Investigating the Mechanistic Basis of Mutant MutS DNA Repair Protein Malfunction in Lynch Syndrome

by

Emily Kessler
Class of 2018

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Molecular Biology & Biochemistry

Middletown, Connecticut April, 2018
ACKNOWLEDGEMENTS

First and foremost, I want to thank my wonderful parents for their continual love, support, and guidance, particularly throughout this past year. Thank you for always believing in me and for putting up with my endless late night phone calls; I could not have done this without you. To Joanna, Ethan, Misha, and Serene (forever an honorary member of 76 Lawn), your friendship and support has meant the world; I am so grateful to have you all in my life. Alex and Dacheng, thank you for always making me laugh, filling me in on the latest MB&B drama, and keeping me company during late nights in lab. Thank you to Lisa for dropping everything to come visit me when I needed you most; I am so thankful to have a friend like you. And to my day one, Amélie, I can’t believe how lucky I am to have been randomly assigned to room with you four years ago. Thank you for being there these last few weeks as we stuck out this writing process, for always having my back, and for never ceasing to make me laugh.

To the members of the Hingorani lab: Juan, thank you for your patience, guidance, and everything you have taught me while working on this project. Brandon, where do I even begin? Thank you for answering my endless questions, teaching me so much of what I know, keeping me sane throughout this process (even though you don’t believe it), and most importantly, for fueling my procrastination and caffeine addiction with frequent trips to Pi. Thank you to Helena for starting this project and setting the standards high, to Bo for the comradery while writing our theses, and to Miho for all the laughs and help when I was first learning.
To my thesis readers, Professors Ishita Mukerji and Erika Taylor, thank you for sparking my love of science; your passion and enthusiasm is contagious. And finally, to Manju, without whom this project could not have happened, thank you for convincing me to write this thesis and for all of your time, mentorship, and advice (thesis and otherwise) along the way—despite some technological difficulties. Thank you for everything you have taught me and for inspiring my passion for research. I cannot imagine my experience at Wesleyan without you.
ABSTRACT

DNA mismatch repair (MMR) is the highly conserved process responsible for identifying and correcting errors generated during DNA replication and is therefore essential for maintaining genome stability in nearly all living organisms. Consequently, MMR malfunctions result in increased spontaneous mutagenesis and the formation of cancer in mammals, highlighted by Lynch syndrome (LS), a hereditary predisposition to colorectal and other cancers. MMR is initiated by MutS, a protein responsible for recognizing post-replicative errors in DNA and signaling to downstream repair proteins. This study examines cancer-linked single amino acid mutants of the human MutS protein (using Thermus aquaticus MutS as a model system) in order to understand how these changes alter protein structure and function and thereby disrupt MMR. This work focuses on monitoring the timing and conformational dynamics of MutS as it utilizes ATP to work on DNA in order to construct a complete mechanism of action for each mutant protein. Preliminary kinetic analysis of seven T. aquaticus MutS mutants (T113R, G222D, F243S, Y244A, I400G, G434R, and L533R) homologous to human MutSα variants commonly found in LS patients was carried out to assess their effect on the MutS mechanism. Subtle changes in activity could impact how well the mutant proteins recognize errors and initiate MMR. Several fluorescence-based assays were utilized to monitor specific steps in the MutS mechanism and identify any dissimilarities between the mutant proteins and wildtype. So far, a few key differences are discernible: Y244A and I400G are compromised in their ability to form the sliding clamp conformation, a critical intermediate in downstream signaling of repair, which indicates the
uncoupling of their DNA binding and ATPase activity; T113R exhibits increased non-specific binding to DNA, potentially conferring reduced mismatch specificity; and G434R binds mismatched DNA tighter and with a higher affinity, possibly altering its initiation of MMR. This work advances our proof-of-principle study showing that detailed analysis of the structure, dynamics, and catalytic activities of individual MutS mutants is both feasible and can reveal critical information for understanding the molecular basis of Lynch syndrome.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. Materials and Methods</td>
<td>58</td>
</tr>
<tr>
<td>III. Results and Discussion</td>
<td>94</td>
</tr>
<tr>
<td>IV. Conclusions</td>
<td>158</td>
</tr>
<tr>
<td>V. Future Directions</td>
<td>159</td>
</tr>
<tr>
<td>Appendices</td>
<td>161</td>
</tr>
<tr>
<td>Appendix A. MutS sequence alignment</td>
<td>161</td>
</tr>
<tr>
<td>Appendix B. MutS primer sequences</td>
<td>164</td>
</tr>
<tr>
<td>Appendix C. DNA substrate sequences</td>
<td>165</td>
</tr>
<tr>
<td>Appendix D. List of buffers</td>
<td>166</td>
</tr>
<tr>
<td>Appendix E. Abbreviations and symbols</td>
<td>167</td>
</tr>
<tr>
<td>References</td>
<td>172</td>
</tr>
</tbody>
</table>
## List of Figures

