UNIVERSITY OF CALIFORNIA SAN DIEGO

Recruitment and Function of Mlh1-Pms1 in DNA Mismatch Repair

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Biomedical Sciences

by

Matthew DuPrie

Committee in Charge:

Professor Richard Kolodner, Chair
Professor Kevin Corbett
Professor Arshad Desai
Professor Andres Leschziner
Professor Lorraine Pillus

2020

The dissertation of Matthew DuPrie is approved, and it is in acceptable quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego

2020

DEDICATION

I dedicate this to my parents and family, and my wonderful girlfriend Kat. I dedicate this to all of those who've chosen different lives; I did this for you.

EPIGRAPH

THE STUDENT Tokusan used to come to the master Ryutan in the evenings to talk and to listen. One night it was very late before he was finished asking questions.

"Why don't you go to bed?" asked Ryutan.

Tokusan bowed, and lifted the screen to go out. "The hall is very dark," he said.

"Here, take this candle," said Ryutan, lighting one for the student.

Tokusan reached out his hand, and took the candle.

Ryutan leaned forward, and blew it out.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge first and foremost Richard Kolodner, for letting me in his lab and letting me draw from his deep well of experience and wisdom. Finding his laboratory is probably one of the most fortunate events in my life. His support of my research and personal growth is something I will always be indebted to, and I move forward with my career with the fondest memories.

I would like to thank Chris Putnam, a man of indefatigable enthusiasm and a source of brilliance and creativity. His conversations have made me smarter, and I am very happy to have met him.

I would like to thank Anjana Srivatsan, whom I sat next to in the lab for many years, and whose attention to details and logic may have rubbed off on me a little over the years, to my benefit. I appreciate being able to meet such a kind soul, and sincerely wish her well.

I would like to thank all of my lab mates from the Kolodner lab, both past and present: Cathy, Sarah, Sara, Nikki, Eric, Bill, Eva, Binzhong, Dafne, Elaine, Kat, Young Stephen, Computer Bill, Katie and others. I'd also like to thank all of my connections I've made at the Ludwig Institute for Cancer Research in San Diego, including members from the labs of Frank Furnari, Bing Ren, Huilin Zhou, Don Cleveland, Arshad Desai, Karen Oogema, Karen Arden, Kevin Corbett, and the small molecule group on the second floor. I'd like to further thank all of the connections I've made through graduate school, all of the other students I've come to know and have enlightened me and made me smarter. I'd like to thank my friends from college, who have kindly kept in touch and kept me inspired with goings on in the non-science world.

I would like to acknowledge my close friends in the graduate program, Danny, Tina, Navarre, Steve, and Sam, whom have all greatly impressed me, and I will be eternally grateful to have gotten the chance to know. By this point they have all graduated or have moved on, I hope to one day express how much their friendship meant to me.

I would like to acknowledge my family, Jeannine and Harold DuPrie, as well as my younger brother Patrick. I have truly had the great fortune to have been born into a family that is supportive of my ambitions and dreams, and have given me their unconditional love. I think of them when I think of why my work has meaning.

I would like to thank the Canyonview Pool, for being a place of solace and refuge, calmly swimming on beautiful days, on cloudy days, quiet days or days when the entire swim team was practicing. I spent many afternoons contemplating my own life, meaning, experiments, and its deep turquoise waters with dark green forest and electric blue sky will always remain for me a place of mental relaxation, even if we are currently barred from going there due to the virus, and I may never swim in it again.

The material in Chapter 2 is currently being prepared for submission for publication: DuPrie M., Hargreaves V.V., Calil F., Putnam C.D., Kolodner R.D. "Interaction between the shared subunits Mlh1 and Msh2 mediates recruitment between eukaryotic MutS and MutL homologs during DNA mismatch repair." The dissertation author was a primary researcher and author of this material.

VITA

2010 Bachelor of Science, University of Michigan, Ann Arbor

2020 Doctor of Philosophy, University of California, San Diego

Publications

"Downregulation of EZH2 decreases growth of estrogen receptor-negative invasive breast carcinoma and requires BRCA1" *Oncogene 2009 Feb 12;28(6):843-53.* PMID: 19079346. ME Gonzalez, X Li, **M Duprie**, AC Ventura, M Banerjee, M Ljungman, SD Merajver, CG Kleer.

"Histone methyltransferase EZH2 induces Akt-dependent genomic instability and 21406404. *Cancer Res. 2011 Mar 15;71(6):2360-70.* ME Gonzalez, **ML DuPrie**, H Krueger, SD Merajver, KA Toy, CG Kleer.

"TNFR-associated factor 6 and TGF-β-activated kinase 1 control signals for a senescence response by an endosomal NK cell receptor." *J Immunol. 2014 Jan 15;192(2):714-21.* S Rajagopalan, EC Lee, **ML DuPrie**, EO Long.

ABSTRACT OF THE DISSERTATION

Recruitment and function of Mlh1-Pms1 in DNA Mismatch Repair

by

Matthew DuPrie

Doctor of Philosophy in Biomedical Sciences

University of California San Diego, 2020

Professor Richard Kolodner, Chair

Mismatch Repair (MMR) is a highly conserved DNA repair pathway that repairs base-base mispairs and small insertion/deletion loops that frequently arise during DNA replication. Defects in MMR cause cancer, and studies in *Saccharomyces cerevisiae* and *Escherichia coli* have helped elucidate the function of MMR in humans. In *E. coli*, homodimeric MutS initially recognizes a DNA mispair, which then recruits homodimeric MutL and further downstream proteins. In *S. cerevisiae*, MMR is initiated by MutS-homologous complexes Msh2-Msh6 or Msh2-Msh3 binding a DNA mispair, which then recruit the MutL-homologous complex Mlh1-Pms1 to DNA. This interaction between

these complexes is essential for MMR, though due to its dynamic and transient nature it has been difficult to study. Using Deuterium Exchange/Mass Spectrometry (DXMS) of *E. coli* MutS and MutL, I identified a putative interaction interface in the N-terminal hetero-dimer region of Mlh1-Pms1, though further biochemical analysis indicated that this region was likely not the interface region. While performing these studies, a published low-resolution crosslinked structure of the *E. coli* MutS and MutL identified a new putative interface region between MutS and MutL. I used this information to elucidate the interface between Msh2-Msh6 or Msh2-Msh3 and Mlh1-Pms1. Genetic assays determined mutations in this interface region of *MLH1*, but not *PMS1*, caused a null MMR phenotype, and biochemical assays concluded that it is the Mlh1 subunit that interacts with Msh2 and functions in the initiation of MMR. I also characterized mutations affecting the flexible linker region of Mlh1. I found point mutations and deletions that lead to a complete loss of MMR, and found that purified Mlh1-Pms1 mutants containing amino acid substitutions in this region are defective for endonuclease activity, revealing an unknown essential functional for this region of Mlh1 in MMR. Together, my work increases our understanding of the initiation of eukaryotic MMR, as well as how Mlh1-Pms1 leads to downstream repair.

Chapter 1

Introduction

1.1 Global Context and Introduction to Mismatch Repair

All of life is stress. It exists in a medium of constant flux and likely bears these scars in its origin. More than simply containing certain biomolecules, life can be described as a process or system of storing and disseminating information through these biomolecules [1, 2]. Deoxyribonucleic acids (DNA) became the main vehicle of information storage for organisms, and so maintaining the fidelity of a DNA sequence while propagating it has become a defining characteristic of life [3-5].

Therefore, an organism strives to minimize DNA damage which would threaten the integrity of its sequence. These threats come from both within and without. DNA integrity suffers spontaneous base deamination, abasic sites, methylation, hydrolysis with local water molecules as well as exposure to reactive oxygen species [6]. The chemical bonds holding DNA together can be broken and rearranged by ionizing and ultraviolet rays, chemicals such as alkylating agents, aromatic amines, polycyclic aromatic hydrocarbons and other reactive electrophiles, toxins, and environmental stressors such as extreme heat or cold, hypoxia, and oxidative stress [6]. In addition to exogenous sources, errors during DNA replication can also lead to the formation of mispairs, which become mutations after the next round of replication if left unrepaired [7]. Mismatch repair (MMR) is the highly conserved DNA repair pathway that repairs mispairs that form in the genome [7].

This has implications for human disease since a defect in MMR increases the mutational burden on the genome. Loss of MMR occurs in many spontaneous tumors, and an inherited defect in a single copy of an essential MMR gene is the cause of Lynch Syndrome (or hereditary non-polyposis colon cancer) [8-17]. Lynch syndrome

is a genetic predisposition to develop an array of cancers throughout the body, though predominantly of the colorectum and endometrium, with a mean onset between 43-46 years of age [17, 18]. Inheriting biallelic mutations in any essential MMR gene causes constitutional mismatch repair deficiency (CMMR-D), which leads to early onset pediatric cancers [18].

1.2 Eukaryotic DNA Replication and the formation of DNA mispairs

Nuclear DNA replication begins at replication origins. Replication is typically bidirectional from origins, with most leading and lagging strand synthesis being catalyzed by two multi-subunit polymerases, Pol δ and Pol ε [19]. Both contain 3'-exonuclease proofreading activity [19]. Studies suggest that Pol ε primarily synthesizes the leading strand while Pol δ synthesizes the lagging strand [20].

Occasionally DNA polymerases make mistakes when incorporating new nucleotides while replicating DNA [7]. Based solely on the difference in free energy between correct and incorrectly paired bases, a misincorporation error should occur every 100 to 1000 base pairs [21]. A typical high-fidelity polymerase (such as Pol δ or ε) imposes selectivity and has an error rate of 10^{-4} to 10^{-5} per nucleotide, and including their proofreading exonuclease activity these polymerases can achieve an error rate of 10^{-7} per nucleotide [21-23]. Misincorporated bases that escape surveillance by the proofreading polymerases are called DNA mispairs, which can either be a non-Watson Crick base pair or a small insertion/deletion loop of nucleotides, which typically occur in sequences with low sequence complexity, such as homopolymeric tracts.

MMR functions to recognize and repair DNA mispairs that escape proofreading by DNA polymerases. Globally, MMR reduces the 10^{-7} per nucleotide error rate of the DNA polymerase activity to 10^{-9} to 10^{-10} per nucleotide [7, 24, 25]. If mispairs are uncorrected, they become fixed as mutations in the genome after the next cycle of replication [7, 26, 27]. Short repetitive DNA sequences of 1-4 base pairs are subject to frameshift mutations at a higher rate due to polymerase slippage, and MMR is the main DNA repair pathway that repairs the insertion/deletion loops created by these errors [28, 29]. This phenotype of high frameshift mutation rates in repetitive DNA regions is called microsatellite instability [30, 31].

1.3 Methyl-Directed MMR

Mismatch repair was first conceived in 1964 by Robin Holliday studying gene conversion in fungi and Evelyn Witkin investigating brominated nucleotide processing in *Escherichia coli* [32-34]. Gene conversion is the non-reciprocal transfer of genetic information between two or more gene copies as a result of a recombination event [35]. Holliday predicted the existence of the Holliday junction as a DNA intermediate in homologous recombination that could form mispaired bases, and that the processing of these mispaired bases would lead to gene conversion [32, 36, 37].

The first organism in which MMR genes were fully elucidated was *E. coli* [38-43]. Genetic studies identified *mutS*, *mutL*, *mutH*, *uvrD* and *dam* as essential genes for MMR [40-42, 44-46]. MMR in *E. coli* is described as methyl-directed, since the methylation status of the DNA strands determines which strand serves as a template and which strand is excised in MMR [47]. This system is found only in a set

Mismatch Repair in *E. coli*

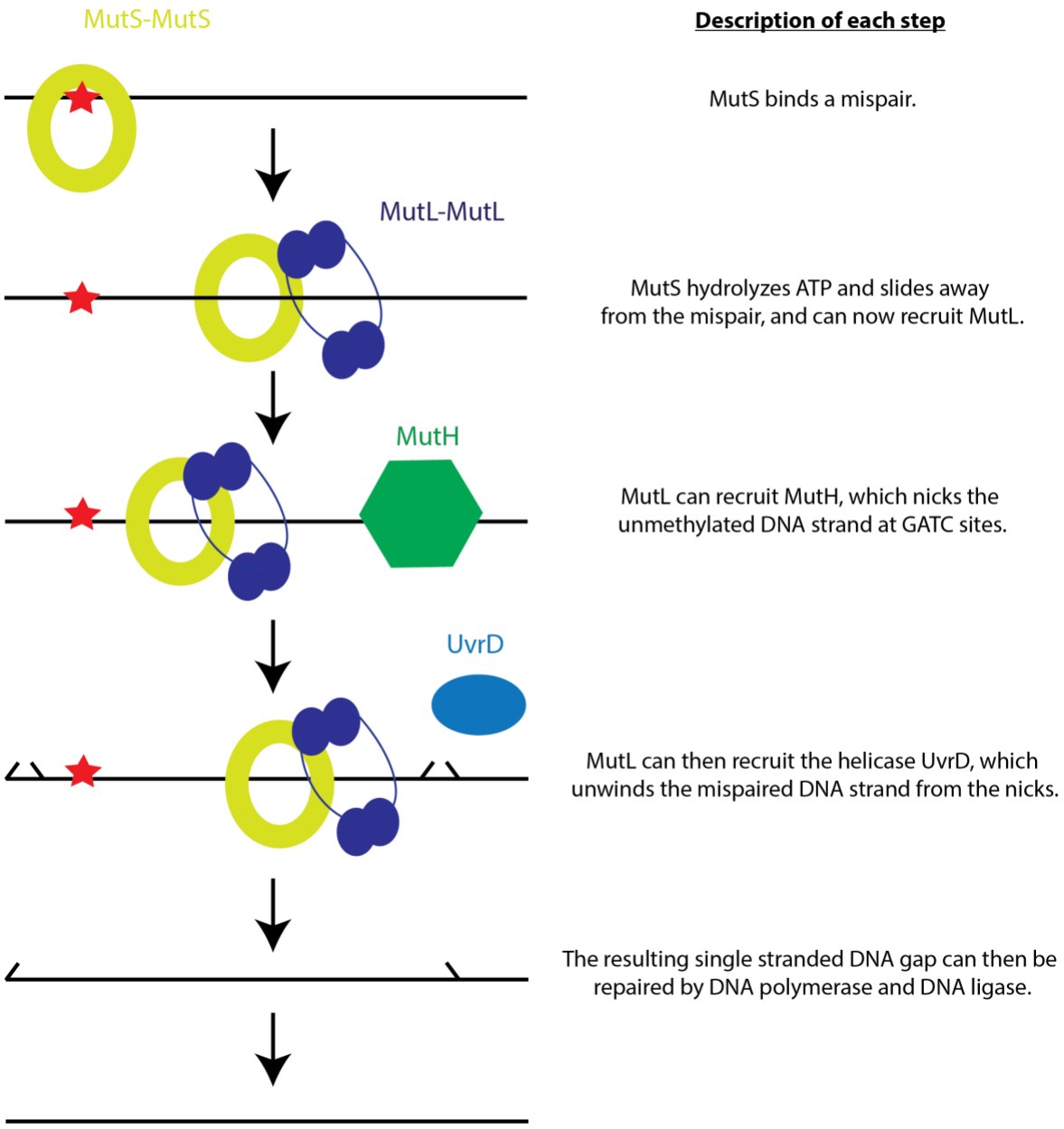


Figure 1.1. Model of methyl-directed mismatch repair in *E. coli*. The model begins at the top left, and follows down vertically through the repair process. The DNA mispair is shown as a red star. Each step is indicated by a black arrow, and a description of each step is given on the right. Greater description is given in the text.

of closely related gammaproteobacteria [48], and what serves as the strand discrimination signal in other organisms is still unclear.

Mispairs are recognized by MutS, which is a homodimeric ABC-family ATPase [49-51]. Crystal structures of MutS from *E. coli* and *Thermus aquaticus* show that the mispair recognition complex binds DNA at the site of the mispair, bends the DNA by ~60 degrees, and opens the DNA base stack so that one face of the base in the mispair is extruded from the helix and is stabilized by π-stacking with a conserved phenylalanine side chain [52-54]. The mispair recognition complex is functionally asymmetric; only one of the two subunits directly interacts with the mispaired bases. Importantly, the mispaired base that is extruded from the DNA helix is not necessarily the one on the strand that will be excised and resynthesized. This is also true for the eukaryotic MutS homologs. Thus, the functional asymmetry of the mispair recognition complex does not act as the MMR strand discrimination mechanism in MMR.

Many studies have examined how MutS searches the genome and identifies mispairs. Single molecule studies on *T. aquaticus* MutS show that while searching for a mispair it exhibits both repeated dissociation from and rebinding to DNA and 1-dimensional rotational diffusion while in continuous contact with the helical backbone of the DNA, much like a nut rotating on a screw [55]. Mismatch recognition of MutS is influenced by the sequence of mispaired nucleotides as well as those of the nearest neighbor nucleotides [56, 57]. Studies have shown that mismatched nucleotides increase DNA flexibility by weakening nucleotide base stacking interactions, which can then introduce a deformation in the DNA helical structure [58-60]. The function of the conserved phenylalanine side chain that is found coordinating the DNA mispair could

then be to kink DNA searching for local regions of increased flexibility. Mispairs are then recognized by these local structural alterations, rather than specific recognition of the mismatched nucleotide [19, 61], which could explain the wide variety of mispairs and lesions that can be recognized.

After MutS binds a mispair, an ADP to ATP exchange is induced concurrent with a conformational change to a "sliding clamp" structure, which releases the bend on DNA and which becomes more stably associated with DNA [62-67]. MutS sliding clamps show diffusion with free rotation and discontinuous contact with the DNA backbone, and its diffusion on DNA is independent of ATP hydrolysis [55, 68-71]. This conformational change is essential for the recruitment of the downstream protein MutL, which functions in signal amplification. ATP binding by MutS is sufficient for the recruitment of MutL; however, ATP hydrolysis is required by *E. coli* MutS in the complete *in vivo* MMR reaction [72].

MutL is also a homodimeric ATPase, which is recruited to DNA by MutS [62, 73-76]. The N-terminal ATPase domain belongs to the GHKL family [77], and is separated from the C-terminal domain by an unstructured linker. The C-terminal domains of MutL are constitutively dimerized, whereas the N-terminal domains dimerize only upon ATP binding to form a ring [78, 79]. MutL homologs in organisms without methyl-directed MMR have a C-terminal domain with an endonuclease activity that generates single-stranded breaks in DNA [80-85]. In organisms with methyl-directed MMR (including *E. coli*), the C-terminal domains have similar folds, but the endonuclease motif is missing [86], and is involved in activating downstream proteins [87].

MutH is recruited by MutL and activated to nick the DNA strand that is to be resynthesized [88, 89]. Strand discrimination in *E. coli* and related bacteria is novel since these bacteria contain the DNA methyl transferase Dam. The Dam methyltransferase catalyzes post-replication methylation of the adenosine N_6 position at palindromic d(GATC) sites using S-adenosylmethionine as a substrate [90]. DNA replication of a fully methylated template gives rise to hemi-methylated d(GATC) sites, in which the template strand is methylated and the newly synthesized strand is unmodified. This transient hemi-methylation status only lasts on the order of minutes [90, 91]. MutH makes a single stranded nick 5' of the G in the unmethylated strand of hemi-methylated d(GATC) Dam sites [88].

These MutH generated nicks can then serve as entry points for displacement of the newly synthesized strand by the UvrD helicase and potential degradation of the displaced strand by a single-stranded DNA exonuclease [92]. The 3' -> 5' UvrD DNA helicase is bound and activated by MutL [93, 94]. While *in vitro* UvrD is inefficient and displays very low processivity [95, 96], single molecule imaging studies have shown that ATP-bound MutL captures UvrD near strand breaks and greatly increases its unwinding processivity [97]. The MMR system can be bi-directional *in vitro*, proceeding from a 3' or 5' nick [89, 98]; however, the direction of repair *in vivo* always proceeds towards the replication fork [98]. A combination of the UvrD helicase and one of four single stranded DNA specific exonucleases (Exo1, ExoIV, RecJ, ExoX) excise the nicked strand past the mispair and the resulting singled-stranded gap is filled in by DNA polymerase III, single strand DNA binding protein and DNA ligase [37, 92, 99, 100], though recent work suggests that UvrD alone that is required to unwind DNA

between nicks present on either side of the mispair and that the exonucleases are not needed for MMR [97]. This model is depicted in **Figure 1.1**.

1.4 Eukaryotic MMR Proteins

Initially the existence of MMR in eukaryotes was questioned, because the *S. cerevisiae* genome lacked all DNA methylation and the human genome lacked adenine methylation, which precluded an *E. coli*-like methyl-directed repair system [36, 101, 102]. In 1986 our laboratory showed that a MMR system functioned in *S. cerevisiae* through the repair of mismatch-containing plasmids constructed *in vitro* [103]. Seymore Fogel identified the *pms1* ("Post-Meiotic Segregation 1") mutation in *S. cerevisiae*, which caused an increase in post-meiotic segregation events [104]. In 1987 our laboratory for the first time demonstrated that this mutant had a defect in repairing mismatch-containing plasmids during mitotic growth [105]. The Fogel laboratory then cloned the gene, which turned out to be a homolog of *E. coli mutL* [106].

In eukaryotes, gene duplication and specialization led to the emergence of multiple MutS and MutL homologs which interact to form heterodimers. In fungi and mammals, the MutS complexes acting in MMR are the Msh2-Msh6 and Msh2-Msh3 heterodimers [7, 107]. The common subunit Msh2 does not directly interact with mispairs, and deletion of the gene that encodes Msh2 results in a complete loss of MMR. A loss of *MSH3* or *MSH6* on its own does not lead to a complete loss of MMR [107-110]. Both Msh6 and Msh3 directly recognize mispairs and have distinct but

overlapping specificities for DNA mispairs [7, 107, 111-113]. Msh2-Msh6 preferentially binds and recognizes individual base-base mispairs and small insertion deletion loops, while Msh2-Msh3 primarily recognizes larger insertion/deletion loops [114]. Exchanging the mispair-binding domains of Msh3 and Msh6 but keeping the rest of the protein the same switches these proteins mispair specificities [115]. These complexes and mispair specificities hold true for human Msh2-Msh6 and Msh2-Msh3 [116-118]. Further genetic evidence in humans supports the importance of both *MSH3* and *MSH6* in functional MMR [119-123].

The initial steps of eukaryotic MMR are similar to MMR in *E coli,* where MutS homologs initially bind to a mispair and then recruit MutL homologs [34, 36]. The MutL homologs form heterodimers, and in *S. cerevisiae* Mlh1-Pms1 is the most important heterodimeric complex involved in MMR (in humans the homologous complex is called Mlh1-Pms2) [124, 125]. Atomic force microscopy and size-exclusion chromatography suggest that Mlh1-Pms1 may hydrolyze ATP to undergo a cycle of changing structure, from an extended conformation with just the C-termini dimerized to a compact globular conformation with both the N- and C-termini dimerized and little space between the domains [78, 126]. Yeast and human Mlh1-Pms1 have an endonuclease activity located in the C-terminal domain that is not present in *E. coli* MutL [82, 83, 127, 128]. This endonuclease is weakly active on supercoiled DNA, but can be further stimulated by RFC and PCNA to be highly specific for nicking a DNA strand that is already nicked [82, 83, 127, 129]. The mechanism for this strand specific nicking is unknown.

Other downstream mismatch repair associated proteins identified include the 5'-3' exonuclease Exo1 [130], the replication clamp Proliferating Cell Nuclear Antigen

(PCNA) [82, 131], the replication clamp loader Replication Factor C (RFC) [132], the single-strand binding protein Replication Factor/Protein A (RPA), DNA ligase, and DNA polymerases δ and ε [112, 133-135].

