2, pICSH-MSH2, pICSH-MLH1, pICSH-MSH6-3, pICSH-MSH6-2 and pICSH-MSH6-1 contain *T. thermophila* MMR genes. The pICSH vector used in this study is based on large *T. thermophila* double rDNA origin stretches thereby ensuring a stable propagation in *T. thermophila* cells upon transformation. The rRNA genes (rDNA) in *T. thermophila* have a unique organization. The organization is related to the nuclear dimorphism of ciliated protozoa, where the new MAC and MIC are formed during conjugation from the diploid micronucleus. The MAC development involves fragmentation, rearrangement, partial elimination, and endoreplication of the zygotic genome. During this macronuclear development, the single-copy micronuclear 21-kb rDNA is amplified into 9,000 extrachromosomal, palindromic dimers, compered with other chromosomes which are maintained at an average of 45 copies per MAC. The rDNA plasmids take advantage of this vast amplification of the rDNA genes during *anlagen*/MAC development (Pan *et al.,* 1995). Thus, it is important that the rDNA based plasmids are transformed during the *anlagen*/MAC development so the mass amplification of this plasmid can take place. The rDNA contains a single base substitution that confers resistance to paromomycin, thereby allowing phenotypic selection of transformed cells. The episomal retention of these plasmids depends on the presence of antibiotics in the culture medium. The disadvantage to these plasmids is that they may recombine homologously and non-directionally into the *T. thermophila* endogenous rDNA (Weide *et al.,* 2007).

4.7 Tetrahymena DNA Transformations

There is several reliable DNA transformation methods developed for *Tetrahymena thermophila*. They include: microinjection, conjugant electroporation (CET), and biolistic bombardment. (Tondravi and Yao, 1986; Gaertig and Gorovsky, 1992; Cassidy-Hanley, 1997). Both micronuclear and macronuclear DNA transformation are now being achieved (Hai and Gorovsky, 1997). Furthermore, gene knockouts can be generated in both types of nuclei by homologous recombination.

Conjugant electrotransformation (CET) and biolistic bombardment are preferred over microinjection because they are less labor-intensive and produce higher transformation frequencies; both approaches were taken in this study.

4.7.1 Biolistic Bombardment of *T. thermophila*

Biolistic bombardment can be used for efficient transformation of vegetative cells. In this study it was used for transforming *T. thermophila* with pKKMSH2 and pKKTDG. The mutant strain CU522 was used as a transformation host for biolistic transformations. It contains a single substitution (Lys350Met) in the β-tubulin *BTU1* gene, conferring increased resistance to several microtubule depolymerizing drugs such as oryzalin and colchicines, and increased sensitivity to a microtubule-stabilizing agent, paclitaxel (Gaertig *et al.,* 1999). *T. thermophila* has two BTU genes designed BTU1 and BTU2. By knocking out the BTU1 harboring the mutation in the β-tubulin BTU1 gene one can restore the cells resistance to paclitaxol in the presence of wild-type β-tubulin gene, BTU2. Thus, paclitaxol selection was used for the selection of *T. thermophila* transformants. Paclitaxol selection proved to be challenging and not very reliable, thus may not be the most efficient way.

Biolistic transformation is an attractive way of transforming *T. thermophila* cells as opposed to CET method as vegetative cells can be used as opposed to mating cells which are required for CET method, making the biolistic method quicker and easier technically. The biolistic method can not be used for the rDNA type of vectors as they require the amplification step in MAC development which takes place approximately 10 hours into mating. Furthermore, for future studies the biolistic method has a major advantage of delivering the DNA into either the macronucleus or the micronucleus, the later leading to heritable germline transformations.