### Introduction

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A minimal model of methyl-directed MMR in <em>E. coli</em> and other closely related bacteria</td>
<td>8</td>
</tr>
<tr>
<td>1.2</td>
<td>A minimal model of nick-directed MMR in eukaryotes</td>
<td>12</td>
</tr>
<tr>
<td>1.3</td>
<td>Crystal structure of <em>T. aquaticus</em> MutS</td>
<td>15</td>
</tr>
<tr>
<td>1.4</td>
<td>MutS forms specific contacts with heteroduplex DNA, kinking it by approximately 60°</td>
<td>17</td>
</tr>
<tr>
<td>1.5</td>
<td>Crystal structure of <em>E. coli</em> MutS in complex with the N-terminal domain of MutL</td>
<td>23</td>
</tr>
<tr>
<td>1.6</td>
<td>The ‘lever-arm’ is a long α-helix that connects domains IV and V</td>
<td>26</td>
</tr>
<tr>
<td>1.7</td>
<td>Sixteen LS-linked variants of MutS were identified for this study and mapped onto the <em>Taq</em> MutS crystal structure</td>
<td>40</td>
</tr>
<tr>
<td>1.8</td>
<td>Thr113 is located in a β-sheet in domain I and the mutation to arginine likely disrupts a hydrogen bonding interaction</td>
<td>43</td>
</tr>
<tr>
<td>1.9</td>
<td>Gly222 is located in an α-helix in domain II adjacent to several charged residues</td>
<td>45</td>
</tr>
<tr>
<td>1.10</td>
<td>Phe243 is located in a small β-sheet in the loop between domains II and III</td>
<td>47</td>
</tr>
<tr>
<td>1.11</td>
<td>Tyr244 has been proposed to form a contact potentially important for allosteric signaling</td>
<td>49</td>
</tr>
</tbody>
</table>
1.12 Ile400 is positioned in a loop between the α-helices that connect domains III and IV

1.13 Gly434 is located close to the DNA backbone at the α-helix in domain IV

1.14 Leu533 is located in the long lever arm in domain III at the three-domain junction

MATERIALS AND METHODS

Page

2.1 WT MutS and the seven mutant proteins were purified 66

2.2 WT MutS C42A/M88C was labeled with IAANS dye 71

2.3 A 2-AP fluorophore selectively reports MutS binding to +T DNA 77

2.4 Each MutS monomer was labeled with an IAANS molecule at the M88C position 80

RESULTS AND DISCUSSION

Page

3.1 Binding of WT MutS to 2-AP +T DNA 96

3.2 MutS mutants T113R, G222D, F243S, Y244A, and L533R appear to bind 2-AP +T DNA with WT-like affinity 97

3.3 MutS mutants I400G and G434R bind 2-AP +T DNA with higher affinity than WT 98

3.4 Binding kinetics of WT MutS to 2-AP +T DNA 102
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Binding kinetics of MutS Y244A, I400G, G434R, and F243S to 2-AP +T DNA show WT-like rates of association</td>
</tr>
<tr>
<td>3.6</td>
<td>Binding kinetics of MutS T113R, G222D, and L533R to 2-AP +T DNA show faster rates of association than WT</td>
</tr>
<tr>
<td>3.7</td>
<td>Binding kinetics of WT MutSIAANS to +T DNA</td>
</tr>
<tr>
<td>3.8</td>
<td>Binding kinetics of MutSIAANS Y244A, I400G, L533R, F243S to +T DNA show WT-like rates of complex formation</td>
</tr>
<tr>
<td>3.9</td>
<td>Binding kinetics of MutSIAANS T113R and G434R to +T DNA indicate a tighter initial collision complex</td>
</tr>
<tr>
<td>3.10</td>
<td>Binding kinetics of MutSIAANS G222D to +T DNA</td>
</tr>
<tr>
<td>3.11</td>
<td>Kinetics of WT MutSIAANS conformational changes driven by ATP binding and hydrolysis</td>
</tr>
<tr>
<td>3.12</td>
<td>Kinetics of conformational changes of MutS mutants driven by ATP binding and hydrolysis</td>
</tr>
<tr>
<td>3.13</td>
<td>Kinetics of MutS Y244A conformational changes reveal differential activity in the presence of ATP</td>
</tr>
<tr>
<td>3.14</td>
<td>Kinetics of the conformational changes of WT MutS bound to +T DNA in the presence of ATP</td>
</tr>
<tr>
<td>3.15</td>
<td>Kinetics of the conformational changes of MutS mutants bound to +T DNA in the presence of ATP</td>
</tr>
<tr>
<td>3.16</td>
<td>Kinetics of the conformational changes of Y244A and I400G bound to +T DNA with ATP show altered activity</td>
</tr>
<tr>
<td>3.17</td>
<td>Kinetics of WT MutS phosphate release</td>
</tr>
</tbody>
</table>
3.18 Kinetics of phosphate release by MutS Y244A and I400G

3.19 Kinetics of phosphate release by WT MutS bound to +T DNA

3.20 Kinetics of phosphate release by MutS Y244A and I400G bound to +T DNA

3.21 Seven LS-linked variants of MutS were included in this study
LIST OF TABLES

INTRODUCTION

Page

1.1 InSiGHT variant classification scheme for Lynch syndrome risk 33
1.2 Sixteen LS-linked variants of MutS were identified for this study 39

MATERIALS AND METHODS

Page

2.1 Percent ammonium sulfate needed to induce precipitation of MutS mutants 63

RESULTS AND DISCUSSION

Page

3.1 The affinity of MutS mutants for +T insertion DNA was determined under equilibrium conditions 99
3.2 The MutS concentration dependence of $k_{obs}$ yields a bimolecular association rate for +T DNA 107
3.3 The +T DNA concentration dependence of $k_i$ yields an apparent $K_D$ and $k_i(max)$ for the binding of MutSIAANS to +T DNA 118
3.4 Kinetic parameters of MutS conformational changes in the presence of ATP 128
3.5 Conformational changes in MutS complexed with a +T insertion were monitored in the presence of ATP 137
3.6 MutS phosphate release in the presence and absence of +T DNA 147
I. INTRODUCTION

Overview of DNA mismatch repair

The DNA mismatch repair (MMR) pathway is highly conserved and essential for maintaining the integrity and stability of the genome of nearly all living organisms. MMR serves as the primary guardian against errors that arise during DNA replication, significantly improving its fidelity and ensuring that the genome passed from parent to daughter cell is correctly copied. Thus, the loss of MMR function has drastic costs for cells, resulting in the accumulation of mutations and an increased propensity for cancer formation in mammals [1, 2]. The existence of MMR was first postulated in 1964 and has been extensively studied in the decades following [3, 4]. The understanding of this complex process has made great advances in recent years and the discovery of a link between mismatch repair defects and the formation of cancer considerably accelerated its investigation [1, 5].