1.5 Model for Eukaryotic MMR

The current model for eukaryotic MMR is that Msh2-Msh6 or Msh2-Msh3 recognize a mispair while in the ADP-bound or nucleotide-free state. Upon binding a mispair, the Msh complex exchanges ADP for ATP and adopts a sliding clamp formation, that both slides away from a mispair and is proficient for recruitment of one or more Mlh1-Pms1 complexes. PCNA activates the endonuclease activity of Mlh1-Pms1 and stimulates nicking on the newly synthesized DNA strand. This endonuclease activity of Mlh1-Pms1 is essential for MMR, though why multiple strand nicks are required is still unknown [81, 83, 136]. Interestingly, while partial defects in this endonuclease activity can cause a mild to moderate defect in MMR, they can synergize with a loss of the exonuclease Exo1 leading to a near full defect in MMR [137-139]. As the loss of Exo1 has little effect on the MMR on its own, these results suggest that efficient nicking by Mlh1-Pms1 can bypass the role of Exo1 *in vivo*, and as Mlh1-Pms1 becomes less efficient the role of Exo1 becomes more integral. This has led to the model that MMR can proceed down two pathways after the recruitment of Msh2-Msh6 and Mlh1-Pms1: Exo1-dependent and Exo1-independent repair [139]. In the Exo1-dependent pathway, Exo1 is recruited by Msh2-Msh6 and Mlh1-Pms1 to excise the

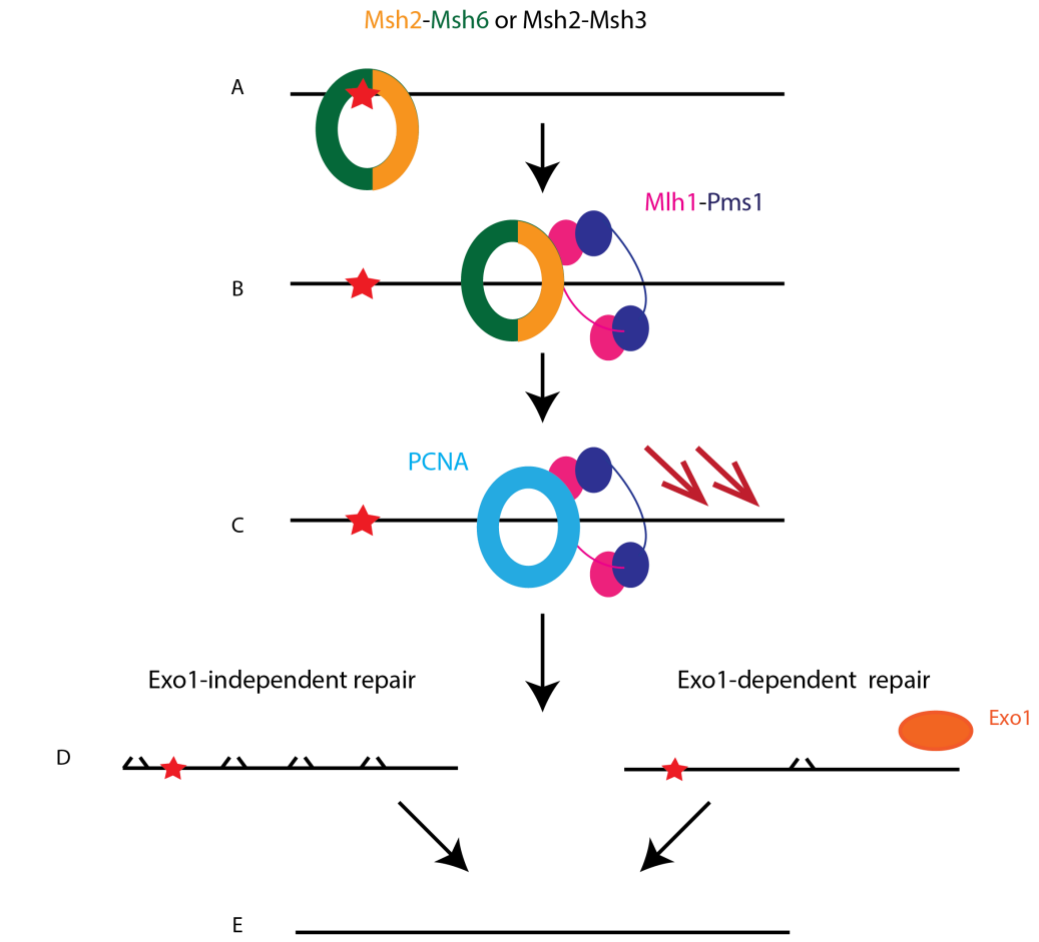


Figure 1.2. Mismatch repair in eukaryotes. This model shows a simplified schematic of our current mechanistic understanding of mismatch repair in eukaryotes. **A** shows the initiation step similar to *E. coli*, where either Msh2-Msh6 (shown) or Msh2-Msh3 will initially bind the mispair. In **B**, Msh2-Msh6 will hydrolyze ATP and form a sliding clamp, allowing the recruitment of Mlh1-Pms1 to DNA. **C** shows the endonuclease function of Mlh1-Pms1, stimulated by its interaction with PCNA, will nick the DNA strand that is to be excised multiple times, and that this activity is essential for repair. This results in **D**, where multiple paths to repair have been proposed. Adequate nicking function by Mlh1-Pms1 can generate enough nicks where the recruitment of the exonuclease Exo1 is non-essential; however, if this nicking activity is attenuated, then recruitment of Exo1 to the DNA becomes important for repair. Final repair shown in **E** can then be achieved in exo1-independent repair by strand displacement synthesis by a DNA polymerase recruited to the nick, or in exo1-dependent repair by exonuclease digestion of the mispair containing strand by Exo1, which is recruited by Msh2-Msh6 and Mlh1-Pms1 to the DNA. The resulting ssDNA gap can then be filled by DNA polymerase and DNA ligase.

mispair-containing strand in a 5' to 3' direction until the mispair is converted to a gap [140]. The mechanism of termination of excision by Exo1 is unknown, but could be mediated after excision passes the mispair by weak processivity of Exo1 in the absence of recruitment by Mlh1-Pms1 or Msh2-Msh6 [141]. In the Exo1-independent pathway, the genetic evidence is consistent with the requirement of multiple nicking events by Mlh1-Pms1. Processive nicking by Mlh1-Pms1 could lead to excision of the newly synthesized strand in the vicinity of a mispair [138] or could generate 3' ends that could initiate strand displacement synthesis by DNA polymerase δ [142]. Both pathways can lead to mutation avoidance [139]; however, *in vivo* studies of fluorescently tagged Pms1, which forms foci that are MMR repair intermediates, suggest that the Exo1-dependent pathway is faster and therefore more frequently used [143]. These studies also showed that Msh2-Msh6 is recruited to DNA by PCNA [143]. There is evidence in humans that Msh2-Msh6 is recruited by a chromatin modification during S-phase through a deuterostome-specific PWWP domain present in the otherwise unstructured Msh6 N-terminal tail [144-148]. This model is depicted in **Figure 1.2**.

1.6 *In vitro* Reconstitution Assays

In vitro reconstitution assays have also been developed to further understand MMR, both in human *and S. cerevisiae*. These studies have defined two types of

mispair-dependent excision/repair reactions. In one reaction, a combination of Msh2-Msh6 or Msh2-Msh3, Exo1, DNA Pol δ, the single stranded DNA binding protein RPA, PCNA and RFC promotes the repair of a circular mispaired substrate containing a nick on the 5' side of the mispair. In this reaction, the mispair recognition factors and other proteins promote excision by Exo1 past the mispair, followed by repair DNA synthesis to fill in the resulting gap repairing the initial mispair [112, 149, 150]. In a second reaction, the combination of Msh2-Msh6 or Msh2-Msh3, Mlh1-Pms1, Exo1, DNA pol δ, RPA, PCNA and RFC promotes the repair of a circular mispaired substrate containing a nick on the 3' side of the mispair. In this reaction, the Mlh1-Pms1 endonuclease is activated to generate nicks 5' to the mispair, which leads to excision and subsequent gap filling either via an Exo1-dependent 5' excision and subsequent gap filling or strand-displacement synthesis by Pol δ [127, 142, 149]. In both of these reactions, mispair recognition by Msh2-Msh6 and Msh2-Msh3 appears to stimulate excision by Exo1, in part by overcoming inhibition of Exo1 by RPA [151]. Further studies by our laboratory have shown that pol ε can also substitute for pol δ in these reactions [135]. Two remarkable discrepancies have been identified between *in vivo* and *in vitro* MMR models: (1) the ATPase activity of Msh2-Msh6 is required *in vivo* but only partially *in vitro* [152] and (2) the endonuclease activity of Mlh1-Pms1 is required *in vivo* but not for many *in vitro* reconstituted assays. These discrepancies argue that the *in vitro* reconstituted reactions do not recapitulate all of the complexities of the *in vivo* MMR reaction.

1.7 Link Between MMR and DNA Replication as Possible Strand Discrimination Signals

A key question for non-methyl-directed MMR mechanisms is how the newly synthesized strand is distinguished from the template strand. This strand discrimination is crucial for MMR, as excision-repair on the template strand would serve to fix the error caused by the mispair into the genome. *In vitro* reconstituted reactions are strand-specific, and this strand specificity is dictated by the presence of a pre-existing nick in the substrate [82, 83, 127, 150]. If nicks are a signal, then DNA replication itself could provide the MMR strand discrimination signal [153, 154]. Although pre-existing nicks are more common on the lagging strand than on the leading strand due to Okazaki fragment formation, recent reports indicate the presence of long-lived nicks on the leading strand [155]. This hypothesis relies on a close coordination between MMR and replication machinery, and there is much evidence to support this association.

Many studies have linked eukaryotic MMR to the replication machinery. Msh3 and Msh6 bind to PCNA [156-158], which is a processivity factor for DNA polymerase δ [159, 160]. Live cell imaging in *S. cerevisiae* shows that Msh2-Msh6 form foci in essentially 100% of S-phase cells in a mispair-independent, PCNA-Msh6 interaction dependent fashion and these foci colocalize with replication fork components, suggesting that Msh2-Msh6 is a constitutive component of replication factories where it acts to detect mispairs formed during replication [143]. Another study showed that by temporally restricting the availability of Msh6 to different phases of the cell cycle by fusing it to cell-cycle specific cyclins in an *msh3Δ S. cerevisiae* strain, MMR in eukaryotes was restricted to a narrow post-replicative window of opportunity similar to

E. coli MMR [161]. This study showed that the window of time for MMR proficiency was no more than 15 minutes after the region of the genome was replicated, while heteroduplex rejection, another process dependent upon Msh2-Msh6, was not restricted to the same temporal window [161]. These studies support the theory that replication-generated DNA structures could serve as strand-discrimination signals for MMR.

Other hypotheses suggest that nicks might be indirect signals by proposing that asymmetric binding of PCNA on DNA or asymmetric loading of PCNA on replication-generated nicks could be the strand discrimination signal [154, 158, 162]. There is evidence for this *in vitro* [163, 164], however there is still no mechanistic proof for this nicking specificity. These models also further underscore our lack of understanding of the role the endonuclease activity of Mlh1-Pms1 plays in MMR, given that there are already nicks present in the DNA.

Misincorporated ribonucleotides from replication excised by RNase H2 have also been suggested to be a source of the DNA nicks leading to strand discrimination in MMR [165, 166]. However, the mutator phenotype caused by RNase H2 defects are orders of magnitude lower than those expected to be caused by a universal strand discrimination signal defect [167]. Further, mutations in RNase H2 defective strains are completely dependent on topoisomerase 1, indicating that these mutations do not reflect a MMR defect.

1.8 Understanding the MMR initiation complex

Much is still not understood of the biochemical initiation of MMR, and how Mlh complexes are recruited by Msh complexes to DNA. The formation of a ternary

complex between MutS and MutL complexes on DNA has largely been inferred from biochemical and genetic data, though structurally uncharacterized in part due to its dynamic nature [168]. The MutS-MutL-DNA ternary complex and the eukaryotic Msh-Mlh-DNA complex are dynamic and exhibit rapid dissociation [62, 73, 168-170], which may explain the failure of most large-scale physical interaction studies to identify the *S. cerevisiae* ternary complex [171-175]. Both ATP binding and mispair binding are required for the interaction of MutS with MutL and Msh2-Msh6 with Mlh1-Pms1 [62, 168-170]. These cofactors likely transiently induce conformational changes required for ternary complex formation. One paper identified mutations in the N-terminus of Mlh1 that eliminated its interaction with Msh2-Msh6, although it is unclear whether the mutations affect a region directly involved in complex assembly [176].

Multiple models have been proposed as to how the interaction between MutS and MutL homologs leads to the initiation of MMR [177]. Early models hypothesized that MMR proteins form a nucleoprotein filament on the DNA from the mismatch to the initiation site, although no evidence for such a filament has emerged [45]. A static transactivation model was proposed which had static MutS bound to a mispair recruiting MutL, and that this static complex communicates strand breakage by MutH via looping of the intervening DNA [178]. This model was disproven with evidence that artificial blocks on DNA between the mispair and GATC site inhibited MMR, suggesting some DNA translocation is necessary [179]. This evidence also disproved other trans-initiation models [75]. An ATP-dependent hydrolysis model was proposed where MutS homologs actively hydrolyzed ATP to continue sliding away from a mispair to communicate with MutL homologs [180, 181].

The Molecular-Switch model incorporates more modern understanding of the biochemistry of MutS homologs (Msh) and MutL homologs (Mlh). This model incorporates the evidence that mismatch recognition by Msh induces and ADP to ATP exchange much like a G-protein molecular switch [65, 66]. This sliding clamp formation can then recruit Mlh, and these complexes diffuse together along DNA to the strand scission site driven by thermal fluctuation [62, 65, 66, 182]. Uncertainty still surrounds the dynamics of this interaction between these protein complexes and how they lead to downstream MMR signaling. Msh and Mlh proteins could remain bound together as they slide, or they may dissociate allowing Msh to catalytically load multiple Mlh proteins. In addition, Mlh proteins have been proposed to aggregate together on DNA to form a polymer that is loaded by an Msh protein bound to the mismatch to the distant excision-initiation site. Previous *in vivo* fluorescence microscopy in *S. cerevisiae* done by our laboratory showed that Msh2-Msh6 and Mlh1-Pms1 form distinct foci during S-phase [143]. Experiments with *T. aquaticus* MutS and MutL appeared to show multiple MutL proteins trapping MutS at or near a mismatch after the recognition by MutS [183]. Studies have also suggested that Mlh1-Pms1 may bind DNA cooperatively as a polymer [184], and have recently shown that *S. cerevisiae* MutL homolog complex Mlh1-Mlh3 has endonuclease activity activated by polymer formation [185]. All these studies suggest that Mlh proteins form their own separate intermediates during MMR processing but do not prove that they must function as polymers on DNA.

Alternatively, single molecule imaging of *S. cerevisiae* Msh2-Msh6 and Mlh1-Pms1 using quantum dots showed Mlh1-Pms1 being recruited to DNA but forming separate sliding clamps following recruitment [70]. More recent single molecule

imaging experiments of *E. coli* MutS, MutL, and MutH showed that the processes of mismatch binding transmission appeared totally stochastic [186]. This study found that MutS recruited MutL after diffusing away from the mispair, and that MutL could form its own stable sliding clamp that could diffuse rapidly on DNA. MutS and MutL oscillated in their association with each other. This study found the lifetime of an ATP-bound MutL sliding clamp on DNA was around 800 seconds, while ATP-bound MutS had a lifetime of around 200 seconds. This study suggested multiple MutL sliding clamps can be loaded by an ATP-bound MutS, supporting the both a catalytic-loading model of Msh proteins and a sliding clamp functional model for Mlh proteins [186].

To further characterize the interaction between MutS and MutL, in 2009 our laboratory used hydrogen/deuterium exchange mass spectrometry of main-chain amides to examine solvent accessibility of MutS in the presence of ATP and a 71-nt DNA substrate containing a central GT mispair either with or without MutL [187]. This study identified two regions in *E. coli* MutS that showed substantial protection from deuteration in the presence of MutL, one in the second "connector" domain and one in the ATPase domain. Testing the homologous region in *S. cerevisiae* Msh2 and Msh6 showed that only mutations in this region of Msh2 caused a defect in Mlh1-Pms1 binding *in vitro* and an *in vivo* mutator phenotype.

In Chapter 2 I describe the efforts by myself and our laboratory to perform the reciprocal DXMS experiments to identify regions on MutL that showed increased protection from deuteration only in the presence of the connector domain of MutS. While I was pursuing these studies, a low-resolution structure of crosslinked single-cysteine mutant MutS and MutL was published in 2016 [188]. This structure provided

insight into the sliding clamp structure of MutS and identified putative MutS-MutL interaction interfaces. Here I describe my efforts to pursue both of these studies to illuminate the implications for *S. cerevisiae* MMR and identify what residues of Mlh1-Pms1 are involved with its interaction with Msh2-Msh6 and Msh2-Msh3.

1.9 The role of the flexible linker region of Mlh1

As more work has been done to fully understand the role of MutL and its homologs in MMR, more questions have been raised as to what, if any, is the functional role of the unstructured flexible linker region connecting the N- and C-terminal domains. In *S. cerevisiae*, the Mlh1 linker region is around 150 amino acids, and that of Pms1 is 250 amino acids [189]. It has been hypothesized that the long flexible linker arms allow Mlh proteins to form a ring-like structure around DNA [86], with the N-termini transiently dimerizing via their ATPase activity [126]. Single-molecule studies of *S. cerevisiae* Mlh1-Pms1 support this model, while also showing that Mlh1-Pms1, but not Msh2-Msh6, is capable of bypassing nucleosomes while remaining bound to DNA, suggesting the long linker arms are providing this capability [70, 190]. More recent studies single molecule studies on MutL have also shown the importance in liker arm length in bypassing blocks on DNA, showing that if the linker arms become too short MutL loses this ability to bypass obstructions on DNA [191, 192].

Other data suggests the linker regions are more than just passive tethers. Systematic mutagenesis of the entire *S. cerevisiae MLH1* gene identified point mutations in the Mlh1 linker region in *S. cerevisiae* that completely disrupt MMR *in vivo*, but do not affect the function of Mlh1-Pms1 in mediating meiotic crossover events

[193]. The linker region of Mlh1 is much more sensitive to mutation or deletion than the linker region of Pms1 [189], and switching the linker regions of Mlh1 and Pms1 causes a complete defect in MMR [192]. Inserting characterized or random amino acid sequences as a substitute for deleted sequences did not restore function to defective Mlh1 and Pms1 mutants [192]. These studies indicate that the sequence identity of these linker regions is important.

Conserved residues within the linker region may have roles in MMR beyond providing flexibility and length that allow Mlh1-Pms1 to bypass nucleosomes, such as potentially mediating protein interactions [189]. The linker region of Mlh1 is predicted to be unstructured, and studies have shown that unstructured regions of significant size (>50 residues), called intrinsically disordered regions (IDRs), are quite common and can impart function [194, 195]. A region is defined as intrinsically disordered if it does not fold spontaneously into well-organized globular structures in the absence of stabilizing interactions [196], and if it is characterized by low sequence complexity and a biased composition with few bulky hydrophobic amino acids and many charged and hydrophilic amino acids [197]. IDRs can actually mediate many *in vivo* functions and contain conserved sequence motifs that interact with nucleic acids or proteins [194, 196-198], and this could be another role of the Mlh1 linker region.

In Chapter 3 I will describe my efforts to further explore the function of the Mlh1 linker region by analyzing mutations in this region. Work by another laboratory testing deletions in the linker region in Mlh1 for MMR identified a region between amino acids 396-421 that was consistently deleted in all Mlh1 linker region mutations that cause a complete MMR defect [189]. Our laboratory performed a sequence conservation

analysis of 180 fungal *MLH1* sequences and found that this region contains many strongly conserved residues. Mutations that I characterized in this region were completely defective for MMR *in vivo*, and the corresponding mutant proteins failed to support both *in vitro* repair and Mlh1-Pms1 endonuclease reactions. Together, my work identified a novel functional importance for linker region of Mlh1 in MMR.

Chapter 2

Investigating the interface between Msh2-Msh6 and Mlh1-Pms1

Chapter 2 Introduction

Because of the likely transitory nature of the interaction between *E. coli* MutS and MutL and between the eukaryotic homologs of MutS and MutL, defining the interaction between these proteins has been difficult. Our laboratory has previously used DXMS to identify conformational changes of MutS after binding a mispair [199] and to identify that domain II of MutS, also called the "connector" domain, interacts with MutL [187]. This interaction surface was found to be conserved in *S. cerevisiae*; in the Msh2-Msh6 heterodimer, Msh2, but not Msh6, mediates this interaction with Mlh1-Pms1 [187]. Intriguingly, this study also determined that the isolated MutS domain II (MutSDII) could bind MutL independently of both ATP and mispair-containing DNA, which are normally required for the MutS-MutL interaction [187]. These results were also confirmed by protein crosslinking [200] and protein crystallography [188]. To define the complementary interacting surface on *S. cerevisiae* Mlh1-Pms1, I have pursued several different strategies.

2.1 DXMS leads to identification of region in MutL as a potential interaction site with MutS

The first strategy to identify the MutS-interaction region of *E. coli* MutL was initiated when Dr. Victoria Hargreaves, a previous postdoctoral fellow in our laboratory, performed DXMS experiments comparing the hydrogen exchange rates of full-length *E. coli* MutL alone with full-length MutL plus MutSDII in order to identify regions of MutL that were protected upon addition of MutSDII. This analysis demonstrated that amino

acids 292-300 of MutL exhibited a MutSDII-dependent decrease in hydrogen exchange rates (**Figure 2.1**).

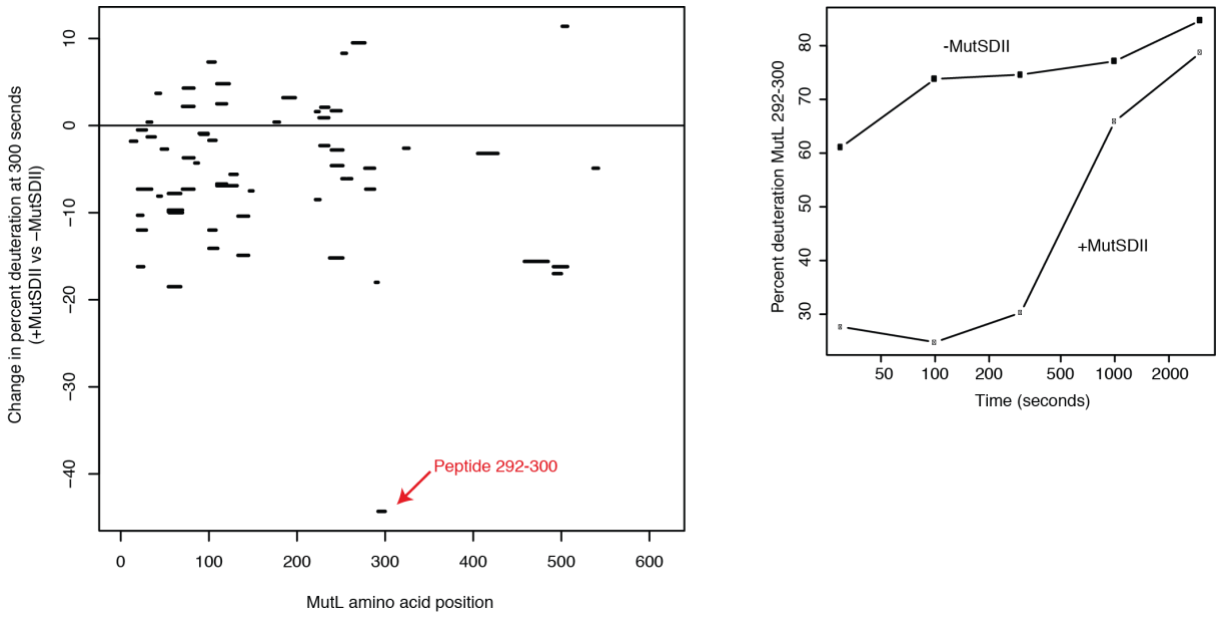
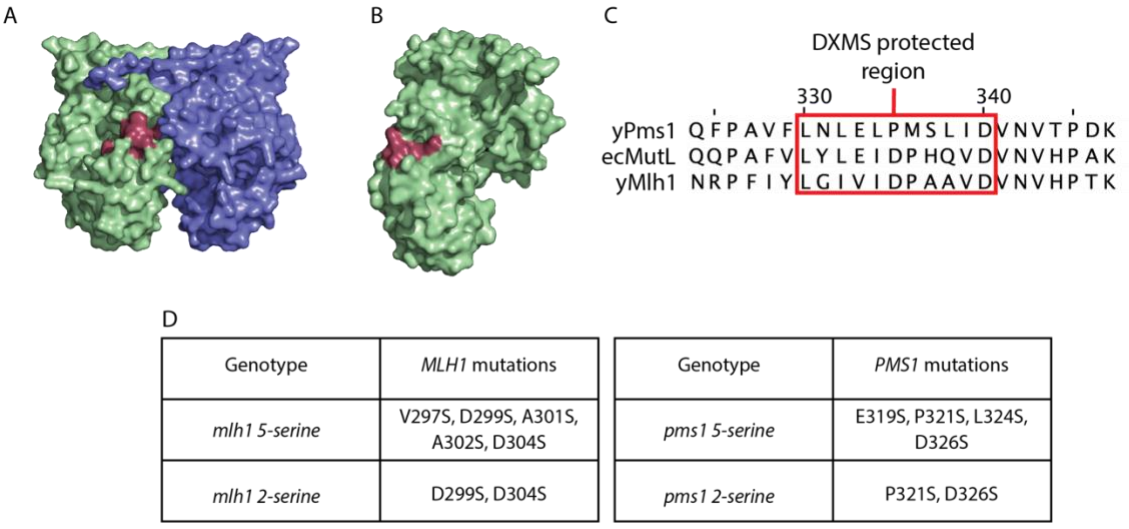


Figure 2.1. DXMS demonstrates that MutSDII causes reduced deuterium exchange for the MutL peptide spanning amino acids 292-300. The graph on the left shows the change in percent deuteration of MutL peptides +/- MutSDII at 300 seconds. The lines indicate individual observations of peptides; the horizontal span of the lines indicate the amino acids spanned by each peptide and the vertical position indicates the percent change in deuteration. The graph on the right shows the percent deuteration in the peptide 292-300 with and without MutSDII over the time course from 30 to 3,000 seconds.