4.7.3 Conjugant Electrotransformation (CET) *T. thermophila*

We used two highly fertile strains with different mating types, such as the heterokaryon strain CU428 [*mpr1-1/mpr1-1* (mp-s, VII)] and the wild-type strain B2086 (II). The conjugant electrotransformation (CET) was carried out for the transformation of rDNA based vectors. It was found that conjugating *Tetrahymena* cells display a period of competence for electrotransformation which corresponds to early stage of macronuclear development (10-11 h after mixing at 30°C) (Gaertig and Gorovsky, 1992). Conjugant electrotransformation delivers DNA to the newly developing macronuclei during the *anlagen*/MAC development where the rDNA is amplified to 9,000 copies. The idea here is that the recombinant rDNA vectors will "swamp out" the contribution of an endogenous rDNA gene with the addition of the drug. The efficiency of transformation is 1000 transformants/μg of replicative vector and 50 transformants/μg of integrative vector (Gaertig et al., 1994a). Selection of transformants relies on a dominant paromomycin mutation in the 17S rRNA sequence. Thus, conjugating cells offer an advantage for transformation because the young developing macronuclei (anlagen) contains only a few copies of endogenous genes, including the rRNA. An exogenous sequence introduced at this stage more easily reaches a selectable level resulting from the amplification of the rDNA. So far, in this study the CET transformants have not been obtained. This is most likely a technical difficultly, one way to improve on the technique is not to rely on time alone for the development of anlagen which approximately takes place 10-11 hours after mixing of the two different mating types (Gaertig and Gorovsky, 1992) but to make a visual check under the florescent microscope with DAPI stained cells for the visualization of the anlagen. This may prove to be a more reliable method as the conjugation stages and time line of *T. thermophila* may vary depending on the strain used as well as conditions.

Since the writing of this thesis the CET transformation of Tetrahymena have been sucessful with the modified method above for the outlined rDNA vectors created in this study.

4.8 Expression of Human TDG Protein in *T. thermophila*

Initial cloning of thymine DNA gylcosylase was done as a control for cloning into the pKK101 vector due to problems which arose with cloning the *T. thermophila* genes. The cloning of TDG into pKK101 was successful as well as the biolistic transformation with the linearized pKKTDG vector. The expression of eukaryotic genes in heterologous systems has become increasingly important both for the production of pharmaceuticals and as a basic research tool (Gaertig *et al.,* 1999). It has recently been shown that small ubiquitin-related modifier (SUMO) is upregulated during sexual reproduction in *Paramecium* and *Tetrahymena* (Matsuda *et al.*, 2006). This is of interest because it has been shown previously by a yeast two-hybrid screening that the human SUMO-1 and SUMO-3 interact with human TDG (Hardeland *et al.,* 2002). It would be of interest to see whether the same mechanism of posttranslational modification by SUMO (SUMOylation) functions in *T. thermophila*.

4.9 Overexpression of *T. thermophila* Proteins: SDS-PAGE and Immunoblot Analysis

The SDS-PAGE experiments were prepared in order to visualize *T. thermophila* proteins. At first it was technically challenging to see high molecular bands but this difficulty was overcome by protocol optimization and as a result high molecular bands were visible. High molecular bands are import in this study because all of the *T. thermophila* MMR proteins are above 80 kDa in size.

There were no visible bands on the immunoblots with an anti-his antibody for *T. thermophila* extracts transformed with pKKMSH2 which should have produced a protein of approximately 93kDa, nor for pKKTDG (figure 27) transformants. A recombinant *E.coli* MutS was used as a control. It is a fusion protein with a histidine tag which as seen in Figure 22 and is detected by an anti His antibody.

The expression of proteins under the β-tubulin promoter by a culture of standard density at 10^6 cell/ml is capable of expressing up to 0.269 mg/L. The *BTU1* mRNA is highly abundant and β-tubulin makes up about 2-3% of total *Tetrahymena* cell protein. The other promoter used in this study is the MTT1 promoter; it can be induced by heavy metals such as zinc, copper, and cadmium (Santovito *et a.l,* 2007), and has been used successfully to regulate expression of genes in *T. thermophila* and in other systems (Shang *et al.,* 2002). When the two promoters were studied together quantitative analysis revealed that the MTT1 promoter is 18-30 times stronger that the BTU1 promoter (Shang *et al.,* 2002). The MTT1 promoter is now the promoter of choice for overexpression of genes in Tetrahymena.