DNA replication is the means by which cells duplicate their genome and transfer genetic information from parent to daughter cell during cell division. The structure of DNA is simple and elegant—two antiparallel polynucleotide strands form a double helix made up of four nucleic acid bases. These bases preferentially pair with Watson-Crick geometry where adenine (A) nucleotides match with thymine (T) nucleotides and guanine (G) nucleotides match with cytosine (C) nucleotides [6]. During DNA replication, the duplex is unwound and each strand of the DNA is used as a template for the synthesis of a nascent strand. DNA polymerases, the enzymes responsible for the incorporation of the correct nucleotide opposite to the template base, work only in the 5’ to 3’ direction [7]. However, differential methods of strand
synthesis allows for bidirectional replication of both DNA strands whereby one strand, the leading strand, is copied continuously in the 5’ to 3’ direction, and the other, the lagging strand, is copied by discontinuous synthesis of small pieces known as Okazaki fragments [2]. The formation of the correct nascent strand relies on the strict Watson-Crick complementary base pairing capability of polymerases. Despite very little energetic difference between Watson-Crick base pairs and other potential pairings, the active site geometry of DNA polymerases accommodates Watson-Crick pairs well, but does not readily allow for the incorporation of a mismatch between a template base and an incoming nucleotide [8, 9]. In addition to the geometric selection of base pairs, all DNA polymerases possess intrinsic or associated 3’ to 5’ exonuclease activity that is activated by the stalling of polymerases at incorrectly incorporated nucleotides, allowing for their efficient removal [2, 10]. Despite the remarkably accurate and efficient replication ability of DNA polymerases, errors typically escape the proofreading mechanism approximately every $10^3$ to $10^6$ nucleotides [11, 12]. Any modifications introduced during replication that evade the repair process are permanently incorporated into the genome and passed on to future generations [1].

Replicative errors are comprised primarily of mismatches, non-Watson-Crick base pairs, and insertion or deletion loops (IDLs), small loops of excess nucleotides that result from slippage of a polymerase on the primer or template strand [1, 13]. A mismatch could potentially form between any two bases, for example if a T is introduced in the nascent strand opposite to a G in the template strand by mistake, it would result in a G-T mismatch. During the following round of DNA replication, one
daughter strand would obtain the correct G-C base pair while the other would be left with an incorrect A-T pair. This type of change could result in the mutation of a non-coding or coding region of the DNA and accordingly affect the expression, structure, and/or function of a protein or RNA, thereby impacting cellular function. If these mutations occur in genes that are essential for the health, survival, or growth of the cell, they can manifest in disease phenotypes, including cancer [2]. Thus, effective repair of DNA damage is essential for the maintenance of genome stability.

DNA mismatch repair serves as the primary mechanism for the identification and removal of such post-replicative errors. In order to successfully initiate repair, the MMR machinery must recognize mismatches, distinguish between the template and daughter DNA strands, and direct repair to the newly replicated strand to remove the error-containing DNA and resynthesize the correct nascent strand [2]. MMR was first identified in *E. coli* and has since been found in nearly all organisms including *S. cerevisiae* and humans, indicating that the process is highly evolutionarily conserved [1]. As a result of MMR, the frequency at which errors are incorporated into the genome is reduced by up to 1,000-fold [4, 14]. Therefore, the loss of MMR function is associated with a higher risk of the increased mutagenesis within the genome. In particular, defects in MMR genes are associated with an increased susceptibility for cancer in humans, best evidenced by Lynch syndrome (LS), a genetic predisposition to several forms of cancer that is linked with the malfunction of MMR [1, 5, 15, 16]. While significant advances have been made in understanding the MMR pathway and its role in Lynch syndrome, little is known about why specific mutations in MMR genes are associated with the disease phenotype. Detailed mechanistic studies of
cancer-linked variants of MMR proteins are thus essential to gain insight into their structure-function effects on DNA repair [5].

The aim of this study was to expand the understanding of the link between mutations in the MutS mismatch repair protein and Lynch syndrome by examining the activities of several MutS point mutants and identifying any alterations in their mechanism of action that may underlie defective MMR. The following sections will review the current state of research on MMR, provide an overview of the structure and function of MutS, detail the relationship between MMR and Lynch syndrome, as well as outline the methods utilized in this study to investigate the activities of a set of MutS mutants in order to determine the basis for their pathogenicity. This work expands upon prior studies of a subset of Lynch syndrome-associated MutS mutations by investigating new variants as well as furthering analysis of previously characterized mutants. The overarching goal of this study is to establish the proof-of-principle concept that detailed biochemical analysis can provide a molecular-level understanding of Lynch syndrome and thereby inform both diagnosis and therapeutic interventions.