2.2 Genetic analysis of the DXMS-implicated region shows that the region in Mlh1, but not Pms1, is required for mismatch repair

Because eukaryotic MMR involves functionally specialized heterodimeric homologs, I designed a series of mutations targeting the sequences homologous to *E.*

coli MutL 292-300 in both *S. cerevisiae* Mlh1 and Pms1 (**Figure 2.2**). As protein-protein interfaces tend to have large buried surface areas, I initially designed a series of serine substitutions that targeted multiple residues in this region (*mlh1-5-serine*, *pms1-5-*serine, mlh1-*2-serine*, *pms1-2-serine*).



D

Genotype	*MLH1* mutations
mlh1 5-serine	V297S, D299S, A301S, A302S, D304S
mlh1 2-serine	D299S, D304S

Genotype	*PMS1* mutations
pms1 5-serine	E319S, P321S, L324S, D326S
pms1 2-serine	P321S, D326S

Figure 2.2. Alignment of MutL DXMS region with S. cerevisiae homologs Mlh1 and Pms1. (A) Surfaces of the dimer of MutL N-termini (PDB: 1B63; [78]), with subunits in green and blue, and the region protected in DXMS experiments shown in red. (B) Rotating only one subunit, from panel A by 90º shows that the protected region extends into the dimer interface. (C) Sequence alignment between *E. coli* MutL and *S. cerevisiae* homologs Mlh1 and Pms1. (D) Mutations engineered in the 5-serine and 2-serine versions of *MLH1* and *PMS1*. Molecular images were generated using Pymol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.) Alignment performed in Clustal Omega and presented with Jalview [201, 202].

S. cerevisiae strains containing chromosomally integrated versions of these mutations were generated using the pop-in/pop-out strategy. These strains were then tested by fluctuation analysis to determine the rates of reversion of the *hom3-10*

frameshift mutation, which restores growth on medium lacking threonine, reversion of the *lys2::InsE-A10* mutation, which restores growth on medium lacking lysine, and forward mutation of the *CAN1* gene, which allows growth on canavanine-containing medium. These rates were compared to those of the MMR-proficient wild-type strain and the completely MMR-deficient strain *msh2Δ*. Both the *mlh1-5-serine* and *mlh1-2-serine* mutations caused mutation rates that were near the MMR-deficient strains, whereas the *pms1-5-serine* and *pms-2-serine* mutations caused only small increases in the mutation rates (**Table 2.1**). Thus, this region in Mlh1 but not Pms1 is essential for MMR.

Table 2.1. Mutation rates in three different mutator assays caused by mutations designed to disrupt the region of Mlh1 and Pms1 implicated by DXMS. Mutations as described in Figure 2.2 were integrated into the wild-type strain RDKY 3686 using the pop-in pop-out technique described in Chapter 5. This analysis indicated that mutations in *MLH1* caused a significant defect in MMR, while those in *PMS1* caused a milder defect or had no effect.

Genotype	*hom3-10* Reversion Rate	*lys2-10A* Reversion Rate	Can^R Mutation Rate
RDKY 3686	1.74[1.31-4.33]E-9 (1)	1.74[1.31-4.33]E-9 (1)	1.74[1.31-4.33]E-9 (1)
RDKY 4236 *mlh1*Δ	5.59[4.70-8.24]E-6 (3213)	1.25[0.86-1.46]E-4 (10593)	2.15[1.80-3.34]E-6 (43)
RDKY4238 *pms1*Δ	4.22[2.84-5.14]E-6 (1513)	2.69[1.47-5.48]E-5 (3233)	3.01[2.09-4.02]E-6 (40)
RDKY3688 *msh2*Δ	4.66[3.95-6.28]E-6 (2678)	8.73[5.71-10.7]E-5 (10593)	2.57[2.01-4.41]E-6 (52)
mlh1-5-serine	1.70[1.08-2.33]E-6 (977)	3.93[3.45-5.13]E-5 (3331)	9.16[7.42-13.6]E-7 (11)
mlh1-2-serine	9.42[8.40-11.8]E-7 (541)	1.98[1.08-2.62}E-5 (1678)	8.12[4.71-18.7]E-7 (10)
pms1-5-serine	5.42[2.98-6.65]E-8 (30)	5.83[4.20-9.47]E-7 (119)	1.27[0.88-1.64]E-7 (2)
pms1-2-serine	4.10[2.59-5.52]E-9 (2)	2.29[0.97-3.18]E-8 (1)	7.00[5.98-11.6]E-8 (1)
mlh1-D304S	1.64[0.90-2.17]E-6 (943)	6.99[3.44-9.97]E-5 (5924)	1.08[0.60-1.72]E-6 (22)
mlh1-D304K	4.54[2.68-8.46]E-6 (2609)	3.86[0.94-5.53]E-5 (4639)	3.12[1.61-5.06]E-6 (41)

To dissect the role of the amino acids in this region of Mlh1, each of the point mutations present in the *mlh1-5-serine* mutation were constructed individually in a replicating plasmid bearing the *MLH1* gene, and the ability of these mutant plasmids to complement a *mlh1Δ* strain were compared to the wild-type plasmid and an empty

vector control. The mutator phenotype of these complemented cells were measured by patch test (**Figure 2.3**). Remarkably, only *mlh1-D304S* (present in both the *mlh1-5-serine* and *mlh1-2-serine* mutations) caused a substantial MMR defect, while the other four mutations caused little or no change in the mutation rates (**Figure 2.3**). These results were confirmed by generating strains with chromosomally integrated version of the *mlh1-D304S* and *mlh1-D304K* mutations. Consistent with the complementation results, the serine substitution and the charge-swapping lysine substitution both caused a complete loss of MMR, similar to a deletion of *MSH2* or *MLH1* (**Table 2.1**).

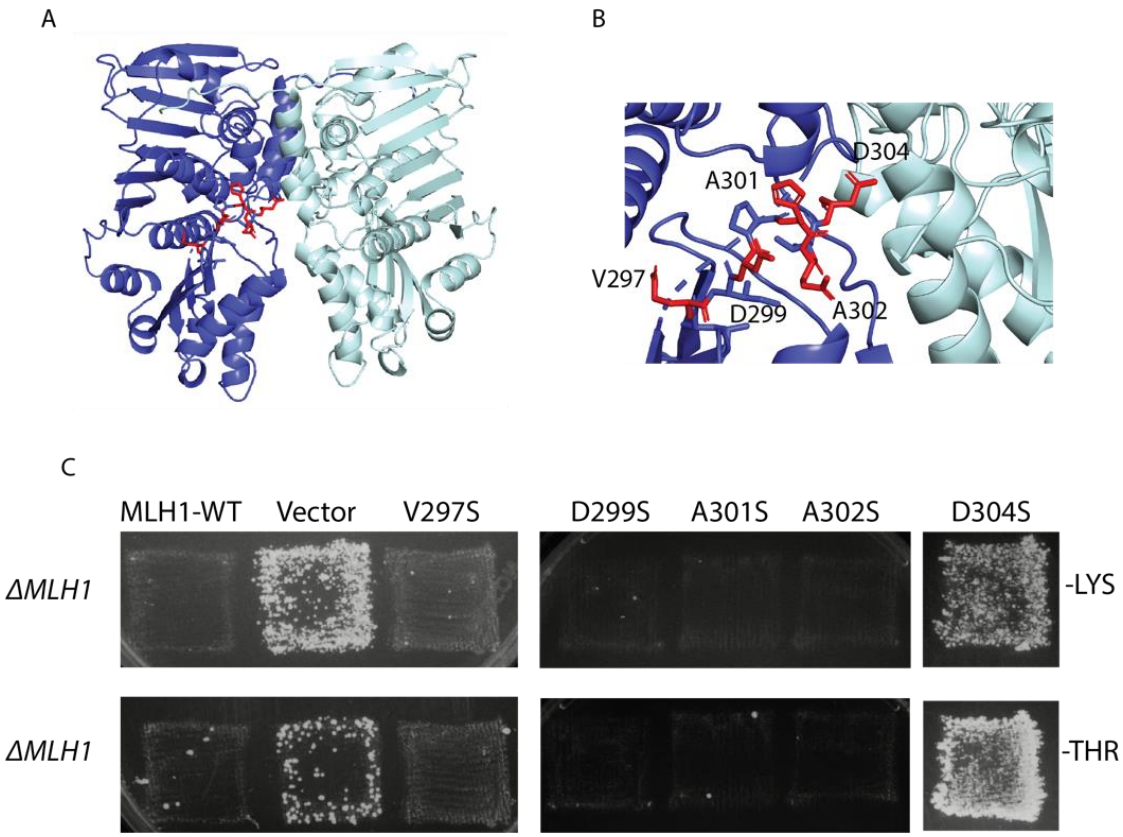


Figure 2.3 *mlh1-D304S* accounts for the entire MMR null phenotype of the *MLH1* mutations made in DXMS region. Each of the serine mutations described in Figure 2.2 in *mlh1-5-serine* were made individually on replicating yeast plasmids. These mutations are depicted on a structure of the MutL N-terminus dimer in (A) and (B). The ability of these plasmids to complement a *mlh1Δ* strain is shown in (C). Increased mutation rates correspond to increased number of papillae.

2.3 Surface plasmon resonance experiments show Mlh1 D304K is proficient for ternary complex formation

One hypothesis consistent with the DXMS and mutation rate results is that Mlh1 D304 is part of the Mlh1-Pms1 interface involved in recruitment by Msh2-Msh6 and Msh2-Msh3, which is required for functional MMR. To test this, wild-type Mlh1-Pms1-

FLAG and mutant Mlh1-D304K-Pms1-FLAG heterodimers were expressed from plasmids and purified from *S cerevisiae* cells with a deletion in *MLH1*, to prevent contamination of the mutant heterodimers with wild-type complexes. I then used a

Surface Plasmon Resonance (SPR) assay previously established in our laboratory to study the ability of Msh2-Msh6 to recruit the Mlh1-Pms1 variants to mispaired DNA [114, 170]. SPR detects small mass changes at the surface of a gold chip, and our assay uses a chip with streptavidin covalently attached to the surface to bind 236 bp biotinylated substrate DNA molecules. While one end of the DNA has a biotin moiety, the other end of the DNA substrate contains a *lacO* sequence so that both ends are blocked when *E. coli* LacI is added; blocking both ends is required as both Msh2-Msh6 and Mlh1-Pms1 can slide off free DNA ends in the presence of ATP [114, 170]. The two 236 bp DNA substrates are identical except for a central base pair that either is a G:C base pair (control) or a G:T mispair. The experiment is performed in two steps: (1) Msh2-Msh6 is added to the immobilized and end-blocked DNA in the presence of ATP, and (2) Mlh1-Pms1 is added in the presence of both Msh2-Msh6 and ATP.

Consistent with previous results [114, 170], Msh2-Msh6 preferentially binds mispaired DNA over basepaired DNA as measured by an increase in the response units, which is proportional to the amount of mass at the surface (**Figure 2.4**). Subsequent addition of wild-type Mlh1-Pms1-FLAG showed that Mlh1-Pms1 is also recruited to mispaired DNA in a mispair-dependent manner (**Figure 2.4**). In contrast to our hypothesis, addition of Mlh1-D304K-Pms1-FLAG mutant showed that this mutant heterodimer is still capable of recruitment to mispair-containing DNA (**Figure 2.4**). This

result suggests that the *mlh1-D304K* mutation does not disrupt mismatch repair through loss of recruitment by with Msh2-Msh6.

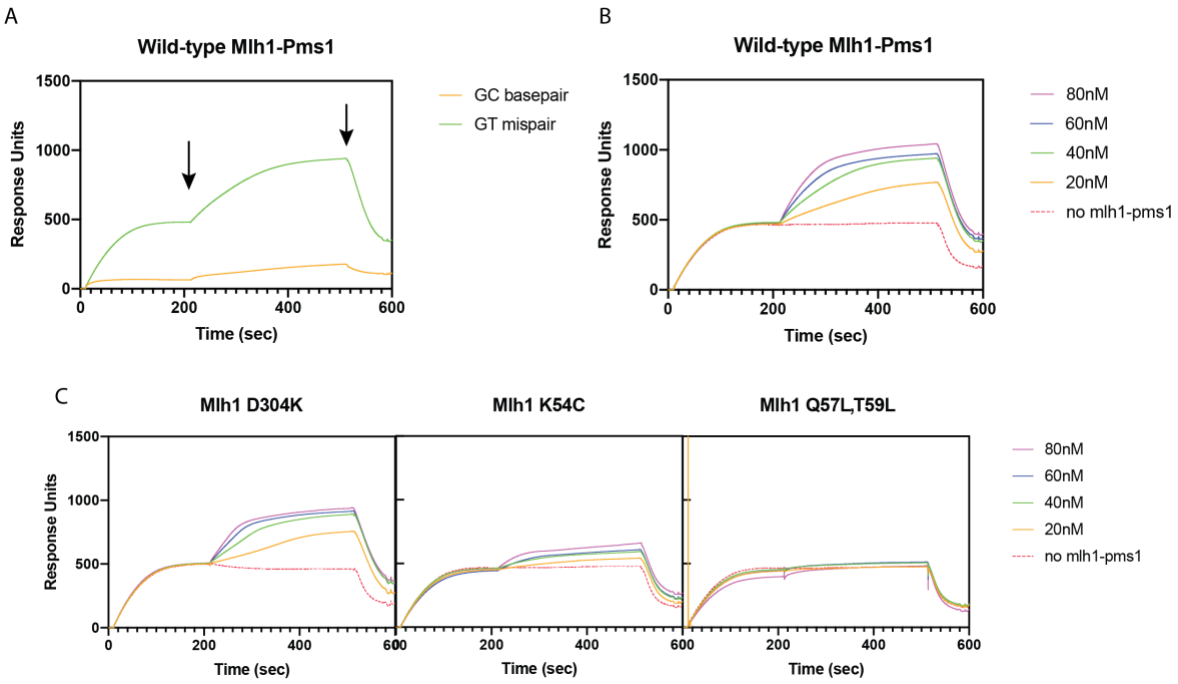


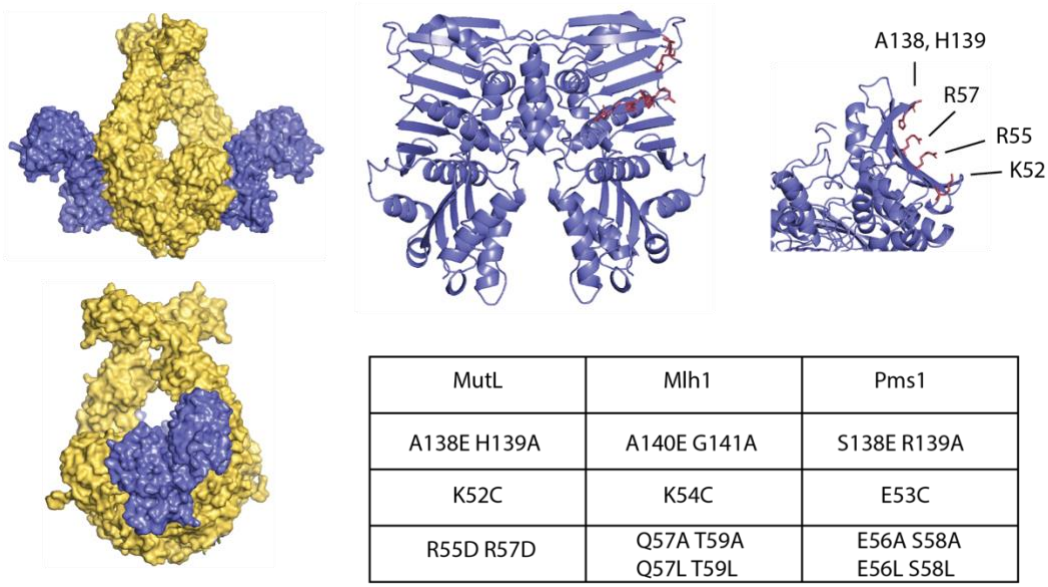
Figure 2.4. Surface Plasmon Resonance (SPR) experiments with Mlh1 mutants. Specific details of SPR method described in Chapter 5. (A) Shows the typical wild-type response curve for ternary complex formation using wild-type Mlh1-Pms1 and Msh2-Msh6. From 0 to 200 seconds, Msh2-Msh6, LacI, and ATP are flowed over surface of SPR chip with a 236 bp DNA substrate with either a central basepair or mispair as indicated. Binding at surface of the chip is detected as arbitrary Response Units. At the first black arrow, Mlh1-Pms1 is added to the previous components. At the second black arrow, all components are removed and only buffer flows over the chip. (B) Shows the response curves for wild-type Mlh1-Pms1-FLAG at varying concentrations of Mlh1-Pms1 as indicated. (C) Response curves for Mlh1 mutants.

2.4 Amino acid substitutions at the interface predicted by the crosslinked MutS-MutL structure in Mlh1, but not Pms1, disrupt MMR *in vivo*

While the studies regarding the DXMS-implicated region of Mlh1 were in progress, the structures of a MutS-MutL complex in three different crystal forms at 4.7, 6.6, and 7.6 Å resolution were published [188]. The conformation of these low resolution structures were consistent; however, these structures were generated by using chemical crosslinkers based on our laboratory's previous determination of the interacting domain of MutS by DXMS [187]. The crosslinking raises the possibility that the observed complexes might not represent the true MutS-MutL interface. Importantly, the MutL surfaces implicated in the interface were quite different than that predicted on the basis of the DXMS experiment (**Figure 2.5**). To investigate the relevance of this crosslinked conformation to eukaryotic MMR, I designed homologous mutations based on this crystal structure that were predicted to abrogate the interaction between Msh2-Msh6 and Mlh1-Pms1. Since the bacterial homologs are homodimers, mutations were separately engineered in both the Mlh1 and Pms1 homologs.

I therefore constructed the *MLH1* mutations (*mlh1-A140E,G141A*, *mlh1-K54C*, *mlh1-Q57A,T59A*, *mlh1-Q57L,T59L*) and *PMS1* mutations (*pms1-S138E,R139A*, *pms1-E53C*, *pms1-E56A,S58A*, *pms1-E56L,S58L*) in replicating yeast plasmids by site directed mutagenesis and transformed them into strains lacking *MLH1* or *PMS1*, respectively. I then tested the ability of these mutant genes to complement the mismatch repair defect of a chromosomal deletion by patch test (**Figure 2.6**). All of the *mlh1* mutations disrupted MMR to some degree; *mlh1-A140E,G141A*, *mlh1-Q57A,T59A*, and *mlh1-Q57L,T59L* mutations caused a total loss of MMR, and *mlh1-*

K54C caused a more modest disruption of MMR. In contrast, the majority of pms1 mutations had little or no MMR defect, with the notable exception of *pms1-S138E,R139A*. Given that both *pms1-S138E,R139A* and the



MutL	Mlh1	Pms1
A138E H139A	A140E G141A	S138E R139A
K52C	K54C	E53C
R55D R57D	Q57A T59A Q57L T59L	E56A S58A E56L S58L

Figure 2.5. Crosslinked structure of MutS and MutL identifies putative new interaction region between Msh2-Msh6 or Msh2-Msh3 and Mlh1-Pms1. (Left) The structure of crosslinked *E. coli* MutS (yellow) and MutL (purple) [188]. (Right) Mutations were designed in *S. cerevisiae* homologs *MLH1* and *PMS1* that affected residues conserved with those in the MutS-MutL interface from the crosslinked structure.

homologous *mlh1-A140E,G141A* caused a complete loss of MMR, we interpreted these mutations as potentially causing destabilization of the Mlh1 and Pms1 structure rather than interfering with the Mlh1-Pms1 interaction with Msh2-Msh6, so I focused on the other *MLH1* mutations. I integrated the *mlh1-K54C* and *mlh1-Q57L,T59L* mutations into the yeast genome and tested the strains by fluctuation analysis, which confirmed that *mlh1-K54C* has a moderate MMR defect while *mlh1-Q57L,T59L* has a complete MMR defect (**Table 2.2**).

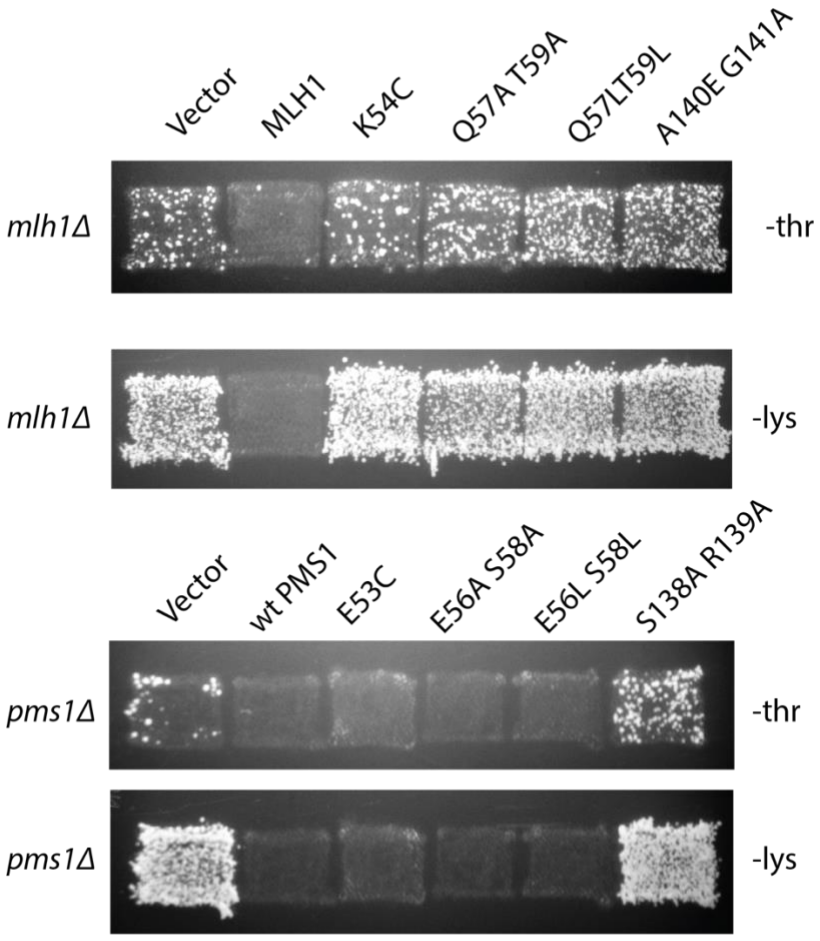


Figure 2.6. Mutations in MLH1, but not PMS1, that affect the putative interface cause loss of MMR. The mutations described in Figure 2.5 were made individually on replicating yeast plasmids. The ability of *MLH1*-bearing plasmids to complement a *mlh1Δ* mutation (top) and the ability of *PMS1*-bearing plasmids to complement a *pms1Δ* mutation (bottom) was assessed using patch tests for the *hom3-10* (-thr) and *lys2::InsE-A10* (-lys) reversion assays.