CHAPTER 5: CONCLUSIONS AND FUTURE INVESTIGATIONS

5.1 Conclusions

Six of *T. thermophila* MMR genes : MLH1, PMS2, MSH2 and three MSH6 homologous were amplified as nested products. Two different cloning strategies were undertaken: a) using a homologous recombination vector to recombine into the β-tubulin loci of *T. thermophila* genome; this was used to obtain pKKMSH2 and pKKTDG, b) using Gateway® technology *Tetrahymena* MMR genes were cloned into rDNA type vector- pICSH creating the following constructs: pICSH-PMS2, pICSH-MSH2, pICSH-MLH1, pICSH-MSH6-3, pICSH-MSH6-2 and pICSH-MSH6-1; these vectors are episomal and the genes are placed under the MMT1 inducible promoter. Furthermore, two different DNA transformations of *T. thermophila* were carried out depending on the type of construct. For the homologous recombination constructs biolistic transformation was performed followed by paclitaxol selection into a mutant strain of *T. thermophila* carrying a negative selectable allele of a β-tubulin gene. The transformants attained were pKKMSH2 but later on lost most likely due to poor selection exhibited by paclitaxol. Conjugant electrotransformation (CET) was the second type of transformation performed on rDNA vectors, no positive transformants have been identified as of yet. Finally immunoblots were carried out on *T. thermophila* samples transformed with of pKKMSH2 and pKKTDG but proteins were not detected.

5.2 Future Investigations

Further work is needed. The transformations of *T. thermophila* with rDNA vectors have been successful since the writing of this thesis. Next, stable transformants need to be selected. Once the selection has been achieved the his-tagged MMR proteins maybe visualized by immunobloting prior to purification. Another way of detecting the recombinant MMR proteins in Tetrahymena is by detecting elevated mRNA levels in the cadmium induced cells since the his-tagged MMR proteins are placed under the control of MTT1promoter in the rDNA vector. Once the MMR proteins are obtained structural and functional studies may proceed, such as the generation of antibodies against *T. thermophila* MutL and MutS homologs. The specific antibodies would be of great benefit for microscopic localizations studies in *T. thermophila.* Furthermore, activity and substrate specificity assay for MutS, MutL homologous and TDG may be investigated. Some of the possible assays to confirm protein activity may include: gel shift and super shift assays with LICOR labeled DNA heteroduplexes (already prepared), these assays are of help in indentifying protein partners as eukaryotic MMR proteins function as heterodimers as well as aid in the identification of their specificities.

Literature Cited

Abuin A, **Zhang H, Bradley A** (2000) Genetic analysis of mouse embryonic stem cells bearing Msh3 and Msh2 single and compound mutations. *Mol Cell Biol* **20**:149-57.

Acharya S, Wilson T, Gradia S, Kane MF, **Guerrette S, Marsischky GT, Kolodner R**, **Fishel R** (1996) hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci USA* **93**:13629-34.

Baitinger C, Burdett V, Modrich P (2003) Hydrolytically deficient MutS E694A is defective in the MutL-dependent activation of MutH and in the mismatch-dependent assembly of the MutS.MutL.heteroduplex complex. *J Biol Chem* **278**:49505-11.

Bjornson KP, Blackwell LJ, Sage H, Baitinger C, Allen D, Modrich P (2003) Assembly and molecular activities of the MutS tetramer. *J Biol Chem* **278**:34667-73.

Borts RH, Chambers SR, Abdullah MF (2000) The many faces of mismatch repair in meiosis. *Mutat Res* **451**:129−150.

Boyer JC, **Umar A, Risinger JI** (1995) Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res* **55**:6063-6070.

Burdett V , Baitinger C, Viswanathan M, Lovett ST, Modrich P (2001) *In vivo* requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc Natl Acad Sci U S A* **98**:6765–6770.

Bruns PJ, Brussard TB (1981) Nullisomic *Tetrahymena. Science* **213**:549–551.

Carraway M, Marinus MG (1993) Repair of heteroduplex DNA molecules with multibase loops in *Escherichia coli. J Bacteriol* **175**:3972-3980.

Cassidy-Hanley D, Bowen J, Lee JH, Cole E, VerPlank LA, Gaertig J, Gorovsky MA, Bruns PJ. (1997) Germline and somatic transformation of mating *Tetrahymena thermophila* by particle bombardment. *Genetics* **146**:135–147.

Cole ES, Cassidy-Hanley D, Hemish J, Tuan J, Bruns PJ (1997) A Mutational Analysis of Conjugation in*Tetrahymena thermophila. Dev Biol* 189:215-32.

Collins K, Gorovsky MA. (2005) *Tetrahymena thermophila*. *Curr Biol* **15**:R317–R318.

Dutta R, Inouye M (2000) GHKL, an Emergent ATPase/kinase Superfamily. *Trends Biochem Sci* **25**:24–28.

Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, Tallon LJ, Delcher AL, Salzberg SL, Silva JC, Haas BJ, Majoros WH, Farzad M, Carlton JM, Smith RK Jr., Garg J, Pearlman RE, Karrer KM, Sun L, Manning G, Elde NC, Turkewitz AP, Asai DJ, Wilkes DE, Wang Y, Cai H, Collins K, Stewart BA, Lee SR, Wilamowska K, Weinberg Z, Ruzzo WL, Wloga D, Gaertig J, Frankel J, Tsao CC, Gorovsky MA, Keeling PJ, Waller RF, Patron NJ, Cherry JM, Stover NA, Krieger CJ, del TC, Ryder HF, Williamson SC, Barbeau RA, Hamilton EP, Orias E (2006) Macronuclear genome sequence of the ciliate *Tetrahymena thermophila,* a model eukaryote. *PLoS Biol* **4**:e286.

Fink D, Zheng H, Nebel S, Norris PS, Aebi S, Lin TP (1997) *In vitro* and *in vivo* resistance to cisplatin in cells that have lost DNA mismatch repair. *Cancer Res* **15**:1841–1845.

Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R (1993) The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**:1027-1038.

Fjerdingstad EJ, Nicolas Schtickzelle, Manhes P, Gutierrez A, Clobert J (2007) Evolution of dispersal and life history strategies – *Tetrahymena* ciliates. *BMC Evol Biol* **7**:133.

Friedberg EC, McDaniel LD, Schultz RA (2003) The role of endogenous and exogenous DNA damage and mutagenesis. *Curr Opin Genet Dev* **14**:5:10.

Gaertig J, Gorovsky MA (1992) Efficent mass transformation of Tetrahymena thermophila by electroporation of conjugants. *Proc Nat Acad Sci USA* **89**:9196-9200.

Gaertig J, Thatcher TH, Gu L , Gorovsky MA (1994 a) Electroporation-mediated replacement of a positively and negatively selectable beta-tubulin gene in *Tetrahymena thermophila*. *Proc Natl Acad Sci* **91**:4549–4553.

Gaertig J, Gao Y, Tishgarten T, Clark TG, Dickerson HW. (1999) Surface display of a parasite antigen in the ciliate *Tetrahymena thermophila*. *Nat Biotechnol* **17**:462–465.

Geisse S, Gram H, Kleuser B, Kocher HP. (1996) Eukaryotic expression systems: a comparison. *Protein Expr Purif* **8**:271-282.

Gorbunova V, Seluanov A, Mao Z, Hine C (2007) Changes in DNA repair during aging. *Nucleic Acid Res* 1-9.

Gracy RW, Talent JM, Kong Y,Conrad CC (1999) Reactive oxygen species: The unavoidable environmental insult? *Mutat Res* **428**:17-22.

Grilley M, Griffith J, Modrich P (1993) Bidirectional excision in methyl-directed mismatch repair. *J Biol Chem* **268**:11830-11837.

Guarne S, Ramon-Maiques EM, Wolff R, Ghirlando, Hu X, Miller JM, Yang W (2004) Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J* **23**:4134–4145.

Hanley-Cassidy D, Bowen J, Lee JH, Cole E, VerPlan LA, Gaertig J, Gorovsky MA, Burns PJ (1997) Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. *Genetics* **146:**135-147.

Hardeland U, Steinacher R, Jiricny J, Schär P (2002) Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *EMBO J* **2**:1456-64.

Harfe BD, Jinks-Robertson S (2000) Sequence composition and context effects on the generation and repair of frameshift intermediates in mononucleotide runs in *Saccharomyces cerevisiae. Genetics* **156**:571-8.

http://www.rndsystems.com/mini_review_detail_objectname_MR03_DNADamageResponse.aspx

Iyer RR, Pluciennik A, Burdett V, Modrich P (2006) DNA mismatch repair: functions and mechanisms. *Chem Rev* **106:**302–323.

Jaroudi S, SenGupta S (2006) DNA repair in mammalian embryos. *Mutat Res* **635**:53.

Jiricny J (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* **7:**335-346.

Jun SH, Kim TG, Ban C (2006) DNA mismatch repair system. Classical and fresh roles. *FEBS J* **27**:1609-19.

Junop MS, Yang W, Funchain P, Clendenin W, Miller JH (2003) In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst)* **2**:387-405.