**Mismatch repair in *E. coli* and closely related prokaryotes**

The highly conserved MMR pathway has been well characterized in *E. coli* and related prokaryotes and serves as a foundational model for MMR in higher order organisms, including humans [14, 17]. The first protein in the pathway, MutS, is a homodimeric ATPase that is tasked with scanning along newly replicated duplex DNA in search of mismatches and short IDLs (Figure 1.1). As it searches for
mispaired bases, MutS probes the DNA nonspecifically and, upon encountering a mismatch, it forms a stable recognition complex and kinks the DNA by approximately 60° [13]. The mismatch-bound complex of MutS then recruits the downstream protein, MutL, another homodimeric ATPase, which is loaded onto the DNA. The MutS-MutL complex is capable of activating MutH, a protein whose endonuclease activity is responsible for nicking the error-containing strand either 3’ or 5’ from the mismatch [14]. This nick creates an entry point for DNA Helicase II (UvrD helicase), which is preferentially loaded in the presence of a MutS-MutL complex and is capable of unwinding the duplex DNA [14, 18]. This allows the approach of an exonuclease, which resects the error-containing strand beginning at the strand break and proceeding past the mismatch [14]. When the error is located 5’ from the site of the nick, the hydrolytic activity of the 3’ to 5’ exonuclease I excises the strand, while errors located 3’ from the nick are removed by either 5’ to 3’ exonuclease VII or Rec J exonuclease [19]. Single-stranded DNA binding proteins (SSB) bind the single-stranded template to prevent the formation of secondary structures following excision of the nascent strand. The final step involves gap repair by DNA polymerase III holoenzyme, which correctly resynthesizes the nascent strand, followed by DNA ligase sealing the nicks to produce an error-free daughter strand [13, 14].

In order to prevent the permanent introduction of a mutation into the genome, incorrectly paired bases must be identified and corrected prior to the subsequent round of replication [1]. An important component of mismatch repair is the necessity for efficient and accurate discrimination between the template strand and the newly
synthesized, error-containing, daughter strand. The mechanism of strand
discrimination in *E. coli* MMR relies on the methylation of the adenine nucleotide in
GATC sequences by the protein Dam methyltransferase [20, 21]. Prior to methylation
of the newly synthesized daughter strand, there is a lag time in which the recently
replicated nascent strand is unmethylated, whereas the template strand is methylated
at GATC sites. This transient hemimethylated state targets the activity of MutH
endonuclease directly 5’ to the GATC site of the unmodified strand, ensuring that the
newly replicated strand is excised and corrected and the template strand is unaltered
[14, 17, 21]. The presence of a single GATC methylation site up to a kilobase away
from the mismatch has been shown to be sufficient to direct effective repair [14, 17,
22]. Methyl-directed MMR has bidirectional capability, allowing the strand
discrimination signal to be recognized both 5’ and 3’ from the error [14, 17, 21].
Abnormally rapid methylation of newly replicated DNA has been shown to lead to
hypermutable within the genome due to the lack of this strand discrimination ability
[17]. Furthermore, the presence of single-strand breaks in the DNA circumvents the
need for MutH, confirming that the nicks to the nascent strand provide the signal for
strand discrimination, rather than MutH itself. Additionally, this indicates a potential
method for targeting MMR to newly repaired strands in organisms that do not have
post-replicative methylation marks broadly distributed across the genome [23].

The short-lived nature of the strand discrimination signal points to a
coordination between DNA replication and MMR [2]. It has been proposed that these
two functions are linked through the β or proliferating cell nuclear antigen (PCNA)
clamp proteins, processivity factors of replicative polymerases that stimulate the
processing of Okazaki fragments [24-26]. Several studies indicate that an association of MutS proteins with β or PCNA serves as a possible tether to recruit MMR proteins to the replication fork in order to localize repair activities [27-29].
Figure 1.1. A minimal model of methyl-directed MMR in *E. coli* and other closely related bacteria. In *E. coli*, MMR is initiated by MutS, which scans duplex DNA in search of a mismatch or insertion deletion loop (IDL). Upon encountering an error, it binds tightly and undergoes an ATP-driven conformational change that allows it to interact with downstream protein, MutL. The subsequent MutS-MutL complex is responsible for activating MutH endonuclease activity that specifically nicks the nascent strand, as identified by the absence of methylation at GATC sites. The nick creates an entry point for the helicase UvrD, which unwinds the error-containing DNA, allowing the resection of the error by an exonuclease. The template strand is subsequently bound by single stranded binding proteins (SSB) and the daughter strand is then resynthesized by the DNA polymerase III holoenzyme; the nicks are sealed by DNA ligase.
Eukaryotic mismatch repair

Eukaryotic MMR is functionally very similar to the methyl-directed MMR pathway of *E. coli* with homologous proteins and a similar order of events; however, there are some key differences [17]. Eukaryotic cells possess several MutS and MutL homologs and the combination of subunits is responsible for dictating the substrate specificity and cellular function of the proteins [30]. Like prokaryotic MMR, eukaryotic MMR is initiated by MutS homologs (MSH), of which there are eight identified thus far that are structurally similar to the bacterial protein (Figure 1.2) [20, 31]. In humans, the two homologs known to function in MMR are hMutSα (MSH2-MSH6) and hMutSβ (MSH2-MSH3), which are heterodimeric ATPases similarly tasked with searching duplex DNA for errors. MutSα is responsible for repairing base-base mismatches and small IDLs, while MutSβ is associated with the repair of longer IDLs up to sixteen nucleotides in length [13, 14]. Following mismatch identification, homologs of MutL, MutLα (human MLH1-PMS2/yeast Mlh1-Pms1) or MutLγ (human MLH1-PMS1/yeast Mlh1-Mlh3), are recruited to the site of the mismatch to form complexes with MutS homologs. Unlike in *E. coli* MMR, eukaryotic MutL homologs possess latent endonuclease activity and are thought to replace the role of MutH by nicking the nascent strand at or near the site of the mismatch [14]. Once the MutS-MutL complex is formed and the nascent strand is nicked, a DNA helicase is loaded at the nick to unwind the duplex DNA, rendering it available for the loading of the appropriate exonuclease. Depending on the location of the nick with respect to the error, a 3’ to 5’ or 5’ to 3’ exonuclease is then recruited to excise the strand. This bidirectional excision ability indicates that either the 5’ or 3’
end of a DNA break is efficient at directing repair [25]. The template is then bound by single stranded DNA binding protein, Replication Protein A (RPA), to stabilize the excision intermediate, while the gap repair is completed by DNA polymerase δ or DNA polymerase ε and the product is sealed by DNA ligase [4, 32].