Table 2.2. Fluctuation Rates of mutations made based on crosslinked MutS-MutL structure. Interface mutations described in figure 2.5 were integrated into the wild-type yeast strain RDKY 5964 using the pop-in pop-out method described in Chapter 4.

Genotype	*hom3-10* Reversion Rate	*lys2-10A* Reversion Rate	Can^R Mutation Rate
RDKY 5964	1.77[0.79-4.32]E-9 (1)	5.94[1.31-13.6]E-9 (1)	5.26[3.93-7.62]E-8 (1)
*msh2*Δ	4.20[3.39-7.87]E-6 (2,373)	1.36[1.24-4.65]E-4 (22,896)	2.04[1.56-4.22]E-6 (39)
*mlh1*Δ	4.17[3.30-6.95]E-6 (2,356)	1.42[1.10-1.59]E-4 (23,906)	2.24[1.77-5.42]E-6 (43)
*pms1*Δ	4.70[3.14-7.23]E-6 (2,655)	9.43[7.33-11.7]E-5 (15,875)	2.59[2.35-3.74]E-6 (49)
mlh1-Q57L,T59L	3.33[2.29-3.84]E-6 (1,881)	9.23[5.35-13.2]E-5 (15,539)	1.81[1.27-3.70]E-6 (34)
mlh1-K54C	4.89[4.08-9.28]E-8 (28)	2.88[2.18-3.89]E-6 (485)	1.24[0.97-1.98]E-7 (2)
mlh1-R401A,D403A	2.26[1.93-3,66]E-6 (1277)	6.40[5.43-11.1]E-5 (10774)	1.84[1.21-4.87]E-6 (35)
*mlh1-396-421*Δ	3.58[2.24-3.88]E-6 (2023)	8.32[6.30-12.6]E-5 (14007)	1.82[1.31-2.72]E-6 (35)

2.5 Mlh1-K54C-Pms1 and Mlh1-Q57L,T59L-Pms1 are defective for ternary complex formation

I next purified both Mlh1-K54C-Pms1-FLAG and Mlh1-Q57L,T59L-Pms1-FLAG complexes as described in methods and tested these mutant heterodimers for their ability to be recruited to mispaired DNA by Msh2-Msh6 using SPR. In contrast to Mlh1-D304K-Pms1-FLAG, both of these mutant heterodimers had a defect in ternary complex formation (**Figure 2.4**). Consistent with the magnitude of the defects observed in my previous genetic studies, the Mlh1-Q57L,T59L-Pms1-FLAG heterodimer had a more substantial ternary complex formation defect than Mlh1-K54C-Pms1-FLAG.

2.6 Mlh1-Q57,T59L-Pms1 and Mlh1-D304K-Pm1 are completely defective for *in vitro* repair, and Mlh1-K54C-Pms1 is only partially defective

As described in Chapter 1, our laboratory has previously described a reconstituted Mlh1-Pms1-dependent 3′ nick-directed MMR reaction requiring Msh2-Msh6 (or Msh2-Msh3), exonuclease 1 (Exo1), replication protein A (RPA), RFC, PCNA, and DNA polymerase δ, requiring Mg^{2+} and Mn^{2+} for optimal activity [127, 150]. We tested the Mlh1 mutant heterodimers Mlh1-D304K-Pms1-FLAG, Mlh1-Q57L,T59L-Pms1-FLAG, and Mlh1-K54C-Pms1-FLAG in this assay. The results showed that using wild-type Mlh1-Pms1-FLAG in this reaction leads to the repair of 24.4% of the initial substrate, while Mlh1-Q57L,T59L-Pms1-FLAG and Mlh1-D304K-Pms1-FLAG are completely defective for *in vitro* repair. Interestingly, Mlh1-K54C-Pms1-FLAG had around 40% the activity of wild-type (**Figure 2.7**). Since these assays are optimized for repair and may hide moderate defects in repair due to high protein concentrations, the

concentration of Mlh1-Pms1 was titrated down to compare the repair between wild-type Mlh1-Pms1-FLAG and Mlh1-K54C-Pms1-FLAG. The results showed that that the mutant Mlh1-K54C-Pms1-FLAG had ~30% wild-type repair activity (**Figure 2.7**).

2.7 Mlh1-Pms1 endonuclease assay shows that all Mlh1 mutants are proficient for endonuclease activity

The complete reconstituted MMR reaction reflects the mispair-dependent recruitment of Mlh1-Pms1 and its nick-directed endonuclease activity. The endonuclease assay tests only the endonuclease function of Mlh1-Pms1. Stimulated by the presence of PCNA and RFC, Mlh1-Pms1 will nick supercoiled DNA [83]. Therefore, the endonuclease activity of the different mutant Mlh1-Pms1 heterodimers mutants can be assayed by detecting the emergence of nicked DNA from supercoiled DNA as visualized on an agarose gel, since these DNA species will run at significantly different speeds. This analysis determined that Mlh1-K54C-Pms1-FLAG, Mlh1-Q57L,T59L-Pms1-FLAG, and Mlh1-D304K-Pms1-FLAG all have near wild-type levels of nicking capability (Figure 2.4).

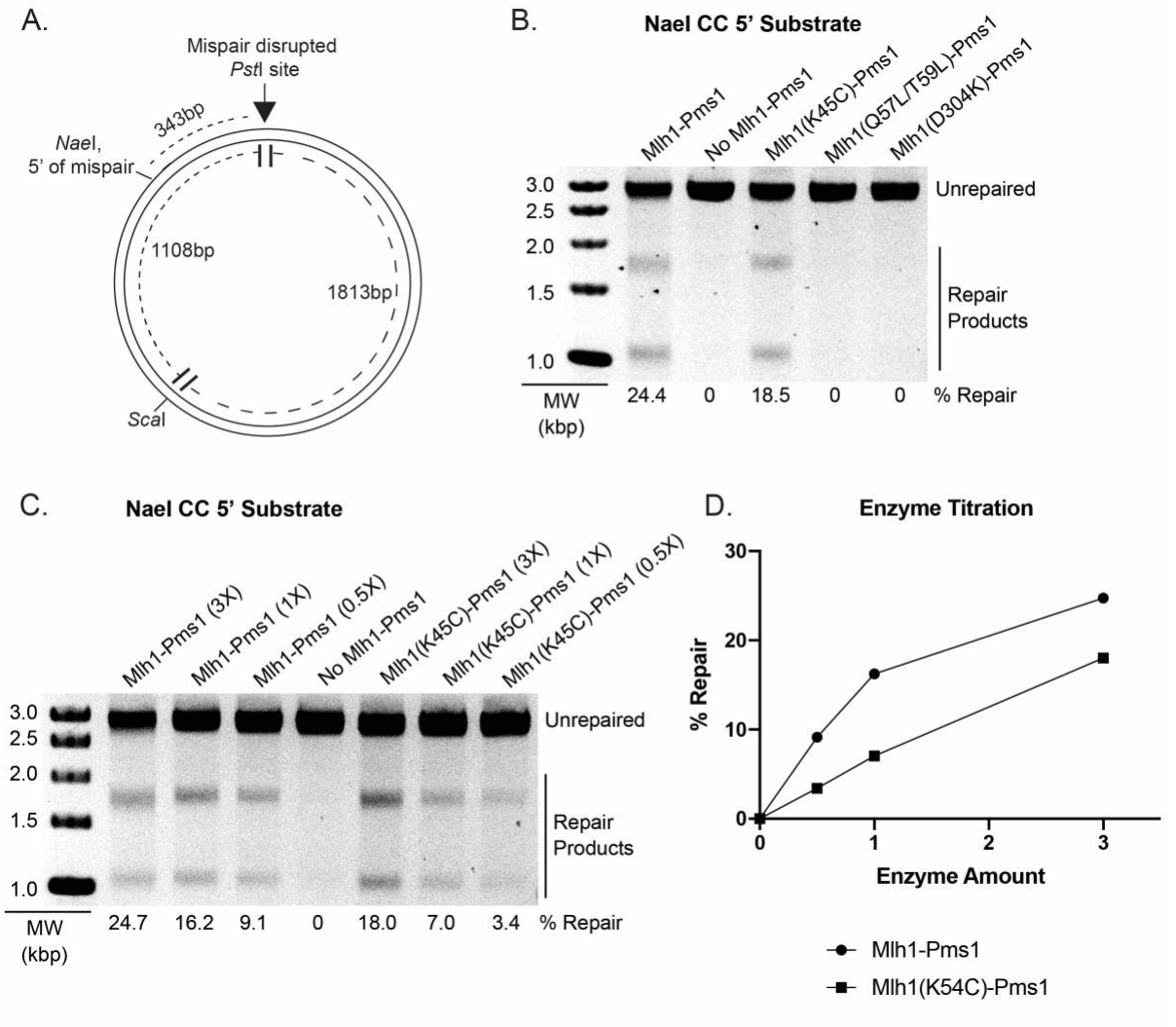


Figure 2.7. Reconstituted in vitro repair assay in *S. cerevisiae*. (A) The in vitro repair substrate contains a 5' nick at the *NaeI* site and a C:C mispair-disrupted *PstI* site. Repair of the nick strand restores the *PstI* site, and digestion of the repaired plasmid by both *ScaI* and *PstI* generates two fragments. Repair is shown in (B), as full DNA repair would lead to the restoration of the *PstI* cut site and the visualization of 2 bands. (C) Titration of wild-type Mlh1-Pms1-FLAG and Mlh1-K54C-Pms1-FLAG, graphically represented in (D).

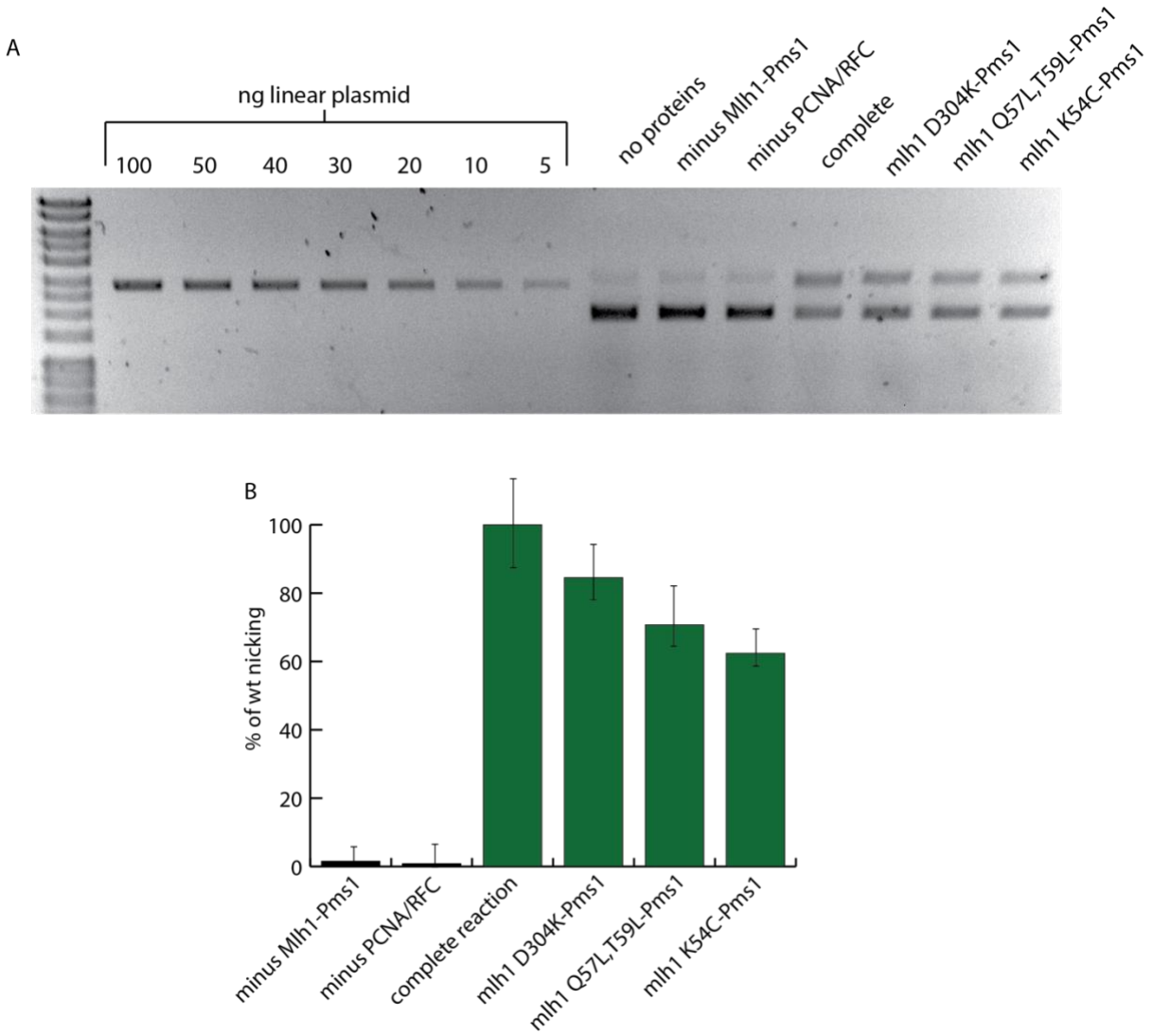


Figure 2.8. Mlh1-Pms1 endonuclease assay. Supercoiled DNA nicking assay as described in Chapter 5. (A) Supercoiled. plasmid was incubated with wild-type or mutant versions of Mlh1-Pms1-FLAG along with PCNA and RFC. Substrate DNA was then run on an agarose gel. In lanes corresponding to the reaction, the lower band is used as a standard (B) Mutant Mlh1-Pms1 mutants nicking activity as a percentage of wild-type Mlh1-Pms1 nicking.

2.8 Integrating *mlh1-D304K* mutation into the *PMS1-4GFP* strain background causes significant increase in the percent of cells with Pms1-GFP foci

Our laboratory has previously characterized a *S. cerevisiae* strain with endogenous Pms1 tagged at the C-terminus with 4 GFP molecules to allow visualization of Mlh1-Pms1 foci using fluorescent microscopy [143]. The 4GFP tag does not interfere with MMR, and accumulation of Pms1-4GFP foci are associated with an increase of repair intermediates: (1) their formation requires Msh2-Msh6, (2) their levels increase with mutations that increase the levels of mispairs, and (3) their levels increase when downstream MMR pathways are disrupted [143]. To determine what effect the *mlh1-D304K* mutation may have on the formation of Mlh1-Pms1-4GFP, I used the pop-in/pop-out technique to integrate the *mlh1-D304K* mutation into a strain expressing the Pms1-4GFP fusion (RDKY7544; *ura3-52, leu2Δ1, trp1Δ63, hom3-10, his3Δ200, lys2::InsE-A10, PMS1-4GFP::kanMX6*) and a strain containing the Pms1-4GFP and a deletion of *EXO1* (RDKY7588; *ura3-52, leu2Δ1, trp1Δ63, hom3-10, his3Δ200, lys2::InsE-A10, exo1::hphNT1, PMS1-4GFP::kanMX6*) backgrounds. Fluorescence microscopy was used to image the cells and determine the percentage of cells containing GFP foci. The percentage of wild type cells with Pms1 foci was ~1%; the reduction of foci relative to the initial published report is due to the difference in the ability to detect foci when using fluorescence vs. confocal microscopy. Deletion of *EXO1* increased the percent of cells with foci to ~4% (**Figure 2.9**). The strain with *mlh1-D304K* mutation integrated had ~28% of cell with foci, and the strain combining the *mlh1-D304K* mutation with the deletion of *EXO1* had ~23% of cells with foci (**Figure 2.9**). This result is in agreement with the biochemical studies on Mlh1-D304K-Pms1-

FLAG, which showed this mutant is still capable of being recruited to DNA by Msh2-Msh6, which is necessary for the formation of foci [143]. The large increase of foci in the *mlh1-D304K PMS1-4GFP* cells is consistent with the hypothesis that *mlh1-D304K* causes an increase in the levels of unresolved repair intermediates, and suggests that the MMR defect caused by *mlh1-D304K* occurs downstream of Mlh1-Pms1 recruitment to DNA.

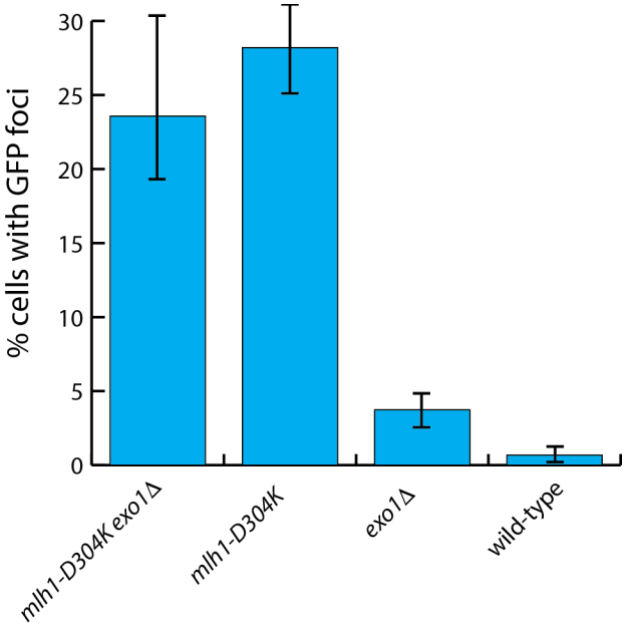


Figure 2.9 *mlh1-D304K* mutation in *PMS1-4GFP* foci strain. Going left to right, n = 661,1218, 639, and 860 cells counted as part of 3 separate image fields. The result shows the percent of cells in a given image field that are positive for at least one GFP foci.

The material in Chapter 2 is currently being prepared for submission for publication: DuPrie M., Hargreaves V.V., Calil F., Putnam C.D., Kolodner R.D. "Interaction between the shared subunits Mlh1 and Msh2 mediates recruitment between eukaryotic MutS and MutL homologs during DNA mismatch repair." The dissertation author was a primary researcher and author of this material.

Chapter 3

A conserved functional motif in the Mlh1 interdomain linker

Chapter 3 Introduction

In biochemistry, function follows structure. In eukaryotes, the Msh2-Msh6 and Msh2-Msh3 heterodimers form rings on DNA, a structure that is efficient in searching for mispairs on DNA, and once a mispair has been detected, these rings can undergo a conformational change to a ring tightly bound to DNA. *S. cerevisiae* Mlh1-Pms1 is a much smaller heterodimer, and composed of two domains connected by long flexible interdomain linkers. Why does Mlh1-Pms1 have this structure, and what advantage does this structure give to its function? The N-terminal domains belong to the GHKL ATPase family and may use ATP binding and hydrolysis to mediate cycles of dimerization like other GHKL ATPase domain-containing proteins [54]. The C-terminal domains mediate constitutive dimerization and contain an endonuclease active site [83]. The importance of the flexible linker connecting these two domains is less clear.

Previous studies have investigated the effects of point mutations in the interdomain linkers of *S. cerevisiae* Mlh1 and Pms1 [193], as well as larger deletions and insertions in these linkers in order to investigate the importance of the linker length [189, 192]. Multiple deletions in the Mlh1 linker caused MMR defects, which were interpreted as an absolute length requirement [189]. Our laboratory noticed that all of these deletions that caused a MMR defect overlapped the same linker region, amino acids 396-421. I therefore sought to investigate whether the MMR defect was caused by a deletion of a previously unrecognized functional region of Mlh1.

3.1 Conservation analysis identifies conserved residues in the Mlh1 interdomain linker

To investigate the region of the Mlh1 interdomain linker that was consistently lost in MMR-defective linker deletions [189], 180 protein sequences of *S. cerevisiae* and other fungal Mlh1 proteins were aligned using Clustal Omega [202], and the sequences that aligned with *S. cerevisiae* MLh1 amino acids 396-421 was analyzed. Remarkably, this analysis identified that the amino acids 390-414 had a substantial conservation as compared to the adjacent linker sequences and that several of the conserved residues in the mostly flexible and hydrophilic linker were hydrophobic (seq logo). These results suggested that the conserved region might represent a previously unrecognized functional region of Mlh1.

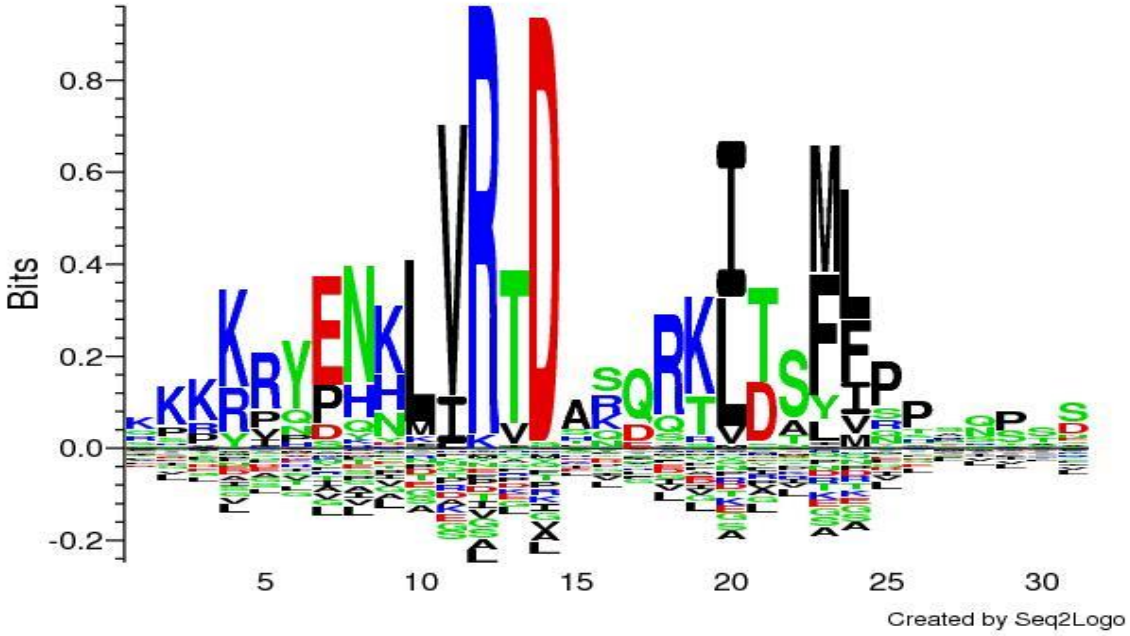


Figure 3.1. Sequence logo of the conserved region in the Mlh1 interdomain linker. Amino acids above line at zero are ones that are frequently observed; the level of conservation is shown by the height of the letter.

3.2 Both *mlh1-R401AD403A* and *mlh1-Δ396-421* are completely defective for MMR *in vivo*

To test if the conserved region of the Mlh1 interdomain linker is important for MMR, I engineered the double point mutation, *mlh1-R401A,D403A*, as well as a complete deletion of the conserved region, *mlh1⊿396-421*, into a pRS306/*MLH1* vector and integrated these mutations into a wild-type yeast strain containing the *hom3-10* and *lys2::InsE-A10* frameshift reversion assays and a wild-type *CAN1* using the pop-in/pop-out strategy. The mutation rates of these strains were then determined by fluctuation analysis. The results showed that all of these mutations caused a complete loss of MMR *in vivo* (**Table 2.2**).