Kadyrov FA, Dzantiev L, Constantin N, Modrich P (2006) Endonucleolytic function of MutLα in human mismatch repair. *Cell* **126**:297–308.

Karrer KM. (2000) Tetrahymena genetics: two nuclei are better than one. *Methods Cell Biol* **62**:127–186.

Kosinski J, Steindorf I, Bujnicki JM, Giron-Monzon L, Friedhoff P (2005) Analysis of the Quaternary Structure of the MutL C-terminal Domain. *J Mol Biol* **351:**895-909.

Kunkel KA and Erie DA (2005) DNA Mismatch Repair. *Annu Rev Biochem* **74**:681-710.

Lahue RS, Au KG, Modrich P (1989) DNA mismatch correction in a defined system. *Science* **245**:160–164.

Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* **407***:*711-717.

Loft S, Poulsen HE (1996) Cancer risk and oxidative DNA damage in man. *J Mol Med* **74**:297-312.

Lu AL, Clark S, Modrich P (1983) Methyl-directed repair of DNA base-pair mismatches in vitro. Proc Natl Acad Sci USA 80: 4639-43.

Matsuda A, Forney JD (2006) The SUMO pathway is developmentally regulated and required for programmed DNA elimination in *Paramecium tetraurelia. Eukaryot Cell* **5**:806-15.

McCulloch SD, Gu L, Li GM (2003) Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *J Biol Chem* **278**:3891-3896.

Mochizuki K, Gorovsky MA (2004a) Small RNAs in genome rearrangement in Tetrahymena. *Curr Opin Genet Dev* **14**:181–187.

Mochizuki K, Gorovsky MA (2004b) RNA Polymerase II Localizes in *Tetrahymena thermophila* Meiotic Micronuclei When Micronuclear Transcription Associated with Genome Rearrangement Occurs. *Eukaryot Cell* **3**:1233–1240.

Modrich P, Lahue R (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem* **65**: 101−133 Li GM (2003) DNA mismatch repair and cancer. *Front Biosci* **8**:997−1017.

Modrich P (2006) Mechanisms in eukaryotic mismatch repair. *J Biol Chem* **281:**30305-30309.

Natrajan G, Lamers MH., Enzlin JH., Winterwerp HH., Perrakis A, Sixma TK (2003) Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res* **31**:4814-21.

Obmolova G, Ban C, Hsieh P, Yang W (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407:**703-710.

Peltomaki P (2001) DNA mismatch repair and cancer. *Mutation Res* **488**:77.

Peltomaki P (2003) Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* **21**:1174–1179.

Peltomaki P (2005) Lynch syndrome genes. *Fam Cancer* **4**:227–232.

Peterson DS, Gao Y, Asokan K, Gaertig J (2002) The circumsporozoite protein of *Plasmodium falciparum* is expressed and localized to the cell surface in the free-living ciliate *Tetrahymena thermophila. Mol Biochem Parasitol* **122**:119–126.

Pinheiro M,Power M,Butler BJ, Dayeh V, Slawson R, Lee LJM, Lynn DH, Bols NC (2007) Use of *Tetrahymena thermophila* To Study the Role of Protozoa in Inactivation of Viruses in *Water Appl Environ Microbiol* **73**:643-9.

Power ME, Sotornik D, Pinheiro M, Dayeh V, Butler BJ, Slawson RM, Lee LJM, Lynn DJ, Bols NC (2006) Development of a Fluorescent Multiwell Assay for Evaluating the Capacity of the Ciliated Protozoan Tetrahymena for Bacterivory in Water Samples. *Water Quality Research Journal of Canada* **41**:307 – 315.

Reese JS, Liu L, Gerson SL (2003) Repopulating defect of mismatch repair-deficient hematopoietic stem cells. *Blood* **102**:1626-33.

Sambrook J *et al.* (2000) Molecular Cloning: A Laboratory Manual (Second Ed). *Cold Spring Harbor Laboratory Press.* Cold Spring Harbor, NY USA.

Santovito G, Formigari A, Boldrin F, Piccinni E (2007) Molecular and functional evolution of Tetrahymena metallothioneins: new insights into the gene family of *Tetrahymena thermophila. Comp Biochem Physiol C Toxicol Pharmacol* **44**:391-7.