The key distinguishing step between prokaryotic and eukaryotic MMR is the method of strand discrimination for the preferential nicking of the nascent strand. Unlike in *E. coli*, eukaryotes do not have a mechanism for the methylation of recently replicated DNA at GATC sites and likewise do not possess MutH homologs [4]. Consequently, the methylation pattern is not used as a strand discrimination signal; however, eukaryotic MMR still preferentially repairs the daughter strand and is capable of repairing all eight possible mismatches [20]. It has been hypothesized that instead of the methylation signal, single-strand breaks in recently replicated DNA may serve as a strong determinant for the bias of MMR to the nascent strand, known as nick-dependent MMR [14, 33]. This hypothesis is supported by the efficiency of repair of lagging strand replication products that are made up of short Okazaki fragments, indicating that the presence of nicks aids in strand discrimination [25, 34]. Additional findings suggest that the presence of nicks introduced by RNase-H2 during the removal of ribonucleotides incorporated in error may help to direct MMR in the leading strand [35]. Like the methylation signal in *E. coli*, a single strand break can sufficiently direct repair up to a kilobase away and either 3’ or 5’ from the site of the mismatch [14, 36, 37].

Nick-directed repair necessitates that MutL endonuclease activity be directed specifically to nick-containing nascent strands over fully ligated template strands. The
mechanism of this strand discrimination has been suggested to involve the replication clamp, PCNA [26]. In the presence of single-strand breaks, PCNA clamps are loaded at the break junction in a specific orientation with the proximal face toward the DNA terminus [26, 38, 39]. As described earlier, MutS and MutL homologs interact with PCNA and the resulting complex therefore also has a fixed geometry on the DNA [26, 40]. In conjunction with the presence of endonuclease activity in only one subunit of MutL, the interaction with PCNA directs MutL homolog nicking specifically to the nascent strand [1, 26, 40].
Eukaryotic MMR is initiated by MutS homologs (human MSH2-MSH6 and MSH2-MSH3) which scan duplex DNA in search of mismatches or insertion deletion loops (IDLs). Upon encountering an error, they bind tightly and undergo an ATP-driven conformational change that allows the interaction with downstream MutL homologs (human MLH1-PMS2 and MLH1-PMS1). The presence of nicks in the nascent strand and the PCNA clamp directs MutL endonuclease activity to preferentially incise the daughter strand. The nicked site creates an entry point for DNA helicase, which unwinds the DNA and the error-containing strand is resected by an exonuclease. The template strand is subsequently bound by single stranded binding protein, Replication Protein A (RPA). The DNA is then resynthesized by DNA polymerase δ or ε and the nicks are sealed by DNA ligase.
**Structure and function of MutS**

In recent years, several crystal structures of MMR proteins in different complexes with and without DNA and ATP have led to a deeper understanding of the molecular-level mechanism of MMR and the coordinated activities of the proteins. The crystal structures of *E. coli* and *Thermus aquaticus* (Taq) MutS as well as the human MutS homologs have been solved and provide great insight into how the evolutionarily conserved structure of MutS relates to its function in MMR [41, 42].

The concurrently published crystal structures of the MutS protein from *E. coli* and Taq reveal that the protein is made up of two subunits, each containing five domains, that encircle the DNA in a clamp-like dimer (Figure 1.3) [41, 42]. Domain I, the DNA binding domain, and domain IV, the clamp domain, are together responsible for the primary contacts with the DNA and encircle the DNA to form a channel. Domain II is the connector domain that interacts with MutL and is necessary for the formation of the functionally important ternary complex between MutS and MutL [43]. Domain V is the ATPase domain and also serves as the dimerization interface [13, 42]. Domain III is the lever, which connects the ATPase domain with the DNA binding domains, and is thought to play a role in the signaling between the two regions [42]. Prokaryotic MutS proteins form homodimers with two identical subunits, while eukaryotic MutS homologs are heterodimers made up of several MutS homologs. Despite the identical sequences of the prokaryotic MutS subunits, the proteins behave as functional heterodimers with asymmetric DNA binding and ATPase sites [41, 44, 45]. In MutS homologs MutSα and MutSβ, the unique subunit,
Msh6 and Msh3 respectively, are likely functionally homologous to the Subunit “A” of *Taq* MutS, while Msh2 serves a similar role to the Subunit “B” [42].