3.3 Biochemical analysis indicates that Mlh1-R401A,D403A-Pms1-FLAG is proficient for ternary complex formation, but deficient for *in vitro* repair and endonuclease function

Given that the mutations affecting the conserved regions caused complete MMR defects, I sought to understand the mechanistic defect using biochemical assays. I engineered an expression vector for the Mlh1-R401A,D403A-Pms1-FLAG mutant protein and overexpressed and purified the mutant protein in a *S. cerevisiae* strain which contained a deletion of the chromosomal copy of *MLH1* in order to prevent contaminating the mutant protein preparations with wild-type Mlh1-Pms1-FLAG. I tested the mutant heterodimer in the reconstituted repair assay and the endonuclease assay (Figure 3.2). I also tested for ternary complex by SPR, and while my data suggested that the complex is still proficient for ternary complex formation, the experiment must be repeated. These results suggest that Mlh1-R401A,D403A-Pms1-FLAG is recruited to mispairs by Msh2-Msh6, but is completely defective for endonuclease activity and Mlh1-Pms1 dependent *in vitro* MMR. Remarkably, this conserved region of the interdomain linker has not been predicted to interact with the C-terminal endonuclease domain nor has the linker been previously implicated in endonuclease activity.

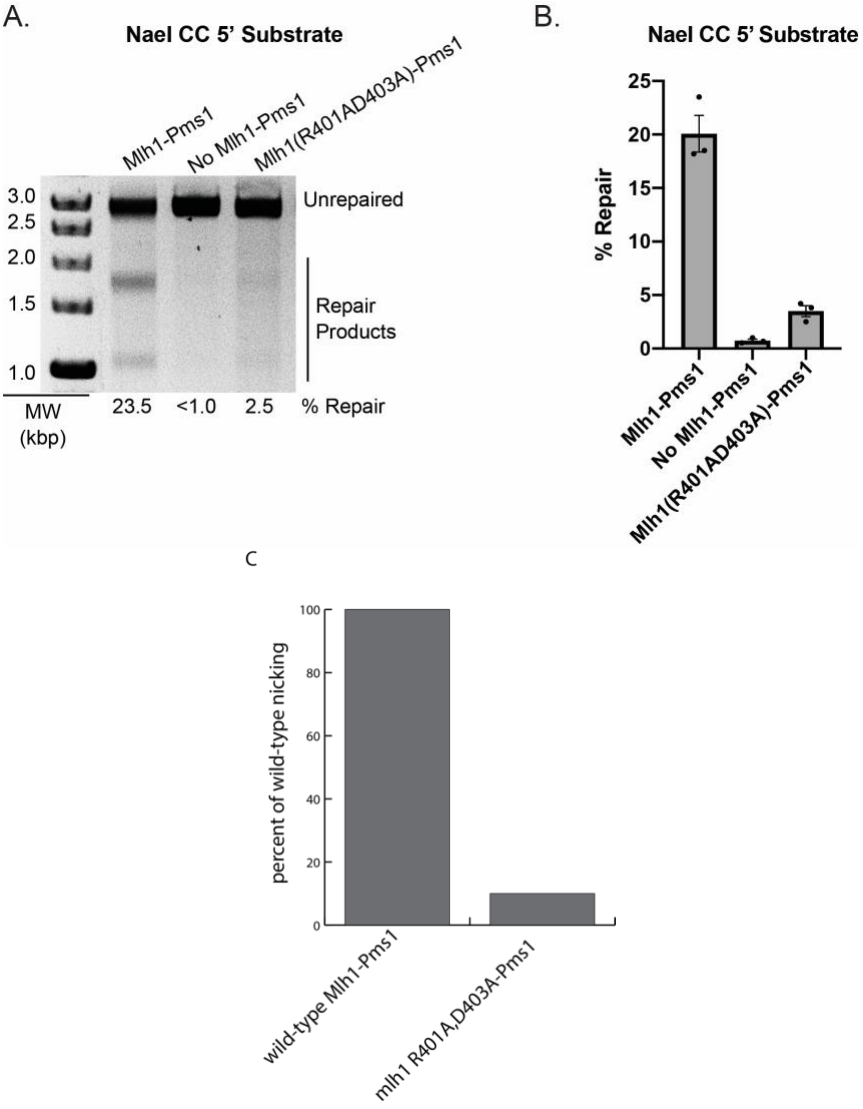


Figure 3.2. Characterizing Mlh1 R401A,D403A-Pms1 in reconstituted *in vitro* repair assay, endonuclease assay. (A) DNA product gel of in vitro repair assay with mlh1 R401A,D403A-Pms1-FLAG. (B) Histogram quantifying results in (A). (C) Results from supercoiled DNA Mlh1-Pms1 endonuclease assay. Results given as total nicking of substrate DNA as a percent of wild-type Mlh1-Pms1 nicking. This experiment is the result of only one experiment and must be repeated to be shown significant.

Chapter 4 Discussion and Analysis

4.1 Interpreting the DXMS studies

Identifying the interface between MutS and MutL has proven to be difficult, since this interaction is known to be so transient and dynamic. Recently, the Sixma group was able to crystallize a mispair- and ATP-dependent MutS-MutL structure, but only after experimenting with numerous cysteine mutants and crosslinking reagents [188, 200, 203]. Previously, our laboratory used DXMS to identify regions of *E. coli* MutS that are protected from deuterium exchange when MutL, mispair-containing DNA, and ATP were present [187]. The slower exchange could be caused by (1) formation of an inter-protein interface or (2) local conformational changes that shield regions of the protein from solvent or that induce local increases in secondary structure. Importantly, the exchange rates that are observed by DXMS are hydrogens on main chain amides and not those on side chains, hence much of the potential protein-protein interface may be invisible to DXMS. Consistent with these complications, of the two regions of MutS identified as being protected by MutL, only one of the two was confirmed to prevent the MutS-MutL interaction by mutagenesis [187]. Additionally, this study further identified that this region is conserved in Msh2, but not Msh6, in *S. cerevisiae* [187].

The data presented here followed up on these studies with the intention of identifying the regions of MutL that interact with MutS, and identifying whether these regions were conserved in Mlh1, Pms1, or both. We measured the rate of deuterium exchange of different regions of MutL in both the presence and absence of MutS connector domain. Using only the known interacting domain of MutS removed the requirements of mispair-containing DNA and ATP, which are required for the interaction of full-length MutS with MutL. In addition, this decreased the peptide

complexity for analysis of the resulting peptides by mass spectrometry. Reduction of the peptide complexity also led us to investigate the *E. coli* homodimers rather than the eukaryotic heterodimers, in spite of the fact that functional asymmetry in the homodimers might cause a reduced DXMS signal.

The DXMS data for MutL identified that the amino acid region 292-300 had slower deuterium exchange in the presence of the MutS connector domain (**Figure 2.1**). MutL residues 292-300 include part a flexible loop with many conserved residues that is disordered in the absence of ATP [78, 79]. This loop extends between the two ATPase domains of MutL, and interacts with both the ATP "lid" and the flexible loop of the first 20 amino acids. The conserved K307 on this loop is inserted into the ATP-binding site and directly contacts the γ-phosphate of the ATP analog in the structure. This loop is essential for both ATP binding and dimer formation, and based on this structure it would seem likely that this region becomes protected in the DXMS assay due to N-terminal dimerization. Victoria Hargreaves, the postdoctoral fellow who performed these studies followed up by purifying mutants of MutL that contained mutations in this region identified by DXMS. These MutL mutants were defective for recruitment to DNA by MutS, consistent with the hypothesis that the DXMS-identified region could be part of the MutS-MutL interface. These results suggested that the region of MutL identified by DXMS could be a region that is interacting with MutS.

I first pursued a genetic analysis by designing mutations in the homologous regions of Mlh1 and Pms1 to the region 292-300 of MutL. I tested these mutations using frameshift reversion assays and fluctuation tests. The results indicated that mutations in Mlh1 caused a complete defect in MMR, and mutating Pms1 caused little

effect (**Figure 2.2**). The genetic defect equivalent to *mlh1Δ* was attributed to mutating one residue in Mlh1, D304. The SPR ternary complex formation assay showed that the mutant Mlh1-D304K-Pms1-FLAG, while a null mutation *in vivo*, is still competent for recruitment to DNA by Msh2-Msh6. Therefore, this region is likely not at the interface of Msh2-Msh6 and Mlh1-Pms1.

Given the fact that Mlh1 D304 is at the predicted N-terminal dimerization interface and associated with the loop that interacts with bound ATP, I investigated the hypothesis that this mutation could cause a MMR defect due to a defect in N-terminal dimerization. MutL is extended in its native conformation and elutes in size-exclusion chromatography with a Stokes radius equivalent to a 300 kDa globular protein. Adding ATP or a non-hydrolyzable ATP analog causes the protein to adopt a more compact conformation with a Stokes radius equivalent to a 180 kDa globular protein [78, 79]. If Mlh1-D304K-Pms1-FLAG had a N-terminal dimerization defect, it would be predicted to elute more slowly than wild-type Mlh1-Pms1 in the presence of a non-hydrolyzable ATP analog. I tested the elution of wild-type Mlh1-Pms1-FLAG and Mlh1-D304K-Pms1-FLAG used a Superdex 3.2/300 size-exclusion chromatography, which should resolve these two conformations. Unfortunately, these experiments were frequently plagued by protein aggregation, problems with pressure in the HPLC system, and show large protein aggregates eluting in the void volume, and eventually pressure issues would compromise the integrity of the column. This was unfortunate, however looking back I realized a number of modifications I could try in the future. As a running buffer I was using only Tris and salt in water, a buffer a previous lab member used to test the protein PCNA, which is much more stable than Mlh1-Pms1. Looking at other

experimental setups where size-exclusion chromatography is used to test MutL or Mlh1-Pms1, the running buffers all contain detergents or glycerol, agents that can be used to reduce aggregation [78, 79]. MutL in particular is known to have aggregation issues. In the future, I would repeat my experiments with different buffers to try to prevent the aggregation issues.

Our laboratory has previously characterized a yeast strain with endogenous Pms1 tagged at the C-terminus with 4 GFP molecules [143], into which I integrated the *mlh1-D304K* mutation. The results showed a dramatic increase in the number of visible Pms1-GFP foci, which again argues that this mutant is capable of binding DNA, and this result is consistent with the generation of unresolved repair intermediates. These results all argue that *mlh1-D304K* interferes with steps downstream of recruitment of Mlh1-Pms1 to DNA.

More work must be done to fully characterize the defect caused by *mlh1-D304K*, as it is still not clearly understood. Although a defect in N-terminal dimerization could be hypothesized, my data showing that it this mutant is still competent for recruitment to mispaired DNA by Msh2-Msh6 would argue that this mutant is therefore still competent for binding DNA. Due to the location of this mutation on a loop that coordinates ATP binding, this mutation could hypothetically be disrupting the ATPase activity of Mlh1. An ATP hydrolysis assay could determine whether this is true.

Single-molecule experiments have recently shown that *E. coli* MutL can form a sliding clamp on DNA similar to MutS, showing marked decrease in dissociation from DNA once recruited [186]. It could be hypothesized that the *mlh1-D304K* mutation could be affecting the stability of this sliding clamp conformation. We have started a

collaboration to test our mutant protein in this setup to determine whether this mutation affects the retention of Mlh1-Pms1 on DNA once recruited.

4.2 Identifying residues on Mlh1 that are involved in binding Msh2-Msh6 and Msh2-Msh3

To obtain a crystal structure of MutS and MutL tethered together, the Sixma group had to replace all normally present cysteine residues in MutS and MutL and then introducing cysteines to find positions were MutS and MutL would crosslink using the flexible linker $BM(PEO)_3$ only when a mispair and nucleotide were present. This approach found that MutS D246C crosslinks to MutL N131C. The C-terminal domains of MutS and MutL were then removed and the protein complex crystallized and studied [188]. The structure showed a MutS dimer each bound to a MutL monomer. The MutS dimer showed a novel structure with the subunits tilted 30º towards each other compared to the mismatch recognition state, the connector domains rotated outwards 160º, and the mispair recognition domain in an unstructured state. The MutL interaction with MutS involved two interfaces, one involving the largest β-sheet of the ATPase domain of MutL and the ATPase and core domains of one subunit of MutS, and the second involved the side of the same β-sheet in MutL and a looped out helix interacting with the flipped out connector domain of the other MutS subunit.

This result provided interesting insight into our laboratory's previous DXMS experiments with MutS, which showed that the connector domain interface on MutS our lab discovered was located in their second MutS-MutL interface [187]. In addition, our laboratory identified a region in the ATPase domain of MutS that become protected

upon MutL interaction, and this structure showed that this is the region that becomes protected from solution on the MutS ATPase domain by the MutS connector domain rotating outwards, validating our conclusion that the region was protected but wasn't involved with directly interacting with MutL.

My work sought to explore the implications of the findings of this paper for MMR in *S. cerevisiae*, and by extension other eukaryotic systems. I made mutations in the homologous residues of Mlh1 and Pms1 to those shown in the crosslinked to be a part of the MutS-MutL interface region. The *mlh1-Q57L,T59L* and *mlh1-K54C* mutations were chosen because the equivalent mutations in *E. coli* MutL caused a MMR defect [188]. In *S. cerevisiae* Mlh1-Pms1, the *mlh1-Q57L,T59L* mutation is completely defective for MMR *in vivo* and the mutant protein is defective for ternary complex formation *in vitro*. The *mlh1-K54C* mutation showed only a partial defect MMR *in vivo*, and similarly showed only a partial defect in ternary complex formation *in vitro*. The weaker effect of *mlh1-K54C* may be due to the fact that it is difficult to disrupt an interface by mutating a single residue. Combined with the crystal structure, these results strongly implicate these residues as interacting with Msh2-Msh6. Making the homologous mutations on the Pms1 subunit caused no defect in MMR at all, furthering the argument that it is only the Mlh1 subunit of Mlh1-Pms1 that mediates its interaction with Msh complexes. Together with the finding that Msh2, the common subunit of the Msh complexes, interacts with the Mlh complexes [187], these results indicate that the common subunit of the Msh complexes recruits the common subunit of the Mlh complexes and that this is how the interaction between the complexes has been preserved even in the face of gene duplication and diversification.

This work is an example of the functional specialization in eukaryotic MutS and MutL homologs that has emerged from the functional asymmetry of bacterial homodimers MutS and MutL. Mispair binding by MutS induces an asymmetrical conformation, with only one subunit contacting the mispair [52, 53]. Similarly, while both subunits of bacterial *Bacillus subtilis* MutL contain an endonuclease motif, research has shown that only one of these subunits will functionally interact with β-clamp and therefore can function as an endonuclease [204]. My work has shown that, while in *E. coli* it is formally possible that a single MutS homodimer can recruit two MutL dimers with each of its two connectors domains, in eukaryotes functional specialization through evolution has led this interaction to be mediated specifically by Msh2 and Mlh1.

4.3 The role of the linker region of Mlh1 in MMR

Previous studies have suggested that the length of the flexible Mlh1 and Pms1 interdomain linkers is important for *in vivo* MMR [191, 192]. Our studies indicate that the linker in Mlh1 may have additional roles. Though these results presented here are preliminary, they are the first to identify and test the roles of the conserved region within the Mlh1 interdomain linker.

My results show that mutating residues in the region of Mlh1 between amino acids 396-421 causes a defect in MMR *in vivo* and disrupts MMR and the Mlh1-Pms1 endonuclease function *in vitro*. This is even more interesting since a group recently tested deletions of the entire Pms1 linker region and found no defect to endonuclease

activity [192], suggesting that it may be only Mlh1 that has a conserved region important for MMR function. We hypothesize that this region may be playing a role in interacting or recruiting other proteins, including PCNA, or it could be that this region is part of the active site of the endonuclease.

In the short term there are future studies that I hope to complete before I leave the laboratory. I am planning to make additional point mutations in the conserved region of Mlh1 and test them by fluctuation analysis. I am also planning to move this conserved region within the Mlh1 linker and potentially to the Pms1 linker to determine if the precise location of the region relative to the N- and C-terminal domain is important. In parallel, another laboratory member has performed the *in vitro* repair assay, using wild-type proteins including peptide fragments derived from the conserved region of the Mlh1 linker. Preliminary results have shown that adding these peptides inhibits *in vitro* repair, though the control experiment must still be completed to interpret these results fully.

4.4 Conclusions and final thoughts

I entered this laboratory because I have always wanted to pursue cancer research, and I am very happy that I have learned much about the field, as well as played my small role in advancing it just a little further. This has always been my life's dream, and I feel happy that I have accomplished it. It has taken much humbleness as well to understand that as my studies have progressed, new data has only opened new questions and new possibilities for experiments and understanding. I've had to realize that many of these questions will have to be solved by those that come after me. I

performed many other studies that may or may not ever see publication. This is the life of a basic scientist as I've found, and to cope I've learned to be enamored in the process, and less attached to results. At this moment at the end of my studies I feel I could make real headway into uncharted territory, but I realize that there will always be more, and fundamentally I must ask what it is all leading towards. I will look back years later and still try to see if we've understood what the strand-discrimination signal is in humans or what Mlh1-Pms1 does in MMR, with full understanding that this is an issue of pertinent interest only to those that directly worked on this problem. What has truthfully been the greatest lesson of my PhD has been to learn the importance of discussing ideas with others.

Chapter 5 Materials and Methods

Deuterium exchange mass spectrometry

Deuterium exchange mass spectrometry (DXMS) is a technique that can be used to determine the solvent accessibility of different regions of a protein by analyzing the ability of protein-bound hydrogens to be exchanged with deuterium in the solvent [205-208]. Proteins of interest are incubated in deuterated water to drive the exchange, quenched with acid at various time points to stop the exchange, digested into peptides with proteases, and finally analyzed using mass spectrometry (MS). Importantly, DXMS can only measure the exchange of hydrogens at the main chain amides, as hydrogen bound directly to carbon essentially do not exchange, and hydrogens on the functional groups of amino acid side chains (e.g. -OH, -SH, -NH2, -COOH) exchange too rapidly for deuteriums to be retained during sample processing. Increased rates of hydrogen exchange provide insight into protein structure; main chain amides hydrogens that are flexible regions of the protein exchange more rapidly (on the order of seconds) than in regions of secondary structure or buried in the interior of the protein (on the order of days).

Full length MutL and ATP with and without Domain II of MutS (MutSDII) were incubated in 10 µL buffer containing 250 µM ATP, 20 mM Tris (pH 8), 4 mM $MgCl_2$, 230 mM NaCl, 4 mM DTT, and 10% glycerol was mixed with 30 µL of D_20 containing 5 mM Tris (pH 8), 4 mM $MgCl_2$, 50 mM NaCl, and 250 µM ATP (final concentrations were 2.5 mg/mL MutL, 4.33 mg/mL MutSDII, 250 µM ATP, and 95 mM NaCl) and incubated for 30, 100, 300, 1,000 and 3,000 seconds at 4ºC. At the indicated times, the sample was added to vials containing 60 µL of quench solution (0.8% formic acid and 0.8M GuHCl) and immediately frozen at 80ºC. In addition, a non-deuterated

sample (incubated in H_20 buffer) and a fully deuterated sample (incubated in D_20 buffer containing 0.5 M GuHCl for 16 hours at 25ºC were prepared.

The 100 µL samples were thawed and immediately passed through an immobilized protease column (66 µL bed volume) of porcine pepsin (Sigma) coupled to a 20AL support (PerSeptive Biosystems) at a flow rate of 100 µL/minute. Proteolytic fragments were collected contemporaneously on a C18 HPLC column (Vydac) and eluted by a linear gradient (5-45% solvent B in 30 minutes, 50 mL/minute: solvent A, 0.05% TFA; solvent B, 80% acetonitrile, 0.01% TFA). Mass spectromic analysis was performed using a Thermo Finnigan LCQ mass spectrometer operated with capillary temperature at 200ºC and spray voltage of 5,000 volts. Deuterium quantification data were collected in MS1 profile mode, and peptide identification data were collected in MS2 mode. The SEQUEST software program (Thermo Finnigan) was used to identify the parent peptide ions. Identified peptides were examined to determine whether the quality of the measured isotopic envelope of peptides was sufficient to allow accurate measurement of the geometric centroid of isotopic envelopes on deuterated samples. Specialized software was used to determine deuterium content in functionally deuterated samples.

S. cerevisiae **strains**

All *S. cerevisiae* strains in this study were derived from S288C. *S. cerevisiae cells* were grown at 30ºC YPD (1% yeast extract, 2% Bacto Peptone, and 2% dextrose) or in drop-out media (0.67% yeast nitrogen base without amino acids, 2% dextrose, and the appropriate amino acid drop out mix). Transformation of DNA fragments for

recombination or plasmids was done using the lithium acetate transformation method [209]. Genomic DNA was isolated from yeast cultures using the Gentra/Purgene yeast/bact kit (Qiagen). The list of plasmids used in these studies is given in **Table 5.1**, and the list of *S. cerevisiae* strains used in these studies is given in **Table 5.2**.

Pop-in/Pop-out Integration of Mutations into the Yeast Genome

Integration of *mlh1* and *pms1* mutations into the genome was performed via the pop-in pop-out integration method [210-212]. Starting with pRDK 1808 and pRDK 1582, containing *MLH1* and *PMS1*, respectively, cloned into the pRS-306 URA3 containing vector [211], the mutations *mlh1-5 serine*, *mlh1-2 serine*, *pms1-5 serine*, *pms1-2 serine*, *mlh1-D304S*, *mlh1-D304K*, *mlh1-Q57LT59L*, *mlh1-K54C*, and *mlh1-R401AD403A* were created by site directed mutagenesis using either a Gene-Art Site Directed Mutagenesis Kit (Invitrogen) or an XL-Quik Change Site Directed Mutagenesis Kit (Agilent) according to the manufacturer's instructions. The sequence of the entire MLH1/PMS1 gene in the integration vector was verified using Sanger sequencing. The verified integration vectors are listed in **Table 5.1.** Next, the verified integration vectors were cut within the MLH1 and PMS1 sequence with Bsu36I and BbvCI, respectively, and purified with a PCR purification kit (Qiagen). The purified linear fragment was transformed into the S. cerevisiae strains RDKY 3686(*MATα ura3-52 leu2Δ1 trp1Δ63 hom3-10 his3Δ200 lys2::lnsE-A10 lys2-10A*) and RDKY 5964(*MATa ura3-52 leu2Δ1 trp1Δ63 hom3-10 his3Δ200 lys2-10A*) as indicated, and

pop-in transformants were selected on CSM-URA yeast media agar plates. Pop-in isolates were screened by PCR amplifying a DNA fragment that surrounds the mutation site and submitting the resulting fragments to Sanger sequencing. As pop-in isolates have both wild-type and mutant copies of the mutagenized region, pop-in isolates were selected on the basis of presence of both wild-type and mutant signals in the sequencing traces. To select for pop-out isolates containing only the mutant copy of the gene, verified pop-in isolates were grown overnight at 30ºC in 5 mL of non-selective YPD, and then 100 μL were plated onto CSM containing 1 mg/mL 5-fluoorotic acid (5-FOA) which selects for pop-outs. Pop-outs are those cells that have lost the *URA3* gene which is present between the mutated and wild-type copies of *MLH1/PMS1*, typically by direct repeat recombination. Pop-out isolates were then screened by PCR amplifying a DNA fragment using primers that surrounded the mutation site and submitting the resulting PCR fragments to Sanger and identifying those that only contain the mutated version of the gene. The resulting strains are listed in **Table 5.2**.

Plasmid Complementation

In order to test the ability of plasmid-borne alleles of MLH1 to complement the chromosomal deletion of *MLH1* in the S. cerevisiae strain RDKY 4236(MATα *ura3-52 leu2Δ1 trp1Δ63 hom3-10 his3Δ200 lys2-10A mlh1::hisG*), mutations were engineered using site-directed mutagenesis into pRDK 1807, which is a version of the yeast replication plasmid pRS316 containing a *URA3* selection marker and a copy of the *MLH1* gene. The presence of only the mutation of interest in the resulting plasmids was verified by Sanger sequencing. The verified complementation vectors are listed in

Table 5.1. Verified plasmids were then transformed into RDKY 4236 and selected on CSM-URA plates. The ability of the mlh1 mutations to complement the chromosomal deletion was measured by fluctuation analysis and patch tests (described below) using media that lacked uracil to maintain selection for the *URA3* marker throughout the experiment.