Sauvant NP, Pepin D, Piccinni E (1999) *Tetrahymena pyriformis*: A tool for toxicological studies: A review. *Chemosphere* **38**:1631-69.

Schofield M.J. and P.Hsieh (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* **57**:579-608.

Shang Y, Song X, Bowen J, Corstanje R, Gao Y, Gaertig J, Gorovsky MA (2002) A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes *in Tetrahymena thermophila. Proc Natl Acad Sci U S A* **99**:3734-9.

Shimodaira H, Yoshioka-Yamashita A, Kolodner RD, Wang JY (2003) Interaction of mismatch repair protein PMS2 and the p53-related transcription factor p73 in apoptosis response to cisplatin. *Proc Natl Acad Sci U S A* **100:**2420–5.

Simon EM, Orias E(1987) Genetic Instability in the Mating Type System of *Tetrahymena pigmentosa. Genetics* **117**:437–449.

Skinner AM, Turker MS (2005) Oxidative mutagenesis, mismatch repair, and aging. *Sci Aging Knowledge Environ.*

Stojic L, Brun R, Jiricny J (2004) Mismatch repair and DNA damage signalling. *DNA Repair (Amst)* **3**:1091-101.

Sweet TM, Carlson G, Cook RG, Nelson D, Allis CD (1997) Phosphorylation of linker histones by a protein kinase A-like activity in mitotic nuclei. *J Biol Chem* **272**:916–923.

Tetrahymena Genome Database - TGB. 2007. http://www.ciliate.org

Turkewitz AP, Orias E, Kapler G (2002) Functional genomics: the coming of age for *Tetrahymena thermophila. Trends Genet* **18**:35–40.

Vaish M (2007) Mismatch repair deficiencies transforming stem cells into cancer stem cells and therapeutic implications. *Mol Cancer* **6**:26

Van Houten B, Eisen JA, Hanawalt PC (2002) A cut above: discovery of an alternative excision repair pathway in bacteria. *Proc Natl Acad Sci USA* **99**:2581.

Wagner R, Meselson M (1979) Repair tracts in mismatched DNA heteroduplexes Proc.Natl.Acad.Sci.USA **73**:4135-4139.

Wang J, Luo MH, Zhang Z, Zhang P, Jiang X, Ma D, Suo R, Zhao L, Qi Q (2007) Clinical and molecular analysis of hereditary non-polyposis colorectal cancer in Chinese colorectal cancer patients. *World J Gastroenterol 13*:1612-7.

Wei K, Kucherlapati R, Edelmann W (2002) Mouse models for human DNA mismatch-repair gene defects. *Trends Mol Med* **8**:346-53.

Weide T, Herrmann L, Bockau U, Niebur N, Aldag I, Laroy W, Contreras R, Tiedtke A, Hartmann MW (2006) Secretion of functional human enzymes by *Tetrahymena thermophila. BMC Biotechnol* **6**:19.

Weide T, Bockau U, Rave A, Herrmann L, Hartmann MWW (2007) A recombinase system facilitates cloning of expression cassettes in the ciliate *Tetrahymena thermophila. BMC Microbiol* **7:** 12.

Wheatley DN, Rasmussen L, Tiedtke A (1994) Tetrahymena: a model for growth, cell cycle and nutritional studies, with biotechnological potential. *Bioessays* **16**:367–372.

Wu J, Zhu B,Yu J, Zhu H, Qiu L, Kindy MS, Gu L,Seidel A, Li GM (2003) In vitro and in vivo modulations of benzo[c]phenanthrene-DNA adducts by DNA mismatch repair system. *Nucleic Acids Res* **31**:6428-34.

Yamaguchi M, Dao V, Modrich P (1998) MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J Biol Chem* **273**:9197-201.

Yang W (2006) Human MutLalpha: the jack of all trades in MMR is also an endonuclease. *DNA Repair (Amst)* **6**:135-9.

Yao MC, Fuller P, Xi X (2003) Programmed DNA deletion as an RNA-guided system of genome defense. *Science* **300**:1581–1584.

Yao MC, Chao JL (2005) RNA-guided DNA deletion in Tetrahymena: an RNAi-based mechanism for programmed genome rearrangements. *Annu Rev Genet* **39**:537–559.

Zaug AJ, Cech TR (1986) The intervening sequence RNA of Tetrahymena is an enzyme. Science **231**:470-475.