One of the crystal structures of *Taq* MutS is presented in Figure 1.3. It shows the homodimer in complex with a 21-base pair heteroduplex DNA containing a +T insertion and bound to two ADP-BeF₃, a non-hydrolyzable ATP analog, at both ATPase active sites (PDB ID: 1NNE) [42]. It is presumed that this structure likely reflects the ATP bound form of MutS in a state prior to the conformational change that enables its interaction with MutL and movement away from the error [42, 46]. Because this structure is thought to represent *Taq* MutS in the mismatch recognition state, a critical step necessary for its initiation of MMR, it was selected for analysis.
Figure 1.3. Crystal structure of *T. aquaticus* MutS. (a) The crystal structure of *Taq* MutS shows the homodimer complexed with +T DNA and ADP-BeF$_3$ bound at both ATPase sites (PDB ID: 1NNE). The Subunit “B” that is homologous to MutSα Msh2 is shown in blue and the Subunit “A” that is homologous to MutSα Msh6 is shown in grey. (b) Each MutS monomer is made up of five domains. Domain I (teal) is the mismatch binding domain, domain II (pink) is the connector domain, domain III (green) is the lever domain, domain IV (orange) is the clamp domain, and domain V (red) is the ATPase domain.
In addition to several nonspecific electrostatic interactions between domains I and IV and the DNA phosphate backbone, the structure of MutS in complex with heteroduplex DNA reveals that two mismatch-specific contacts are formed (Figure 1.4). A highly conserved Phe-X-Glu motif in domain I of one MutS subunit (subunit A in Taq MutS and Msh6 in MutSα) inserts into the minor groove to make specific contacts with the mismatch. The aromatic ring of the Phe39 residue of Taq MutS subunit A approaches the DNA from the minor groove to form specific pi stacking contacts with the mispaired base. The Glu41 residue of the same subunit forms a hydrogen bond with the inserted T and is thought to facilitate the bending of DNA, allowing for a reduction of unfavorable electrostatic interactions between the negatively-charged residue and the phosphate backbone [42, 47]. These base-specific contacts sharply kink the DNA by approximately 45-60˚ toward the major groove at the site of the error [42]. This DNA bending occurs in concert with MutS domains I rearrangement during the formation of the mismatch-specific stable complex between MutS and DNA [48, 49]. The DNA kinking and phenylalanine stacking interactions are general features of MutS-DNA complexes and are similarly observed in the crystal structures of E. coli and eukaryotic homologs of MutS, indicating their conserved role in mismatch recognition [41, 50].
Figure 1.4. MutS forms specific contacts with heteroduplex DNA, kinking it by approximately 60°. Upon encountering a mismatch, MutS utilizes its Phe-X-Glu motif to specifically recognize the error. In Taq MutS, domain I of subunit A (Msh6 in eukaryotic homologs) inserts its Phe39 residue adjacent to the mispaired base to form a pi stacking interaction as well as forming a hydrogen bond between the Glu41 residue and the error. These mismatch-specific contacts facilitate the bending of the DNA by approximately 60° to form the MutS-mismatch recognition complex [42].
The crystal structure further reveals that MutS belongs to the ATP binding cassette (ABC) superfamily containing C-terminal domain Walker A and Walker B motifs in domains V that make up two ATPase active sites within the protein [42]. These active sites are composite in nature, comprised of five residues from one subunit and one from the other subunit, potentially providing a mechanism for coordination between the two subunits [13, 42]. In the structure of the ternary complex of E. coli MutS with DNA in the presence of ADP, only a single ADP molecule was found to be bound in the mismatch binding subunit, indicating that the two ATPase sites are nonequivalent [41]. Although this contrasts the ternary complex of Taq MutS in which both ATPase sites are occupied by ADP-BeF₃, the active sites are known to be asymmetric, binding ATP with different affinities and hydrolyzing ATP sequentially [44, 51, 52]. In addition to their role as the ATPase domains, domains V serve as the primary site for dimerization between the two subunits with a C-terminal helix-turn-helix motif that is required for the dimer stability and interacts directly with the opposite subunit [53]. When the dimerization is disrupted, ATP binding and hydrolysis are diminished and the activity of the protein is severely hindered [53]. The following section aims to highlight the critical components of each step of the MutS mechanism from search to mismatch recognition to downstream signaling as MutS utilizes ATP to initiate MMR.
**Dynamic conformational changes and ATPase activity facilitate mismatch search and recognition by MutS**

Prior to its association with DNA, MutS domains I and IV exist in a dynamic equilibrium between open and closed states, as indicated by their unresolved nature in the crystal structure of apo MutS [42]. Additional FRET experiments with donor and acceptor labels located on the *Taq* MutS domains I further indicate the opening and closing of the domains from 30 to 70 Å apart, likely providing a means for DNA to enter the MutS clamp [42, 49, 54, 55]. In the presence of homoduplex DNA, the previously mobile clamp domains become increasingly ordered, closing around the DNA to keep the protein topologically linked during mismatch search [42, 49, 56]. MutS is capable of identifying errors by random 3D collision-dependent diffusion, which may be important for its ability to bypass blockades on DNA such as proteins and histones, or by 1D sliding along the helical backbone, allowing it to scan the duplex in search of mismatches [55, 57-59]. Domains I remain mobile in this 1D search mode and allow MutS to probe the DNA for reduced stability [49]. It has been proposed that MutS may rely on the local deformability of mismatch-containing regions for its recognition capabilities [41, 42]. The preferential binding of MutS to mismatched DNA may be due to the reduced energetic cost (~2-3 kcal/mol) of bending DNA containing a non-canonical base pair or modified base as opposed to Watson-Crick paired homoduplex DNA [41, 60, 61]. Additionally, MutS is known to bind homoduplex DNA with an affinity that is only approximately 10 to 20-fold lower than that for heteroduplex DNA [62-64]. This relatively small difference in
affinity poses the question of how exactly MutS confers its high specificity for mismatched DNA. The answer may lie in its dependence on ATP [44].

The ATPase cycle of MutS during mismatch search involves the asymmetric affinity of the two MutS subunits for ATP with one high affinity ($K_D \sim 1 \mu M$) site and one low affinity ($K_D \sim 30 \mu M$) site [13, 51, 52, 65, 66]. Kinetic experiments in which the Walker A and Walker B motifs of the ATPase sites of the two subunits were mutated revealed that Subunit A in prokaryotes or Msh6 in eukaryotes is the high affinity ATPase site, while subunit B or Msh2 is the low affinity site [67]. In the absence of DNA and while MutS scans the DNA in search of errors, it binds and catalyzes a rapid burst of ATP hydrolysis at the high affinity site with a slower steady state turnover rate ($k_{cat}$) limited by a step after ATP hydrolysis and phosphate ($P_i$) release. The low affinity site also binds and hydrolyzes ATP but at a much slower rate, which is indistinguishable from $k_{cat}$, potentially indicating that ATP hydrolysis at the low affinity site may be the rate limiting step. Subsequent ADP release from one or both subunits is similarly slow and may also contribute to the slow turnover rate [49, 68]. Previous studies of MutS\(\alpha\) have shown that Msh2 has a high affinity for ADP and thus in mismatch search mode, MutS likely has at least one ADP bound [48, 67]. The fast hydrolysis of ATP at the high affinity site and the slow release of ADP results in an accumulation of an ADP-bound form of MutS that is capable of specifically recognizing a mismatch [49, 69].