Mutation Rate Analysis

Fluctuation analysis was used to determine the mutation rate for the *hom3-10* and *lys2::InsE-A10* frameshift reversion assays and the *CAN1* forward mutation assay for the indicated yeast strains. Briefly, each strain is streaked out on YPD to obtain single colonies, and independent colonies are used to grow 14 overnight cultures in 10 mL YPD at 30ºC. Appropriate dilutions of cells from each culture were then plated onto YPD plates and YPD plates lacking threonine, lysine, or arginine with 60mg /L canavanine added. The number of colonies grown on each plate was scored after 3 days of incubation at 30ºC. For each strain, the average mutation rate is calculated as described by Lea and Coulson.[213]

Patch testing was also used as a semi-quantitative method to determine strain mutation rates. A strain is struck out for single colonies on YPD plates, and individual colonies are used to then draw a small ~1 inch by 1 inch "patch" onto another YPD plate. After allowing to grow for 2-3 days at 30ºC to fill in the patch, the entire plate is then replica plated onto -lysine, -threonine, and +canavanine/-arginine selective plates. These plates are then allowed to grow for 2-3 days at 30ºC, after which the plates were screened for papillation, indicating a mutator phenotype.

Purification of Mlh1-Pms1

The overproduction plasmids used in this study are based upon the pRS42x series of plasmids into which the GAL1-10 upstream activating sequence (GAL1-10 UAS), including the transcriptional start sites for the GAL1 and GAL10 genes, as a 678-nt BamHI-EcoRI fragment, was inserted into the corresponding plasmid polylinker sites, resulting into vectors pRS424-GAL (TRP1) and pRS425-GAL (LEU2) [214, 215]. Our lab has previously generated 2-micron galactose-inducible overexpression vectors for Mlh1 (pRDK573 TRP1 GAL10-MLH1) and Pms1-FLAG(pRDK1099 *LEU2 GAL10-PMS1-FLAG)* [69]. Mutations in these expression vectors were introduced by standard site-directed mutagenesis methods and are listed in Table X.

S. cerevisiae Mlh1-Pms1 was purified from 2 L of culture of the *S. cerevisiae* overproduction strain RDKY 8053 (*MATα ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL mlh1::hph*) transformed with the Mlh1 and Pms1 galactose-inducible overexpression plasmids. Cell growth and induction of Mlh1-Pms1 expression utilized a published lactate-to-galactose shift protocol [216]. The SCGL -trp -leu medium contains per liter: 1.7 g of yeast nitrogen base without amino acids and ammonium sulfate, 5 g of ammonium sulfate, 30 mL glycerol, 20 mL lactic acid, 1 g of glucose, 20g of agar for solid media, 20 mg each of adenine, uracil, histidine, proline, arginine, and methionine, and 30 mg each of isoleucine, tyrosine, and lysine, 50 mg of phenylalanine, and 100 mg each of glutamic acid, aspartic acid, valine, threonine, and serine. Prior to autoclaving, the pH of the media was adjusted to 5-6 with concentrated sodium hydroxide. The YPGL medium contains per liter: 10g yeast extract, 20 g

peptone, 30 mL glycerol, 20 mL lactic acid, 2 g of glucose, and 20 mg of adenine. Prior to autoclaving, the pH of the medium was adjusted to 5-6 with concentrated sodium hydroxide.

Cells are initially struck out on SCGL -try -leu agar plates. Colonies of the expression strain were used to inoculate 200 mL of SCGL -trp -leu media and grown overnight shaking at 30ºC. These cells were pelleted and then used to inoculate 1 L of SCGL -trp -leu media to an optical density of around 0.2 at 600nm. The cells were grown with shaking at 30ºC until the cultures reached an optical density of 1.0 at 600nm. At this point YPGL media was added in a 1:1 volume to volume ration to the cultures to induce expression. The cells were then grown with shaking at 30ºC for 16 hours. The cells were then harvested by centrifugation and immediately lysed to initiate the purification procedure. All purification steps were performed at 4ºC in buffer A (50 mM Tris, pH 8, 1 mM DTT, 10% glycerol, 0.5 M EDTA, 0.01% Igepal, 1 mM PMSF, 1 mM benzamidine, 0.5 mg/liter bestatin, and 1 mg/L each of chymostatin, pepstatin A, aprotinin, and leupeptin) with the indicated concentrations of NaCl, except that 2 mM β-mercaptoethanol was substituted for the 1 mM DTT in the buffers used to run the Heparin and FLAG antibody columns whereas all other buffers contained 1 mM DTT. The cells were lysed by seven passes through a microfluidizer (Microfluidics) in 200 mL buffer A containing 200 mM NaCl. The lysates were clarified by centrifugation for 1 hour at 15,500 rpm in a Beckman-Coulter Avanti JXN-26 Centrifuge, JA25.50 rotor, and the supernatants were loaded onto two 5 mL Hi-Trap heparin columns (GE Healthcare) connected in series. The columns were washed with Buffer A containing 200 mM NaCl and eluted in a single step of 1 M NaCl in Buffer A. The pooled Heparin

column fractions were diluted with Buffer A to obtain a final NaCl concentration of 500 mM prior to being subjected to the next column. Mlh1-Pms1-containing fractions were loaded on a 5 mL column packed with anti-FLAG antibody resin (Sigma), washed with Buffer A containing 200 mM NaCl, and step eluted with Buffer A containing 200 mg/mL FLAG peptide (Sigma). The eluted proteins were pooled and diluted to a final concentration of 100 mM NaCl, loaded onto a 1 mL SP Sepharose column (GE Healthcare), washed with Buffer A containing 100 mM NaCl, and eluted with a 15 mL linear gradient from 100 mM NaCl to 1 M NaCl in Buffer A. The SP Sepharose column fractions were diluted with Buffer A to a final NaCl concentration of 200 mM prior to being loaded onto a 1 ml HiTrap Q column (GE Healthcare) followed by elution with a 100 mM to 1 M linear NaCl gradient run in Buffer A. The HiTrap Q column fractions containing the Mlh1-Pms1 were concentrated and desalted using a Centraprep (Ultracel 30K) spin column. The resulting Mlh1-Pms1 was concentrated to around 100 μL at 200 mM NaCl Buffer A, and was frozen in liquid nitrogen and stored at −80ºC. All columns were run on an Akta Pure system (GE Healthcare). SDS-PAGE analysis indicated the resulting protein preparation was >95% pure and had an equimolar ratio of the two subunits.

Surface Plasmon Resonance

DNA substrates: Biotinylated oligonucleotides were synthesized by Midland Certified Reagent Co. Inc. (Midland, TX). All other oligonucleotides were synthesized by MWG Biotech (www.mwgbiotech.com). The GT mispair-containing substrate and the GC base pair-containing substrates were 236-nucleotide PCR-derived substrates

and were prepared in the following manner. The biotinylated “G” top strand was prepared by PCR amplification of a 236-bp product, using the forward 5′-biotinylated oligonucleotide primer (5′-Biot- ACC ATG ATT ACG CCA AGC TC) and the reverse 5′-phosphorylated oligonucleotide primer (5′-Phos-TCA CAC ATC aat tgt tat ccg ctc aca att GGG TAA CGC CAG GGT TTT C), which has the *lacO1* operator sequence incorporated (lowercase letters) with plasmid pRDK505 as template DNA. This PCR product was purified with a PCR purification kit (Qiagen) and then digested with λ-exonuclease (New England Biolabs) which is specific for the 5'-phosphorylated strand of the double stranded DNA; thus, digestion of this PCR product produced by the biotinylated G top strand. The “C” bottom strand, for the GC base pair-containing substrate, was made by amplifying the plasmid template pRDK505 with the forward 5'-phosphorylated oligonucleotide primer (5'-Phos-ACC ATG ATT ACG CCA AGC TC) and the reverse oligonucleotide primer (5'-TCA CAC ATC aat tgt tat ccg ctc aca att GGG TAA CGC CAG GGT TTT C). This PCR product was purified with a PCR purification kit (Qiagen) and digested with λ-exonuclease using the same reaction conditions described above to degrade the 5'-phosphorylated top strand and generate the bottom C strand. The “T” bottom strand for the GT mispair substrate was made by amplifying the plasmid template pRDK506 with the same primers used to create the C strand. The plasmid pRDK506 has the same sequence as pRDK505 except for a single nucleotide change that places an AT base pair instead of a GC base pair 100 bases from the 5′ end of the forward primer. This PCR product was purified with a PCR purification kit (Qiagen) and digested with λ-exonuclease using the same reaction conditions described above to degrade the 5'-phosphorylated top strand and generate

the bottom C strand. The biotinylated top G and bottom C strands were mixed together and annealed by heating to 95°C and cooling at 60°C for four hours to create the GC base pair-containing substrate, and the G and T strands were similarly annealed to create the GT mispair-containing substrate. The double stranded DNA was then purified using the Qiagen PCR purification kit (Qiagen). Quality control of the DNA substrates was performed by restriction endonuclease digestion analysis as follows. The PCR product that was derived from pRDK505 has an XhoI site is lost upon formation of the mispair with the bottom strand of pRDK506, which has a NsiI site that is lost upon formation of the mispair with the top strand from pRDK505. Therefore, GC base pair-containing substrate DNA has an XhoI site, whereas the GT mispair-containing substrate DNA has neither the XhoI site nor the NsiI site. By restriction enzyme analysis, all GT-mispair containing substrates used were greater than 90% pure. The 150-bp GT-mispair containing substrate and GC base pair-containing substrate substrates used in the immunoprecipitation assays were similarly constructed except that 5' biotinylated or phosphorylated oligonucleotide forward primer had the sequence 5′-ACC ATG ATT ACG CCA AGC TC and the reverse primer, which was 5' phosphorylated for the construction of the top strand, had the sequence 5′-TAC GAC TCA CTA TAG GGC G were used for PCR amplification of the template DNAs.

Running Buffer: The running buffer used in these experiments contained of 25 mM Tris, pH 8.0, 110 mM NaCl, 0.5 mM dithiothreitol, 2% glycerol, 0.05% IGEPAL CA-630 (Nonidet P-40), 25 μM ADP and 10 mM $MgCl_2$.

Biacore Instrument: Protein-DNA and protein-protein-DNA interactions were monitored using a Biacore T100 instrument (GE Healthcare). Approximately 20 ng (100 ± 5 resonance units (RUs)) of different DNA substrates were conjugated to 2 flow cells of streptavidin-coated Biacore SA chips (GE Healthcare), and the 4th flow cell was used as an unmodified reference surface in each experiment. The experiments investigating DNA binding and sliding clamp formation were performed using 50 nM Msh2-Msh6. The experiments investigating recruitment of Mlh1-Pms1 were performed using 20 nM Msh2-Msh6 or Msh2-Msh3 and 40 nM Mlh1-Pms1. In experiments with end blocked DNA substrates, the DNA ends were blocked by including 30 nM LacI in the reactions. ATP or ADP was included in the reaction mixtures at a final concentration of 250 μM as indicated. All experiments were performed at 25°C at a flow rate of 20 μL/min, and data were collected at a frequency of 10 Hz. The data were analyzed using the BiaEvaluation v3.1 (GE Healthcare) and Prism 6 (GraphPad Software, Inc., La Jolla, CA) software.

Reconstituted Mlh1-Pms1-dependent MMR Assays

DNA Substrate: Circular DNA substrates were constructed following previously described methods using previously published mutant derivatives of pBluescript SK+ [115, 134, 217]. The substrates were constructed by annealing the combinations of mutant single-stranded circular pBluescriptSK + DNAs and linearized, denatured mutant double-stranded pBluescript SK+DNAs pRDK507 and pRDK508. Double-stranded pBluescript SK+plasmid DNA was purified using a PlasmidMaxi Kit (Qiagen) from a 100-mL overnight culture of *E. coli* XL1-Blue MRF' or TOP 10F' containing a

pBluescript SK+ mutant that was grown for 14 h in LB media containing 100μg/mL Ampicillin. Single-stranded circular pBluescript SK+DNA was prepared from E. coli. XL1-Blue MRF or TOP 10F'containing a pBluescript SK+ mutant using the helper phage rescue method described by the manufacturer (Stratagene). Fifty micrograms of double-stranded plasmid DNA was digested with 150 units of NaeI in 140-μL reactions for 2 h at 37°C. Heteroduplex substrates were constructed according to a previously published procedure [217], except with the following two modifications: (i)The denatured DNA was neutralized by adding unbuffered 2 M Tris·HCL to a concentration of 400 mM; and (ii) the DNA mixtures were annealed for 2 h at 65 °C. Then, the heteroduplex DNA was purified as follows. The heteroduplex substrate DNA mixture was digested with 660 units of ExoV (RecBCD) in a 1.8-mL reaction for 1 h according to the manufacturer's instructions (NEB) to degrade contaminating single-stranded circular and double-stranded linear DNA. The DNA was then bound to an 0.8-mL BND Cellulose column (Sigma) that was then washed with 4 mL of 300 mM NaCl in 10 mM Tris, pH 8.0, and 1 mM EDTA (TE Buffer), and the bound DNA was eluted with 4 mL of 800 mM NaCl in TE Buffer. The eluted DNA was then concentrated by ethanol precipitation and stored at −20°C in TE Buffer.

Repair Assay: Our lab has previously established the protocols for these assays. [127, 150] Proteins were diluted, if necessary, with 7.5 mM Hepes, pH 7.5, 10% (vol/vol) glycerol, 200 mM KCl, 1 mM DTT, and 0.5 mg/mL BSA. For the MMR reaction the proteins concentration was (unless otherwise noticed) 0.37 fmol of Exo1, 390 fmol of Mlh1-Pms1, 390 fmol Msh2-Msh6, 290 fmol of PCNA (PCNA trimers), 400 fmol of Polε, 80 fmol of Polδ, 220 fmol of RFC-Δ1N (or RFC), and 1,800 fmol of RPA.

Different protein mixtures were combined into 4 µL and mixed with 1 µL of 100 ng/µL of 5′ NaeI-nicked substrate DNA and 5 µL of 33 mM Tris pH 7.6, 75 mM KCl, 2.5 mM ATP, 1.66 mM glutathione, 8.3 mM $MgCl_2$, 80 µg/mL of BSA, and 200 µM each of the dNTPs. The reactions, containing a final concentration of 118 mM KCl, were then incubated at 30 °C for 3 h. The reactions were terminated by the addition of 500 mM EDTA to a concentration of 20 mM followed by the addition of 20 µL of 360 µg/mL of proteinase K and 0.4 mg/mL of glycogen. Reactions were then incubated at 55 °C for 30 min. The DNA products were then purified by phenol extraction and ethanol precipitation and digested with 2.5 units each of PstI and ScaI for 1 h at 37 °C. The DNA products were then separated by electrophoresis on a 0.8% agarose gel run in Tris-acetate-EDTA buffer (BioRad) containing 0.6 µg/mL of ethidium bromide, and the gels were photographed using a BioRad ChemiDoc XP imaging system and Image Lab software, version 4.1. The amounts of DNA in each band were quantified and the amount of DNA present in the repair-specific 1.1-kb and 1.8-kb fragments was then expressed as the percent of total DNA present in the repair-specific 1.1-kb and 1.8-kb fragments and the 2.9-kb substrate fragment. In all assays, 100% repair would be equal to repair of 100 ng or 52.75 fmol of substrate. The S.E. was calculated from the results of three or more independent experiments and is indicated by the error bars in individual figures.

Mlh1-Pms1 endonuclease assay

40 µL reactions containing 1 mM MnSO4, 20 mM Tris pH 7.5, 0.5 mM ATP 0.2 mg/mL bovine serum albumin (BSA), 2 mM DTT and 100 ng pRDK507 were incubated

at 30°C for 30 minutes. Reactions were terminated by incubation at 55°C following introduction of SDS, EDTA, glycerol and proteinase K at concentrations of 0.1%, 14 mM, 8% and 0.5 ug/ml respectively. Mlh1-Pms1, PCNA, or RFC-Δ1N were diluted to the appropriate working concentrations with a buffer comprised of 10% glycerol, 200 mM NaCl, 2 mM DTT and 20 mM Tris pH 7.5. Following termination of the reaction the samples were electrophoresed on a 0.8% agarose gel, the gel was stained with ethidium bromide, extensively destained and then the bands were quantified using a BioRad ChemiDoc XP imaging system. Serial dilutions of XhoI linearized pRDK507 ranging from 10–100 ng were used as a concentration standard for quantification.

Table 5.1. List of plasmids used.

Name		Relevant Genotype	Reference
pRS316		ampR ColE1-orf f1(+)-ori lacZa CEN6 ARSH4 URA3	Sikorski Hieter 1989.
pRDK1807		pRS316 + MLH1	Sandra Martinez
pRDK1667		pRS316 + PMS1	Cathy Smith
pRS306		ampR, ori, URA3	Sikorski Hieter 1989.
pRDK1582		pRS306 + PMS1	
pRDK1099		ampR ori 2-micron LEU2 pGal10-y-PMS1-flag (2nd PMS1 ATG)	
pRDK573		TRP1, 2-micron, pGAL10-MLH1	
pRDK1971		pRCC-K pRS42K Neomycin Amp Cas9	Generoso et al. J Microbiol Methods. 2016
pRDK 1808		pRS306 + MLH1	
pRDK1943		pRS316 + mlh1-V297S	This study
pRDK1944		pRS316 + mlh1-D299S	This study
pRDK1945		pRS316 + mlh1-A301S	This study
pRDK1946		pRS316 + mlh1-A302S	This study
pRDK1947		pRS316 + mlh1-D304S	This study
pRDK1948		pRS316 + pms1-E319S	This study
pRDK1949		pRS316 + pms1-P321S	This study
pRDK1950		pRS316 + pms1-L324S	This study
pRDK1951		pRS316 + pms1-D326S	This study
pRDK1953		pRS306 + mlh1-A297S,D299S,A301S,A302S,D304S	This study
pRDK1954		pRS306 + mlh1-D299S,D304S	This study
pRDK1955		pRS306 + pms1-E319S,P321S,L324S,D326S	This study

Table 5.1 List of Plasmids used (continued).

Name	Relevant Genotype	Reference
pRDK1956	pRS306 + pms1-P321S,D326S	This study
pRDK1957	TRP1, 2-micron, pGAL10-mlh1-D304S	This study
pRDK1958	TRP1, 2-micron, pGAL10-mlh1-D304K	This study
pRDK1959	pRS306 + mlh1-D304S	This study
pRDK1960	pRS306 + mlh1-D304K	This study
pRDK1952	pRS316 + mlh1-A140E,G141A	This study
pRDK1961	pRS316 + mlh1-K54C	This study
pRDK1962	pRS316 + mlh1-Q57A, T59A	This study
pRDK1963	pRS316 + mlh1-Q57L, T59L	This study
pRDK1964	pRS316 + pms1-S138A, R139A	This study
pRDK1965	pRS316 + pms1-E53C	This study
pRDK1966	pRS316 + pms1-E56A, S58A	This study
pRDK1967	pRS316 + pms1 E56L, S58L	This study
pRDK1968	TRP1, 2-micron, pGAL1-mlh1-Q57L T59L	This study
pRDK1969	ampR ori 2-micron LEU2 pGal10-y-pms1-E56L,S58L-flag	This study
pRDK1970	pRDK1971 + mlh1-Q57LT59L gRNA sequence	This study
pRDK1972	pRDK1971 + mlh1-K52 gRNA sequence	This study
pRDK 1973	TRP1, 2-micron, pGAL1-mlh1-K54C	

Table 5.2 List of *S. cerevisiae* strains used.

Name	Relevant Genotype	Reference
RDKY 3686	Mata ura3-52 leu2Δ1 trpΔ63 his3Δ200 hom3-10 lys2::InsE-A10	Amin et al. 2001
RDKY 3688	Mata ura3-52 leu2Δ1 trpΔ63 his3Δ200 hom3-10 lys2::InsE-A10 msh2::hisG	Mendillo et al. 2009
RDKY 4236	RDKY 3686 mlh1::hisG	-
RDKY 4238	RDKY 3686 pms1::hisG	-
RDKY 5964	MATa ura3-52 leu2Δ1 trpΔ63 his3Δ200 hom3-10 lys2::InsE-A10	Hombauer et al. 2011
RDKY 9510	RDKY 4236 + pRDK1943	This study
RDKY 9513	RDKY 4236 + pRDK1944	This study
RDKY 9516	RDKY 4236 + pRDK1945	This study
RDKY 9519	RDKY 4236 + pRDK1946	This study
RDKY 9522	RDKY 4236 + pRDK1947	This study
RDKY 9537	RKDY 4236 + pRS316	This study
RDKY 9541	RDKY3686 + mlh1-V297S,D299S,A301S,A302S,D304S	This study
RDKY 9542	RDKY 3686 + mlh1-V297S,D299S,A301S,A302S,D304S	This study
RDKY 9544	RDKY3686 + pms1-E319S,P321S,L324S,D326S	This study
RDKY 9545	RDKY3686 + pms1-E319S,P321S,L324S,D326S	This study

Table 5.2 List of *S. cerevisiae* strains used (continued).

Name	Relevant Genotype	Reference
RDKH 9546	RDKY 3686 + pms1-P321S,D326S	This study
RDKY 9547	RDKY 3686 + pms1-P321S,D326S	This study
RDKY 9558	RDKY 3686 + mlh1-D299S,D304S	This study
RDKY 9559	RDKY 3686 + mlh1-D299S,D304S	This study
RDKY 9585	RDKY 4236 + pRDK1807	This study
RDKY 9593	RDKY 4238 + pRDK1667	This study
RDKY 9595	RDKY 4238 + pRS316	This study
RDKY 9611	RDKY 3686 + mlh1-D304K	This study
RDKY 9612	RDKY 3686 + mlh1-D304K	This study
RDKY 9619	RDKY 3686 + mlh1-D304S	This study
RDKY 9620	RDKY 3686 + mlh1-D304S	This study
RDKY 9626	RDKY 7544 + mlh1-D304K	This study
RDKY 9628	RDKY 7588 + mlh1-D304K	This study
RDKY 9403	RDKY 4236 + pRDK1952	This study
RDKY 9400	RDKY 4236 + pRDK1961	This study
RDKY 9634	RDKY 4236 + pRDK1962	This study
RDKY 9636	RDKY 4236 + pRDK1963	This study
RDKY 9638	RDKY 4238 + pRDK1964	This study
RDKY 9407	RDKY 4238 + pRDK1965	This study
RDKY 9408	RDKY 4238 + pRDK1966	This study
RDKY 9405	RDKY 4238 + pRDK1967	This study
RDKY 9658	RDKY 5964 + msh2::hisG	This study
RDKY 9659	RDKY 5964 + msh2::hisG	This study
RDKY 9670	RDKY 5964 + mlh1::HisG	This study

Table 5.2 List of *S. cerevisiae* strains used (continued).

Name	Relevant Genotype	Reference
RDKY 9671	RDKY 5964 + mlh1::HisG	This study
RDKY 9674	RDKY 5964 + pms1::HisG	This study
RDKY 9789	RDKY 5964 + mlh1-Q57L,T59L	This study
RDKY 9790	RDKY 5964 + mlh1-Q57L,T59L	This study
RDKY 9791	RDKY 5964 + mlh1-K54C	This study
RDKY 9792	RDKY 5964 + mlh1-K54C	This study
RDKY 9793	RDKY 8053 + pRDK1958 + pRDK1099	This study
RDKY 9794	RDKY 8053 + pRDK1968 + pRDK1099	This study
RDKY 9795	RDKY 8053 + pRDK1973 + pRDK1099	This study

References

1. Cronin, L. and S.I. Walker, *ORIGIN OF LIFE. Beyond prebiotic chemistry.* Science, 2016. **352**(6290): p. 1174-5.

2. Walker, S.I., N. Packard, and G.D. Cody, *Re-conceptualizing the origins of life.* Philos Trans A Math Phys Eng Sci, 2017. **375**(2109).