When MutS encounters a mismatch, an initial weak, nonspecific complex quickly forms with a similar affinity to that of MutS for matched DNA [48, 49, 68]. This complex then undergoes a transformation that involves concerted
conformational changes in MutS and the mismatched DNA, resulting in a stable complex with $10^3$-fold higher affinity than the initial complex [48, 49]. The transition from the weak collision complex to the stably bound MutS-mismatch recognition complex serves as an important proofreading mechanism that ensures MutS binds specifically to mismatched DNA before initiating downstream steps in the MMR pathway [48, 49]. Upon encountering a mismatch, MutS domains I undergo rearrangement in which the DNA binding domains move toward each other to adopt a closed conformation such that one subunit interacts directly with the mispair (see Figure 1.4) [42, 48, 55].

In addition to mismatched DNA triggering a conformational change in MutS and the DNA to form the tight binding complex, it also alters the MutS ATPase mechanism. Mismatch recognition stimulates the release of ADP and the rapid binding of ATP to both ATPase sites [69]. Subunit A or Msh6, the subunit that interacts specifically with the DNA, rapidly binds ATP, followed by a lag in which subunit B/Msh2 releases ADP and binds ATP [49]. This ADP to ATP exchange is a critical component of mismatch recognition and the signaling of downstream repair. In the presence of a mismatch, the rate of ATP hydrolysis at the high affinity site of subunit A is reduced by $\sim$30-fold and the rate-limiting step becomes a step prior to hydrolysis, stabilizing MutS in an ATP-bound state [51, 65, 68]. Following mismatch recognition, the ATP bound MutS undergoes several conformational changes. The first step involves the concerted disengagement of MutS domains I from the site of the mismatch and the unbending of the DNA. This is followed by a second conformational change where domains I are further separated and MutS fully releases
the mismatch [48, 55]. This ATP-induced conformational change results in the ‘sliding clamp’ state of MutS.

In an effort to better understand this MutS sliding clamp conformation and its interaction with MutL, the crystal structure of *E. coli* MutS covalently linked to the N-terminal domain of MutL and bound to a non-hydrolyzable ATP analog, AMP-PNP, was solved (Figure 1.5) [70]. This structure is thought to show MutS in the ATP-bound sliding clamp state. In contrast with other crystal structures of MutS bound to mismatched DNA, this structure shows that the lower DNA binding channel of MutS that forms a tight interaction with mismatches appears to be deformed [70]. The structure indicates that a second upper channel may form in which the domains I and the connector domains are rearranged and the two MutS subunits cross each other by ~30° [54, 70]. Although in this structure there is no DNA present, the domain rearrangement and apparent formation of a novel MutS channel suggests that the DNA may be translocated into a second, upper channel after mismatch recognition, which could allow for a looser topological connection between the protein and DNA [54, 70]. The presence of this looser channel is consistent with biochemical evidence suggesting that following mismatch recognition, the ATP-bound state of MutS attains a conformation in which it freely diffuses along DNA [58, 71-73]. It has been proposed that this ATP-bound mobile sliding clamp conformation of MutS is necessary to signal the activation of downstream repair proteins, however the exact order of events that leads to the downstream interactions remains unresolved [70, 74].
Figure 1.5. Crystal structure of *E. coli* MutS in complex with the N-terminal domain of MutL. The crystal structure of the complex between MutS and MutL (PDB ID: 5AKB) reveals the sliding clamp conformation of MutS in which domains IV appear to cross each other and connector domains II move outward to create an enlarged channel, which may allow for loose contact with DNA. This ATP-bound complex likely represents the conformation of MutS that is capable of sliding along DNA and is necessary for the interaction with MutL. Note that domains I were not resolved in this structure, possibly due to mobility.
As described above, kinetic, biochemical, and structural analyses of MutS have revealed a clear connection between the mismatch binding and ATPase activities. Despite their separation of ~70 Å, ATP binding and hydrolysis at domains V appears to modulate MutS interactions with DNA at the mismatch binding domains. Furthermore, the two subunits behave asymmetrically both in interactions with DNA substrates and nucleotides [42, 51]. This coupling indicates that an allosteric connection may exist as a means for transferring information across this relatively large distance. Experimental evidence suggests that nucleotide binding to the ATPase domains of MutS functions as an allosteric effector, triggering conformational changes in the protein [49, 75]. Although the exact mechanism of allostery in MutS remains unresolved, structural and computational studies have uncovered some key domains and residues that may play an important role in the signaling pathways.

The structural alignment of MutS proteins from a variety of organisms indicates the presence of several highly conserved regions. In particular, the three-domain junction where domains II, III, and V unite, designated the ‘transmitter’ region, as well as two long α-helices that link domains IV and V, known as the ‘lever-arm,’ are thought to aid in the communication between DNA binding and ATPase sites (Figure 1.6) [42]. The presence of nucleotides at domains V is thought to trigger slight conformational changes in the transmitter region which may be amplified by the lever-arm, resulting in the restructuring of the DNA binding domains [44, 76]. These two regions are therefore hypothesized to be involved in the coordination between the DNA binding and ATPase sites [42, 44]. Further computational studies suggest that MutS may utilize a mechanism of long-range
allosteric signaling in which small fluctuations in the structure allow for the efficient communication of information [75]. Recent molecular dynamics (MD) simulations indicate that the correlated motions of clusters of amino acids may propagate through a network of interconnected residues to communicate intramolecular signaling over long distances [46, 75]. These subtle perturbations in the structure potentially play a role in triggering more global conformational changes in MutS that are important for the initiation of mismatch search, successful recognition, and recruitment of downstream proteins.
Figure 1.6. The ‘lever-arm’ is a long $\alpha$-helix that connects domains IV and V.