3. Spitzer, J., G.J. Pielak, and B. Poolman, *Emergence of life: Physical chemistry changes the paradigm.* Biol Direct, 2015. **10**: p. 33.

4. Lauterbur, P.C., *The spontaneous development of biology from chemistry.* Astrobiology, 2008. **8**(1): p. 3-8.

5. Szostak, J.W., *The Narrow Road to the Deep Past: In Search of the Chemistry of the Origin of Life.* Angew Chem Int Ed Engl, 2017. **56**(37): p. 11037-11043.

6. Chatterjee, N. and G.C. Walker, *Mechanisms of DNA damage, repair, and mutagenesis.* Environ Mol Mutagen, 2017. **58**(5): p. 235-263.

7. Kolodner, R.D. and G.T. Marsischky, *Eukaryotic DNA mismatch repair.* Curr Opin Genet Dev, 1999. **9**(1): p. 89-96.

8. Borresen, A.L., R.A. Lothe, G.I. Meling, S. Lystad, P. Morrison, J. Lipford, M.F. Kane, T.O. Rognum, and R.D. Kolodner, *Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas.* Hum Mol Genet, 1995. **4**(11): p. 2065-72.

9. Kane, M.F., M. Loda, G.M. Gaida, J. Lipman, R. Mishra, H. Goldman, J.M. Jessup, and R. Kolodner, *Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines.* Cancer Res, 1997. **57**(5): p. 808-11.

10. Peltomaki, P., *Role of DNA mismatch repair defects in the pathogenesis of human cancer.* J Clin Oncol, 2003. **21**(6): p. 1174-9.

11. Cancer Genome Atlas, N., *Comprehensive molecular characterization of human colon and rectal cancer.* Nature, 2012. **487**(7407): p. 330-7.

12. de la Chapelle, A., *Genetic predisposition to colorectal cancer.* Nat Rev Cancer, 2004. **4**(10): p. 769-80.

13. Kolodner, R.D., N.R. Hall, J. Lipford, M.F. Kane, P.T. Morrison, P.J. Finan, J. Burn, P. Chapman, C. Earabino, E. Merchant, and et al., *Structure of the human MLH1 locus and analysis of a large hereditary nonpolyposis colorectal carcinoma kindred for mlh1 mutations.* Cancer Res, 1995. **55**(2): p. 242-8.

14. Kolodner, R.D., N.R. Hall, J. Lipford, M.F. Kane, M.R. Rao, P. Morrison, L. Wirth, P.J. Finan, J. Burn, and P. Chapman, *Structure of the human MSH2 locus and analysis of two Muir-Torre kindreds for msh2 mutations.* Genomics, 1994. **24**(3): p. 516-26.

15. Hendriks, Y.M., S. Jagmohan-Changur, H.M. van der Klift, H. Morreau, M. van Puijenbroek, C. Tops, T. van Os, A. Wagner, M.G. Ausems, E. Gomez, M.H. Breuning, A.H. Brocker-Vriends, H.F. Vasen, and J.T. Wijnen, *Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome).* Gastroenterology, 2006. **130**(2): p. 312-22.

16. Lagerstedt Robinson, K., T. Liu, J. Vandrovcova, B. Halvarsson, M. Clendenning, T. Frebourg, N. Papadopoulos, K.W. Kinzler, B. Vogelstein, P. Peltomaki, R.D. Kolodner, M. Nilbert, and A. Lindblom, *Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics.* J Natl Cancer Inst, 2007. **99**(4): p. 291-9.

17. Fishel, R., M.K. Lescoe, M.R. Rao, N.G. Copeland, N.A. Jenkins, J. Garber, M. Kane, and R. Kolodner, *The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer.* Cell, 1993. **75**(5): p. 1027-38.

18. Lynch, H.T., C.L. Snyder, T.G. Shaw, C.D. Heinen, and M.P. Hitchins, *Milestones of Lynch syndrome: 1895-2015.* Nat Rev Cancer, 2015. **15**(3): p. 181-94.

19. Kunkel, T.A. and D.A. Erie, *Eukaryotic Mismatch Repair in Relation to DNA Replication.* Annu Rev Genet, 2015. **49**: p. 291-313.

20. Lujan, S.A., A.R. Clausen, A.B. Clark, H.K. MacAlpine, D.M. MacAlpine, E.P. Malc, P.A. Mieczkowski, A.B. Burkholder, D.C. Fargo, D.A. Gordenin, and T.A. Kunkel, *Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition.* Genome Res, 2014. **24**(11): p. 1751-64.

21. Loeb, L.A. and T.A. Kunkel, *Fidelity of DNA synthesis.* Annu Rev Biochem, 1982. **51**: p. 429-57.

22. Kunkel, T.A., *DNA replication fidelity.* J Biol Chem, 2004. **279**(17): p. 16895-8.

23. Kunkel, T.A., *Evolving views of DNA replication (in)fidelity.* Cold Spring Harb Symp Quant Biol, 2009. **74**: p. 91-101.

24. Buermeyer, A.B., S.M. Deschenes, S.M. Baker, and R.M. Liskay, *Mammalian DNA mismatch repair.* Annu Rev Genet, 1999. **33**: p. 533-64.

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

26. Jiricny, J., *Postreplicative mismatch repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012633.

27. Li, G.M., *Mechanisms and functions of DNA mismatch repair.* Cell Res, 2008. **18**(1): p. 85-98.

28. Kroutil, L.C., K. Register, K. Bebenek, and T.A. Kunkel, *Exonucleolytic proofreading during replication of repetitive DNA.* Biochemistry, 1996. **35**(3): p. 1046-53.

29. Liu, D., G. Keijzers, and L.J. Rasmussen, *DNA mismatch repair and its many roles in eukaryotic cells.* Mutat Res, 2017. **773**: p. 174-187.

30. Levinson, G. and G.A. Gutman, *High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K-12.* Nucleic Acids Res, 1987. **15**(13): p. 5323-38.

31. Strand, M., T.A. Prolla, R.M. Liskay, and T.D. Petes, *Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair.* Nature, 1993. **365**(6443): p. 274-6.

32. Holliday, R., *A mechanism for gene conversion in fungi.* Genet Res, 2007. **89**(5-6): p. 285-307.

33. Witkin, E.M. and N.A. Sicurella, *Pure Clones of Lactose-Negative Mutants Obtained in Escherichia Coli after Treatment with 5-Bromouracil.* J Mol Biol, 1964. **8**: p. 610-3.

34. Fishel, R., *Mismatch repair.* J Biol Chem, 2015. **290**(44): p. 26395-403.

35. Santoyo, G. and D. Romero, *Gene conversion and concerted evolution in bacterial genomes.* FEMS Microbiol Rev, 2005. **29**(2): p. 169-83.

36. Kolodner, R.D., *A personal historical view of DNA mismatch repair with an emphasis on eukaryotic DNA mismatch repair.* DNA Repair (Amst), 2016. **38**: p. 3-13.

37. Modrich, P., *Mechanisms in E. coli and Human Mismatch Repair (Nobel Lecture).* Angew Chem Int Ed Engl, 2016. **55**(30): p. 8490-501.

38. Wildenberg, J. and M. Meselson, *Mismatch repair in heteroduplex DNA.* Proc Natl Acad Sci U S A, 1975. **72**(6): p. 2202-6.

39. Wagner, R., Jr. and M. Meselson, *Repair tracts in mismatched DNA heteroduplexes.* Proc Natl Acad Sci U S A, 1976. **73**(11): p. 4135-9.

40. Glickman, B., P. van den Elsen, and M. Radman, *Induced mutagenesis in dam-mutants of Escherichia coli: a role for 6-methyladenine residues in mutation avoidance.* Mol Gen Genet, 1978. **163**(3): p. 307-12.

41. Glickman, B.W. and M. Radman, *Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction.* Proc Natl Acad Sci U S A, 1980. **77**(2): p. 1063-7.

42. McGraw, B.R. and M.G. Marinus, *Isolation and characterization of Dam+ revertants and suppressor mutations that modify secondary phenotypes of dam-3 strains of Escherichia coli K-12.* Mol Gen Genet, 1980. **178**(2): p. 309-15.

43. Siegel, E.C. and V. Bryson, *Mutator gene of Escherichia coli B.* J Bacteriol, 1967. **94**(1): p. 38-47.

44. Radman, M. and R. Wagner, *Mismatch repair in Escherichia coli.* Annu Rev Genet, 1986. **20**: p. 523-38.

45. Modrich, P., *DNA mismatch correction.* Annu Rev Biochem, 1987. **56**: p. 435-66.

46. Karran, P. and M.G. Marinus, *Mismatch correction at O6-methylguanine residues in E. coli DNA.* Nature, 1982. **296**(5860): p. 868-9.

47. Langle-Rouault, F., G. Maenhaut-Michel, and M. Radman, *GATC sequences, DNA nicks and the MutH function in Escherichia coli mismatch repair.* EMBO J, 1987. **6**(4): p. 1121-7.

48. Putnam, C.D., *Evolution of the methyl directed mismatch repair system in Escherichia coli.* DNA Repair (Amst), 2016. **38**: p. 32-41.

49. Parker, B.O. and M.G. Marinus, *Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1730-4.

50. Su, S.S. and P. Modrich, *Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs.* Proc Natl Acad Sci U S A, 1986. **83**(14): p. 5057-61.

51. Hingorani, M.M., *Mismatch binding, ADP-ATP exchange and intramolecular signaling during mismatch repair.* DNA Repair (Amst), 2016. **38**: p. 24-31.

52. Lamers, M.H., A. Perrakis, J.H. Enzlin, H.H. Winterwerp, N. de Wind, and T.K. Sixma, *The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch.* Nature, 2000. **407**(6805): p. 711-7.

53. Obmolova, G., C. Ban, P. Hsieh, and W. Yang, *Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA.* Nature, 2000. **407**(6805): p. 703-10.

54. Groothuizen, F.S. and T.K. Sixma, *The conserved molecular machinery in DNA mismatch repair enzyme structures.* DNA Repair (Amst), 2016. **38**: p. 14-23.

55. Cho, W.K., C. Jeong, D. Kim, M. Chang, K.M. Song, J. Hanne, C. Ban, R. Fishel, and J.B. Lee, *ATP alters the diffusion mechanics of MutS on mismatched DNA.* Structure, 2012. **20**(7): p. 1264-1274.

56. Jones, M., R. Wagner, and M. Radman, *Repair of a mismatch is influenced by the base composition of the surrounding nucleotide sequence.* Genetics, 1987. **115**(4): p. 605-10.

57. Mazurek, A., C.N. Johnson, M.W. Germann, and R. Fishel, *Sequence context effect for hMSH2-hMSH6 mismatch-dependent activation.* Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4177-82.

58. Kool, E.T., *Hydrogen bonding, base stacking, and steric effects in dna replication.* Annu Rev Biophys Biomol Struct, 2001. **30**: p. 1-22.

59. Rossetti, G., P.D. Dans, I. Gomez-Pinto, I. Ivani, C. Gonzalez, and M. Orozco, *The structural impact of DNA mismatches.* Nucleic Acids Res, 2015. **43**(8): p. 4309-21.

60. Vafabakhsh, R. and T. Ha, *Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization.* Science, 2012. **337**(6098): p. 1097-101.

61. Yang, W., *Poor base stacking at DNA lesions may initiate recognition by many repair proteins.* DNA Repair (Amst), 2006. **5**(6): p. 654-66.

62. Acharya, S., P.L. Foster, P. Brooks, and R. Fishel, *The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair.* Mol Cell, 2003. **12**(1): p. 233-46.

63. Snowden, T., K.S. Shim, C. Schmutte, S. Acharya, and R. Fishel, *hMSH4-hMSH5 adenosine nucleotide processing and interactions with homologous recombination machinery.* J Biol Chem, 2008. **283**(1): p. 145-54.

64. Wilson, T., S. Guerrette, and R. Fishel, *Dissociation of mismatch recognition and ATPase activity by hMSH2-hMSH3.* J Biol Chem, 1999. **274**(31): p. 21659-64.

65. Gradia, S., S. Acharya, and R. Fishel, *The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch.* Cell, 1997. **91**(7): p. 995-1005.

66. Gradia, S., D. Subramanian, T. Wilson, S. Acharya, A. Makhov, J. Griffith, and R. Fishel, *hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA.* Mol Cell, 1999. **3**(2): p. 255-61.

67. Hura, G.L., C.L. Tsai, S.A. Claridge, M.L. Mendillo, J.M. Smith, G.J. Williams, A.J. Mastroianni, A.P. Alivisatos, C.D. Putnam, R.D. Kolodner, and J.A. Tainer, *DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals.* Proc Natl Acad Sci U S A, 2013. **110**(43): p. 17308-13.

68. Hargreaves, V.V., C.D. Putnam, and R.D. Kolodner, *Engineered disulfide-forming amino acid substitutions interfere with a conformational change in the mismatch recognition complex Msh2-Msh6 required for mismatch repair.* J Biol Chem, 2012. **287**(49): p. 41232-44.

69. Hargreaves, V.V., S.S. Shell, D.J. Mazur, M.T. Hess, and R.D. Kolodner, *Interaction between the Msh2 and Msh6 nucleotide-binding sites in the Saccharomyces cerevisiae Msh2-Msh6 complex.* J Biol Chem, 2010. **285**(12): p. 9301-10.

70. Gorman, J., F. Wang, S. Redding, A.J. Plys, T. Fazio, S. Wind, E.E. Alani, and E.C. Greene, *Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair.* Proc Natl Acad Sci U S A, 2012. **109**(45): p. E3074-83.

71. Jeong, C., W.K. Cho, K.M. Song, C. Cook, T.Y. Yoon, C. Ban, R. Fishel, and J.B. Lee, *MutS switches between two fundamentally distinct clamps during mismatch repair.* Nat Struct Mol Biol, 2011. **18**(3): p. 379-85.

72. Robertson, A.B., S.R. Pattishall, E.A. Gibbons, and S.W. Matson, *MutL-catalyzed ATP hydrolysis is required at a post-UvrD loading step in methyl-directed mismatch repair.* J Biol Chem, 2006. **281**(29): p. 19949-59.

73. Galio, L., C. Bouquet, and P. Brooks, *ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL.* Nucleic Acids Res, 1999. **27**(11): p. 2325-31.

74. Spampinato, C. and P. Modrich, *The MutL ATPase is required for mismatch repair.* J Biol Chem, 2000. **275**(13): p. 9863-9.

75. Schofield, M.J., S. Nayak, T.H. Scott, C. Du, and P. Hsieh, *Interaction of Escherichia coli MutS and MutL at a DNA mismatch.* J Biol Chem, 2001. **276**(30): p. 28291-9.

76. Selmane, T., M.J. Schofield, S. Nayak, C. Du, and P. Hsieh, *Formation of a DNA mismatch repair complex mediated by ATP.* J Mol Biol, 2003. **334**(5): p. 949-65.

77. Dutta, R. and M. Inouye, *GHKL, an emergent ATPase/kinase superfamily.* Trends Biochem Sci, 2000. **25**(1): p. 24-8.

78. Ban, C., M. Junop, and W. Yang, *Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair.* Cell, 1999. **97**(1): p. 85-97.

79. Ban, C. and W. Yang, *Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis.* Cell, 1998. **95**(4): p. 541-52.

80. Namadurai, S., D. Jain, D.S. Kulkarni, C.R. Tabib, P. Friedhoff, D.N. Rao, and D.T. Nair, *The C-terminal domain of the MutL homolog from Neisseria gonorrhoeae forms an inverted homodimer.* PLoS One, 2010. **5**(10): p. e13726.

81. Pillon, M.C., J.J. Lorenowicz, M. Uckelmann, A.D. Klocko, R.R. Mitchell, Y.S. Chung, P. Modrich, G.C. Walker, L.A. Simmons, P. Friedhoff, and A. Guarne, *Structure of the endonuclease domain of MutL: unlicensed to cut.* Mol Cell, 2010. **39**(1): p. 145-51.

82. Kadyrov, F.A., L. Dzantiev, N. Constantin, and P. Modrich, *Endonucleolytic function of MutLalpha in human mismatch repair.* Cell, 2006. **126**(2): p. 297-308.

83. Kadyrov, F.A., S.F. Holmes, M.E. Arana, O.A. Lukianova, M. O'Donnell, T.A. Kunkel, and P. Modrich, *Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease.* J Biol Chem, 2007. **282**(51): p. 37181-90.

84. Kosinski, J., G. Plotz, A. Guarne, J.M. Bujnicki, and P. Friedhoff, *The PMS2 subunit of human MutLalpha contains a metal ion binding domain of the iron-dependent repressor protein family.* J Mol Biol, 2008. **382**(3): p. 610-27.

85. Mizushima, R., J.Y. Kim, I. Suetake, H. Tanaka, T. Takai, N. Kamiya, Y. Takano, Y. Mishima, S. Tajima, Y. Goto, K. Fukui, and Y.H. Lee, *NMR characterization of the interaction of the endonuclease domain of MutL with divalent metal ions and ATP.* PLoS One, 2014. **9**(6): p. e98554.

86. Guarne, A., S. Ramon-Maiques, E.M. Wolff, R. Ghirlando, X. Hu, J.H. Miller, and W. Yang, *Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair.* EMBO J, 2004. **23**(21): p. 4134-45.

87. Ahrends, R., J. Kosinski, D. Kirsch, L. Manelyte, L. Giron-Monzon, L. Hummerich, O. Schulz, B. Spengler, and P. Friedhoff, *Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry.* Nucleic Acids Res, 2006. **34**(10): p. 3169-80.

88. Welsh, K.M., A.L. Lu, S. Clark, and P. Modrich, *Isolation and characterization of the Escherichia coli mutH gene product.* J Biol Chem, 1987. **262**(32): p. 15624-9.

89. Au, K.G., K. Welsh, and P. Modrich, *Initiation of methyl-directed mismatch repair.* J Biol Chem, 1992. **267**(17): p. 12142-8.

90. Campbell, J.L. and N. Kleckner, *E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork.* Cell, 1990. **62**(5): p. 967-79.

91. Ogden, G.B., M.J. Pratt, and M. Schaechter, *The replicative origin of the E. coli chromosome binds to cell membranes only when hemimethylated.* Cell, 1988. **54**(1): p. 127-35.

92. Lahue, R.S., K.G. Au, and P. Modrich, *DNA mismatch correction in a defined system.* Science, 1989. **245**(4914): p. 160-4.

93. Hall, M.C., J.R. Jordan, and S.W. Matson, *Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD.* EMBO J, 1998. **17**(5): p. 1535-41.

94. Dao, V. and P. Modrich, *Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex.* J Biol Chem, 1998. **273**(15): p. 9202-7.

95. Matson, S.W., *Escherichia coli helicase II (urvD gene product) translocates unidirectionally in a 3' to 5' direction.* J Biol Chem, 1986. **261**(22): p. 10169-75.

96. Comstock, M.J., K.D. Whitley, H. Jia, J. Sokoloski, T.M. Lohman, T. Ha, and Y.R. Chemla, *Protein structure. Direct observation of structure-function relationship in a nucleic acid-processing enzyme.* Science, 2015. **348**(6232): p. 352-4.

97. Liu, J., R. Lee, B.M. Britton, J.A. London, K. Yang, J. Hanne, J.B. Lee, and R. Fishel, *MutL sliding clamps coordinate exonuclease-independent Escherichia coli mismatch repair.* Nat Commun, 2019. **10**(1): p. 5294.

98. Hasan, A.M. and D.R. Leach, *Chromosomal directionality of DNA mismatch repair in Escherichia coli.* Proc Natl Acad Sci U S A, 2015. **112**(30): p. 9388-93.

99. Grilley, M., J. Griffith, and P. Modrich, *Bidirectional excision in methyl-directed mismatch repair.* J Biol Chem, 1993. **268**(16): p. 11830-7.

100. Burdett, V., C. Baitinger, M. Viswanathan, S.T. Lovett, and P. Modrich, *In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair.* Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6765-70.

101. Hattman, S., C. Kenny, L. Berger, and K. Pratt, *Comparative study of DNA methylation in three unicellular eucaryotes.* J Bacteriol, 1978. **135**(3): p. 1156-7.

102. Proffitt, J.H., J.R. Davie, D. Swinton, and S. Hattman, *5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA.* Mol Cell Biol, 1984. **4**(5): p. 985-8.

103. Bishop, D.K. and R.D. Kolodner, *Repair of heteroduplex plasmid DNA after transformation into Saccharomyces cerevisiae.* Mol Cell Biol, 1986. **6**(10): p. 3401-9.

104. Williamson, M.S., J.C. Game, and S. Fogel, *Meiotic gene conversion mutants in Saccharomyces cerevisiae. I. Isolation and characterization of pms1-1 and pms1-2.* Genetics, 1985. **110**(4): p. 609-46.

105. Bishop, D.K., M.S. Williamson, S. Fogel, and R.D. Kolodner, *The role of heteroduplex correction in gene conversion in Saccharomyces cerevisiae.* Nature, 1987. **328**(6128): p. 362-4.

106. Kramer, W., B. Kramer, M.S. Williamson, and S. Fogel, *Cloning and nucleotide sequence of DNA mismatch repair gene PMS1 from Saccharomyces cerevisiae: homology of PMS1 to procaryotic MutL and HexB.* J Bacteriol, 1989. **171**(10): p. 5339-46.

107. Marsischky, G.T., N. Filosi, M.F. Kane, and R. Kolodner, *Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair.* Genes Dev, 1996. **10**(4): p. 407-20.

108. Warren, J.J., T.J. Pohlhaus, A. Changela, R.R. Iyer, P.L. Modrich, and L.S. Beese, *Structure of the human MutSalpha DNA lesion recognition complex.* Mol Cell, 2007. **26**(4): p. 579-92.

109. Reenan, R.A. and R.D. Kolodner, *Isolation and characterization of two Saccharomyces cerevisiae genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins.* Genetics, 1992. **132**(4): p. 963-73.

110. Reenan, R.A. and R.D. Kolodner, *Characterization of insertion mutations in the Saccharomyces cerevisiae MSH1 and MSH2 genes: evidence for separate mitochondrial and nuclear functions.* Genetics, 1992. **132**(4): p. 975-85.

111. Drummond, J.T., G.M. Li, M.J. Longley, and P. Modrich, *Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells.* Science, 1995. **268**(5219): p. 1909-12.

112. Zhang, Y., F. Yuan, S.R. Presnell, K. Tian, Y. Gao, A.E. Tomkinson, L. Gu, and G.M. Li, *Reconstitution of 5'-directed human mismatch repair in a purified system.* Cell, 2005. **122**(5): p. 693-705.

113. Sia, E.A., R.J. Kokoska, M. Dominska, P. Greenwell, and T.D. Petes, *Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes.* Mol Cell Biol, 1997. **17**(5): p. 2851-8.

114. Srivatsan, A., N. Bowen, and R.D. Kolodner, *Mispair-specific recruitment of the Mlh1-Pms1 complex identifies repair substrates of the Saccharomyces cerevisiae Msh2-Msh3 complex.* J Biol Chem, 2014. **289**(13): p. 9352-64.

115. Shell, S.S., C.D. Putnam, and R.D. Kolodner, *Chimeric Saccharomyces cerevisiae Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins.* Proc Natl Acad Sci U S A, 2007. **104**(26): p. 10956-61.

116. Acharya, S., T. Wilson, S. Gradia, M.F. Kane, S. Guerrette, G.T. Marsischky, R. Kolodner, and R. Fishel, *hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6.* Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13629-34.

117. Palombo, F., I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, and J. Jiricny, *hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA.* Curr Biol, 1996. **6**(9): p. 1181-4.

118. Palombo, F., P. Gallinari, I. Iaccarino, T. Lettieri, M. Hughes, A. D'Arrigo, O. Truong, J.J. Hsuan, and J. Jiricny, *GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells.* Science, 1995. **268**(5219): p. 1912-4.

119. Edelmann, W., A. Umar, K. Yang, J. Heyer, M. Kucherlapati, M. Lia, B. Kneitz, E. Avdievich, K. Fan, E. Wong, G. Crouse, T. Kunkel, M. Lipkin, R.D. Kolodner, and R. Kucherlapati, *The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression.* Cancer Res, 2000. **60**(4): p. 803-7.

120. Edelmann, W., K. Yang, A. Umar, J. Heyer, K. Lau, K. Fan, W. Liedtke, P.E. Cohen, M.F. Kane, J.R. Lipford, N. Yu, G.F. Crouse, J.W. Pollard, T. Kunkel, M. Lipkin, R. Kolodner, and R. Kucherlapati, *Mutation in the mismatch repair gene Msh6 causes cancer susceptibility.* Cell, 1997. **91**(4): p. 467-77.

121. de Wind, N., M. Dekker, N. Claij, L. Jansen, Y. van Klink, M. Radman, G. Riggins, M. van der Valk, K. van't Wout, and H. te Riele, *HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions.* Nat Genet, 1999. **23**(3): p. 359-62.

122. Papadopoulos, N., N.C. Nicolaides, B. Liu, R. Parsons, C. Lengauer, F. Palombo, A. D'Arrigo, S. Markowitz, J.K. Willson, K.W. Kinzler, and et al.,

Mutations of GTBP in genetically unstable cells. Science, 1995. **268**(5219): p. 1915-7.

123. Umar, A., J.I. Risinger, W.E. Glaab, K.R. Tindall, J.C. Barrett, and T.A. Kunkel, *Functional overlap in mismatch repair by human MSH3 and MSH6.* Genetics, 1998. **148**(4): p. 1637-46.

124. Prolla, T.A., D.M. Christie, and R.M. Liskay, *Dual requirement in yeast DNA mismatch repair for MLH1 and PMS1, two homologs of the bacterial mutL gene.* Mol Cell Biol, 1994. **14**(1): p. 407-15.

125. Li, G.M. and P. Modrich, *Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs.* Proc Natl Acad Sci U S A, 1995. **92**(6): p. 1950-4.

126. Sacho, E.J., F.A. Kadyrov, P. Modrich, T.A. Kunkel, and D.A. Erie, *Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL alpha.* Mol Cell, 2008. **29**(1): p. 112-21.

127. Smith, C.E., N. Bowen, W.J.t. Graham, E.M. Goellner, A. Srivatsan, and R.D. Kolodner, *Activation of Saccharomyces cerevisiae Mlh1-Pms1 Endonuclease in a Reconstituted Mismatch Repair System.* J Biol Chem, 2015. **290**(35): p. 21580-90.

128. Gueneau, E., C. Dherin, P. Legrand, C. Tellier-Lebegue, B. Gilquin, P. Bonnesoeur, F. Londino, C. Quemener, M.H. Le Du, J.A. Marquez, M. Moutiez, M. Gondry, S. Boiteux, and J.B. Charbonnier, *Structure of the MutLalpha C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site.* Nat Struct Mol Biol, 2013. **20**(4): p. 461-8.

129. Pluciennik, A., L. Dzantiev, R.R. Iyer, N. Constantin, F.A. Kadyrov, and P. Modrich, *PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair.* Proc Natl Acad Sci U S A, 2010. **107**(37): p. 16066-71.

130. Tishkoff, D.X., A.L. Boerger, P. Bertrand, N. Filosi, G.M. Gaida, M.F. Kane, and R.D. Kolodner, *Identification and characterization of Saccharomyces cerevisiae EXO1, a gene encoding an exonuclease that interacts with MSH2.* Proc Natl Acad Sci U S A, 1997. **94**(14): p. 7487-92.

131. Lau, P.J., H. Flores-Rozas, and R.D. Kolodner, *Isolation and characterization of new proliferating cell nuclear antigen (POL30) mutator mutants that are defective in DNA mismatch repair.* Mol Cell Biol, 2002. **22**(19): p. 6669-80.

132. Xie, Y., C. Counter, and E. Alani, *Characterization of the repeat-tract instability and mutator phenotypes conferred by a Tn3 insertion in RFC1, the large subunit of the yeast clamp loader.* Genetics, 1999. **151**(2): p. 499-509.

133. Longley, M.J., A.J. Pierce, and P. Modrich, *DNA polymerase delta is required for human mismatch repair in vitro.* J Biol Chem, 1997. **272**(16): p. 10917-21.

134. Lin, Y.L., M.K. Shivji, C. Chen, R. Kolodner, R.D. Wood, and A. Dutta, *The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair.* J Biol Chem, 1998. **273**(3): p. 1453-61.

135. Bowen, N. and R.D. Kolodner, *Reconstitution of Saccharomyces cerevisiae DNA polymerase epsilon-dependent mismatch repair with purified proteins.* Proc Natl Acad Sci U S A, 2017. **114**(14): p. 3607-3612.

136. Smith, C.E., M.L. Mendillo, N. Bowen, H. Hombauer, C.S. Campbell, A. Desai, C.D. Putnam, and R.D. Kolodner, *Dominant mutations in S. cerevisiae PMS1 identify the Mlh1-Pms1 endonuclease active site and an exonuclease 1-independent mismatch repair pathway.* PLoS Genet, 2013. **9**(10): p. e1003869.

137. Amin, N.S., M.N. Nguyen, S. Oh, and R.D. Kolodner, *exo1-Dependent mutator mutations: model system for studying functional interactions in mismatch repair.* Mol Cell Biol, 2001. **21**(15): p. 5142-55.

138. Goellner, E.M., C.E. Smith, C.S. Campbell, H. Hombauer, A. Desai, C.D. Putnam, and R.D. Kolodner, *PCNA and Msh2-Msh6 activate an Mlh1-Pms1 endonuclease pathway required for Exo1-independent mismatch repair.* Mol Cell, 2014. **55**(2): p. 291-304.

139. Goellner, E.M., C.D. Putnam, and R.D. Kolodner, *Exonuclease 1-dependent and independent mismatch repair.* DNA Repair (Amst), 2015. **32**: p. 24-32.

140. Goellner, E.M., C.D. Putnam, W.J.t. Graham, C.M. Rahal, B.Z. Li, and R.D. Kolodner, *Identification of Exo1-Msh2 interaction motifs in DNA mismatch repair and new Msh2-binding partners.* Nat Struct Mol Biol, 2018. **25**(8): p. 650-659.

141. Jeon, Y., D. Kim, J.V. Martin-Lopez, R. Lee, J. Oh, J. Hanne, R. Fishel, and J.B. Lee, *Dynamic control of strand excision during human DNA mismatch repair.* Proc Natl Acad Sci U S A, 2016. **113**(12): p. 3281-6.

142. Kadyrov, F.A., J. Genschel, Y. Fang, E. Penland, W. Edelmann, and P. Modrich, *A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair.* Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8495-500.

143. Hombauer, H., C.S. Campbell, C.E. Smith, A. Desai, and R.D. Kolodner, *Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates.* Cell, 2011. **147**(5): p. 1040-53.

144. Li, F., G. Mao, D. Tong, J. Huang, L. Gu, W. Yang, and G.M. Li, *The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha.* Cell, 2013. **153**(3): p. 590-600.

145. Li, F., L. Tian, L. Gu, and G.M. Li, *Evidence that nucleosomes inhibit mismatch repair in eukaryotic cells.* J Biol Chem, 2009. **284**(48): p. 33056-61.

146. Javaid, S., M. Manohar, N. Punja, A. Mooney, J.J. Ottesen, M.G. Poirier, and R. Fishel, *Nucleosome remodeling by hMSH2-hMSH6.* Mol Cell, 2009. **36**(6): p. 1086-94.

147. Kadyrova, L.Y., E.R. Blanko, and F.A. Kadyrov, *CAF-I-dependent control of degradation of the discontinuous strands during mismatch repair.* Proc Natl Acad Sci U S A, 2011. **108**(7): p. 2753-8.

148. Schopf, B., S. Bregenhorn, J.P. Quivy, F.A. Kadyrov, G. Almouzni, and J. Jiricny, *Interplay between mismatch repair and chromatin assembly.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1895-900.

149. Constantin, N., L. Dzantiev, F.A. Kadyrov, and P. Modrich, *Human mismatch repair: reconstitution of a nick-directed bidirectional reaction.* J Biol Chem, 2005. **280**(48): p. 39752-61.

150. Bowen, N., C.E. Smith, A. Srivatsan, S. Willcox, J.D. Griffith, and R.D. Kolodner, *Reconstitution of long and short patch mismatch repair reactions using Saccharomyces cerevisiae proteins.* Proc Natl Acad Sci U S A, 2013. **110**(46): p. 18472-7.

151. Genschel, J. and P. Modrich, *Mechanism of 5'-directed excision in human mismatch repair.* Mol Cell, 2003. **12**(5): p. 1077-86.

152. Graham, W.J.t., C.D. Putnam, and R.D. Kolodner, *The properties of Msh2-Msh6 ATP binding mutants suggest a signal amplification mechanism in DNA mismatch repair.* J Biol Chem, 2018. **293**(47): p. 18055-18070.

153. Holmes, J., Jr., S. Clark, and P. Modrich, *Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines.* Proc Natl Acad Sci U S A, 1990. **87**(15): p. 5837-41.

154. Lau, P.J. and R.D. Kolodner, *Transfer of the MSH2.MSH6 complex from proliferating cell nuclear antigen to mispaired bases in DNA.* J Biol Chem, 2003. **278**(1): p. 14-7.

155. Sriramachandran, A.M., G. Petrosino, M. Mendez-Lago, A.J. Schafer, L.S. Batista-Nascimento, N. Zilio, and H.D. Ulrich, *Genome-wide Nucleotide-Resolution Mapping of DNA Replication Patterns, Single-Strand Breaks, and Lesions by GLOE-Seq.* Mol Cell, 2020.

156. Clark, A.B., F. Valle, K. Drotschmann, R.K. Gary, and T.A. Kunkel, *Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes.* J Biol Chem, 2000. **275**(47): p. 36498-501.

157. Kleczkowska, H.E., G. Marra, T. Lettieri, and J. Jiricny, *hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci.* Genes Dev, 2001. **15**(6): p. 724-36.

158. Flores-Rozas, H., D. Clark, and R.D. Kolodner, *Proliferating cell nuclear antigen and Msh2p-Msh6p interact to form an active mispair recognition complex.* Nat Genet, 2000. **26**(3): p. 375-8.

159. Prelich, G., C.K. Tan, M. Kostura, M.B. Mathews, A.G. So, K.M. Downey, and B. Stillman, *Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein.* Nature, 1987. **326**(6112): p. 517-20.

160. Krishna, T.S., X.P. Kong, S. Gary, P.M. Burgers, and J. Kuriyan, *Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA.* Cell, 1994. **79**(7): p. 1233-43.

161. Hombauer, H., A. Srivatsan, C.D. Putnam, and R.D. Kolodner, *Mismatch repair, but not heteroduplex rejection, is temporally coupled to DNA replication.* Science, 2011. **334**(6063): p. 1713-6.

162. Stukenberg, P.T., J. Turner, and M. O'Donnell, *An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps.* Cell, 1994. **78**(5): p. 877-87.

163. Pillon, M.C., J.H. Miller, and A. Guarne, *The endonuclease domain of MutL interacts with the beta sliding clamp.* DNA Repair (Amst), 2011. **10**(1): p. 87-93.

164. Genschel, J., L.Y. Kadyrova, R.R. Iyer, B.K. Dahal, F.A. Kadyrov, and P. Modrich, *Interaction of proliferating cell nuclear antigen with PMS2 is required for MutLalpha activation and function in mismatch repair.* Proc Natl Acad Sci U S A, 2017. **114**(19): p. 4930-4935.

165. Lujan, S.A., J.S. Williams, A.R. Clausen, A.B. Clark, and T.A. Kunkel, *Ribonucleotides are signals for mismatch repair of leading-strand replication errors.* Mol Cell, 2013. **50**(3): p. 437-43.

166. Ghodgaonkar, M.M., F. Lazzaro, M. Olivera-Pimentel, M. Artola-Boran, P. Cejka, M.A. Reijns, A.P. Jackson, P. Plevani, M. Muzi-Falconi, and J. Jiricny, *Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair.* Mol Cell, 2013. **50**(3): p. 323-32.

167. Allen-Soltero, S., S.L. Martinez, C.D. Putnam, and R.D. Kolodner, *A saccharomyces cerevisiae RNase H2 interaction network functions to suppress genome instability.* Mol Cell Biol, 2014. **34**(8): p. 1521-34.

168. Iyer, R.R., A. Pluciennik, V. Burdett, and P.L. Modrich, *DNA mismatch repair: functions and mechanisms.* Chem Rev, 2006. **106**(2): p. 302-23.

169. Blackwell, L.J., S. Wang, and P. Modrich, *DNA chain length dependence of formation and dynamics of hMutSalpha.hMutLalpha.heteroduplex complexes.* J Biol Chem, 2001. **276**(35): p. 33233-40.

170. Mendillo, M.L., D.J. Mazur, and R.D. Kolodner, *Analysis of the interaction between the Saccharomyces cerevisiae MSH2-MSH6 and MLH1-PMS1 complexes with DNA using a reversible DNA end-blocking system.* J Biol Chem, 2005. **280**(23): p. 22245-57.

171. Gavin, A.C., P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L.J. Jensen, S. Bastuck, B. Dumpelfeld, A. Edelmann, M.A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A.M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J.M. Rick, B. Kuster, P. Bork, R.B. Russell, and G. Superti-Furga, *Proteome survey reveals modularity of the yeast cell machinery.* Nature, 2006. **440**(7084): p. 631-6.

172. Gavin, A.C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M.A. Heurtier, R.R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga, *Functional organization of the yeast proteome by systematic analysis of protein complexes.* Nature, 2002. **415**(6868): p. 141-7.

173. Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki, *A comprehensive two-hybrid analysis to explore the yeast protein interactome.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4569-74.

174. Krogan, N.J., G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A.P. Tikuisis, T. Punna, J.M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M.D. Robinson, A. Paccanaro, J.E. Bray, A. Sheung, B. Beattie, D.P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M.M. Canete, J. Vlasblom, S. Wu, C. Orsi, S.R. Collins, S. Chandran, R. Haw, J.J. Rilstone, K. Gandi, N.J. Thompson, G. Musso, P. St Onge, S. Ghanny, M.H. Lam, G. Butland, A.M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J.S. Weissman, C.J. Ingles, T.R. Hughes, J. Parkinson, M. Gerstein, S.J. Wodak, A. Emili, and

J.F. Greenblatt, *Global landscape of protein complexes in the yeast Saccharomyces cerevisiae.* Nature, 2006. **440**(7084): p. 637-43.

175. Uetz, P., L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J.M. Rothberg, *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. **403**(6770): p. 623-7.

176. Plotz, G., C. Welsch, L. Giron-Monzon, P. Friedhoff, M. Albrecht, A. Piiper, R.M. Biondi, T. Lengauer, S. Zeuzem, and J. Raedle, *Mutations in the MutSalpha interaction interface of MLH1 can abolish DNA mismatch repair.* Nucleic Acids Res, 2006. **34**(22): p. 6574-86.

177. Kolodner, R.D., M.L. Mendillo, and C.D. Putnam, *Coupling distant sites in DNA during DNA mismatch repair.* Proc Natl Acad Sci U S A, 2007. **104**(32): p. 12953-4.

178. Junop, M.S., G. Obmolova, K. Rausch, P. Hsieh, and W. Yang, *Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair.* Mol Cell, 2001. **7**(1): p. 1-12.

179. Pluciennik, A. and P. Modrich, *Protein roadblocks and helix discontinuities are barriers to the initiation of mismatch repair.* Proc Natl Acad Sci U S A, 2007. **104**(31): p. 12709-13.

180. Allen, D.J., A. Makhov, M. Grilley, J. Taylor, R. Thresher, P. Modrich, and J.D. Griffith, *MutS mediates heteroduplex loop formation by a translocation mechanism.* EMBO J, 1997. **16**(14): p. 4467-76.

181. Blackwell, L.J., K.P. Bjornson, and P. Modrich, *DNA-dependent activation of the hMutSalpha ATPase.* J Biol Chem, 1998. **273**(48): p. 32049-54.

182. Fishel, R., S. Acharya, M. Berardini, T. Bocker, N. Charbonneau, A. Cranston, S. Gradia, S. Guerrette, C.D. Heinen, A. Mazurek, T. Snowden, C. Schmutte, K.S. Shim, G. Tombline, and T. Wilson, *Signaling mismatch repair: the mechanics of an adenosine-nucleotide molecular switch.* Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 217-24.

183. Qiu, R., V.C. DeRocco, C. Harris, A. Sharma, M.M. Hingorani, D.A. Erie, and K.R. Weninger, *Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling.* EMBO J, 2012. **31**(11): p. 2528-40.

184. Hall, M.C., H. Wang, D.A. Erie, and T.A. Kunkel, *High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer.* J Mol Biol, 2001. **312**(4): p. 637-47.

185. Manhart, C.M., X. Ni, M.A. White, J. Ortega, J.A. Surtees, and E. Alani, *The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans.* PLoS Biol, 2017. **15**(4): p. e2001164.

186. Liu, J., J. Hanne, B.M. Britton, J. Bennett, D. Kim, J.B. Lee, and R. Fishel, *Cascading MutS and MutL sliding clamps control DNA diffusion to activate mismatch repair.* Nature, 2016. **539**(7630): p. 583-587.

187. Mendillo, M.L., V.V. Hargreaves, J.W. Jamison, A.O. Mo, S. Li, C.D. Putnam, V.L. Woods, Jr., and R.D. Kolodner, *A conserved MutS homolog connector domain interface interacts with MutL homologs.* Proc Natl Acad Sci U S A, 2009. **106**(52): p. 22223-8.

188. Groothuizen, F.S., I. Winkler, M. Cristovao, A. Fish, H.H. Winterwerp, A. Reumer, A.D. Marx, N. Hermans, R.A. Nicholls, G.N. Murshudov, J.H. Lebbink, P. Friedhoff, and T.K. Sixma, *MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA.* Elife, 2015. **4**: p. e06744.

189. Plys, A.J., M.V. Rogacheva, E.C. Greene, and E. Alani, *The unstructured linker arms of Mlh1-Pms1 are important for interactions with DNA during mismatch repair.* J Mol Biol, 2012. **422**(2): p. 192-203.

190. Gorman, J., A.J. Plys, M.L. Visnapuu, E. Alani, and E.C. Greene, *Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice.* Nat Struct Mol Biol, 2010. **17**(8): p. 932-8.

191. Mardenborough, Y.S.N., K. Nitsenko, C. Laffeber, C. Duboc, E. Sahin, A. Quessada-Vial, H.H.K. Winterwerp, T.K. Sixma, R. Kanaar, P. Friedhoff, T.R. Strick, and J.H.G. Lebbink, *The unstructured linker arms of MutL enable GATC site incision beyond roadblocks during initiation of DNA mismatch repair.* Nucleic Acids Res, 2019. **47**(22): p. 11667-11680.

192. Kim, Y., C.M. Furman, C.M. Manhart, E. Alani, and I.J. Finkelstein, *Intrinsically disordered regions regulate both catalytic and non-catalytic activities of the MutLalpha mismatch repair complex.* Nucleic Acids Res, 2019. **47**(4): p. 1823-1835.

193. Argueso, J.L., A.W. Kijas, S. Sarin, J. Heck, M. Waase, and E. Alani, *Systematic mutagenesis of the Saccharomyces cerevisiae MLH1 gene reveals distinct roles for Mlh1p in meiotic crossing over and in vegetative and meiotic mismatch repair.* Mol Cell Biol, 2003. **23**(3): p. 873-86.

194. Dunker, A.K., C.J. Brown, J.D. Lawson, L.M. Iakoucheva, and Z. Obradovic, *Intrinsic disorder and protein function.* Biochemistry, 2002. **41**(21): p. 6573-82.

195. Uversky, V.N., *Natively unfolded proteins: a point where biology waits for physics.* Protein Sci, 2002. **11**(4): p. 739-56.

196. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.

197. Wright, P.E. and H.J. Dyson, *Intrinsically disordered proteins in cellular signalling and regulation.* Nat Rev Mol Cell Biol, 2015. **16**(1): p. 18-29.

198. Oldfield, C.J. and A.K. Dunker, *Intrinsically disordered proteins and intrinsically disordered protein regions.* Annu Rev Biochem, 2014. **83**: p. 553-84.

199. Mendillo, M.L., C.D. Putnam, A.O. Mo, J.W. Jamison, S. Li, V.L. Woods, Jr., and R.D. Kolodner, *Probing DNA- and ATP-mediated conformational changes in the MutS family of mispair recognition proteins using deuterium exchange mass spectrometry.* J Biol Chem, 2010. **285**(17): p. 13170-82.

200. Winkler, I., A.D. Marx, D. Lariviere, R.J. Heinze, M. Cristovao, A. Reumer, U. Curth, T.K. Sixma, and P. Friedhoff, *Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in Escherichia coli.* J Biol Chem, 2011. **286**(19): p. 17326-37.

201. Waterhouse, A.M., J.B. Procter, D.M. Martin, M. Clamp, and G.J. Barton, *Jalview Version 2--a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189-91.

202. Madeira, F., Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, and R. Lopez, *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic Acids Res, 2019. **47**(W1): p. W636-W641.

203. Friedhoff, P., L. Manelyte, L. Giron-Monzon, I. Winkler, F.S. Groothuizen, and T.K. Sixma, *Use of Single-Cysteine Variants for Trapping Transient States in DNA Mismatch Repair.* Methods Enzymol, 2017. **592**: p. 77-101.

204. Liu, L., M.C. Ortiz Castro, J. Rodriguez Gonzalez, M.C. Pillon, and A. Guarne, *The endonuclease domain of Bacillus subtilis MutL is functionally asymmetric.* DNA Repair (Amst), 2019. **73**: p. 1-6.

205. Black, B.E., M.A. Brock, S. Bedard, V.L. Woods, Jr., and D.W. Cleveland, *An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes.* Proc Natl Acad Sci U S A, 2007. **104**(12): p. 5008-13.

206. Brock, M., F. Fan, F.C. Mei, S. Li, C. Gessner, V.L. Woods, Jr., and X. Cheng, *Conformational analysis of Epac activation using amide hydrogen/deuterium exchange mass spectrometry.* J Biol Chem, 2007. **282**(44): p. 32256-63.

207. Melnyk, R.A., K.M. Hewitt, D.B. Lacy, H.C. Lin, C.R. Gessner, S. Li, V.L. Woods, Jr., and R.J. Collier, *Structural determinants for the binding of anthrax lethal factor to oligomeric protective antigen.* J Biol Chem, 2006. **281**(3): p. 1630-5.

208. Pantazatos, D., J.S. Kim, H.E. Klock, R.C. Stevens, I.A. Wilson, S.A. Lesley, and V.L. Woods, Jr., *Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS.* Proc Natl Acad Sci U S A, 2004. **101**(3): p. 751-6.

209. Gietz, R.D. and R.H. Schiestl, *High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method.* Nat Protoc, 2007. **2**(1): p. 31-4.

210. Scherer, S. and R.W. Davis, *Replacement of chromosome segments with altered DNA sequences constructed in vitro.* Proc Natl Acad Sci U S A, 1979. **76**(10): p. 4951-5.

211. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae.* Genetics, 1989. **122**(1): p. 19-27.

212. Rothstein, R., *Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast.* Methods Enzymol, 1991. **194**: p. 281-301.

213. Lea, D.E. and C.A. Coulson, *The distribution of the numbers of mutants in bacterial populations.* J Genet, 1949. **49**(3): p. 264-85.

214. Johnston, M. and R.W. Davis, *Sequences that regulate the divergent GAL1-GAL10 promoter in Saccharomyces cerevisiae.* Mol Cell Biol, 1984. **4**(8): p. 1440-8.

215. Christianson, T.W., R.S. Sikorski, M. Dante, J.H. Shero, and P. Hieter, *Multifunctional yeast high-copy-number shuttle vectors.* Gene, 1992. **110**(1): p. 119-22.

216. Gerik, K.J., S.L. Gary, and P.M. Burgers, *Overproduction and affinity purification of Saccharomyces cerevisiae replication factor C.* J Biol Chem, 1997. **272**(2): p. 1256-62.

217. Muster-Nassal, C. and R. Kolodner, *Mismatch correction catalyzed by cell-free extracts of Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7618-22.