The lever-arm (highlighted in red) has been implicated in the allosteric communication pathway of MutS and is thought to amplify the signaling between the DNA binding and ATPase domains.
Despite recent advances in resolving the MutS structure and DNA binding and ATPase mechanisms, the significance and functional outcomes of free diffusion of MutS away from the mismatch following the complex process of mismatch recognition continues to elude researchers. While it is well understood that MutS homologs undergo conformational changes to a mobile sliding clamp that can diffuse away from a mismatch in the presence of ATP, it is hotly debated whether the interaction of this MutS sliding clamp with MutL stabilizes the complex at the site of the mismatch or leads to a mobile MutS-MutL complex, and how these complexes initiate repair.

There are three conflicting models in the literature that have been proposed to explain the initiation of downstream repair events, the active translocation model, the molecular switch model, and the static transactivation model. The active translocation model originated from electron microscopy images showing that ATP hydrolysis promotes the formation of $\alpha$-shaped loops on heteroduplex DNA in which MutS is bound at the base of the loop and the mismatch is contained within the loop [17, 77]. It was proposed that a complex between MutS and MutL assembles at the site of the mismatch and translocates along the DNA utilizing ATP hydrolysis to expand the loops until a downstream strand discrimination signal is encountered [69, 76]. Similarly, the molecular switch model proposes translocation of MutS, however it suggests that following mismatch recognition, ATP binding by MutS allows the formation of a sliding clamp that releases heteroduplex DNA and slides along the DNA in a hydrolysis-independent manner to encounter downstream MMR.
components [64, 69, 76]. In this model, the mismatch serves as an ‘exchange factor’ that induces clamp formation and causes the rapid dissociation of MutS from the mismatch [78]. It has been proposed that the diffusion of MutS away from the mismatch may allow the iterative loading of multiple ATP-bound MutS sliding clamps that are then capable of interacting with MutL [69]. This model is supported by evidence that a double-strand break in the DNA between the mismatch and the nick prevents successful initiation of repair, indicating that the complex must remain in contact with the DNA [21, 79].

In contrast, the static transactivation model postulates that MutS maintains its association with the mismatch or IDL to recruit MutL [76]. The MutS-MutL complex bound at the site of the mismatch is then thought to come into contact with the strand discrimination signal by DNA bending, rather than translocation, to induce endonuclease activity [44, 80]. This model was suggested based on experimental evidence that MutS is capable of activating MutH endonuclease activity at a GATC site on an entirely separate DNA strand without a mismatch and that a physical blockage between a mismatch and the strand discrimination signal in the DNA does not prevent the successful removal of the error [44, 81]. The differences in these models may be partially explained by fundamental differences between methyl-directed repair such as in *E. coli* and MutH-independent repair such as in *Taq* and eukaryotes [74].

Despite the widely-debated nature of the MutS signaling mechanism, the majority of the recent evidence supports a mismatch-induced MutS switching to form static and/or mobile MutS-MutL clamps and that some degree of localization at the
mismatch occurs with MutL nicking preferentially near the site of the mismatch [26, 40, 74]. Structural, single molecule FRET, and bulk kinetic studies of Taq and eukaryotic homologs of MMR proteins indicate that MutL binds and traps MutS on the DNA after the mismatch induced ADP to ATP exchange, but prior to the ATP-induced conformational switch to the sliding clamp [49, 70, 74]. In line with the evidence that MutS sliding clamp formation occurs in two slow steps, MutL has been shown to interact with MutS following the first conformational change in which the MutS domains I release the DNA but prior to the second, slower conformational change to the sliding clamp [48, 70, 74]. Although some older studies found that the presence of ATP activates MutS-MutL complex diffusion away from the mismatch, several recent experiments show that the presence of MutL increases the residence time of the complex at or near the mismatch [29, 74, 82]. One possibility is that both stationary and mobile MutS-MutL clamps are able to form and that their functions may be complementary, allowing both iterative loading of several clamps at the mismatch for signal amplification as well as effective search for downstream strand discrimination signals [49, 70, 74].

*Mutations in eukaryotic MMR homologs are linked to cancer phenotypes*

The inactivation of the genes that encode MMR proteins results in increased rates of spontaneous mutation within the genome and, in mammals, a predisposition for tumor development [14, 30]. The loss of MMR function has been shown to increase the rate of acquiring mutations in important cell growth control and survival genes several hundred-fold [5, 83]. This link between deficiencies in MMR and the
development of cancer is best exemplified by Lynch syndrome (LS), the most common hereditary colorectal cancer condition. LS is characterized by a predisposition to a variety of cancers, primarily colorectal and endometrial, with tumors that exhibit impaired MMR activity [84]. LS is the most prevalent hereditary colorectal cancer predisposition syndrome, responsible for 3-4% of all colon cancer cases [5, 30]. This strong association with the colon cancer phenotype led the initial familial cancer syndrome to be named hereditary non-polyposis colon cancer (HNPCC) syndrome, however the name was changed after the recognition that patients additionally develop numerous extracolonic cancers [84]. The first indication that such a genetically inherited cancer predisposition might exist was in 1895 when a pathologist created a comprehensive record of the clustering of cancer in several families. Although the popular view at this time suggested that cancer was a disease of environmental factors, this discovery of the clustering of colon cancer was foundational for the widespread acceptance of an inherited genetic basis for cancer [84]. The investigation of the familial clustering of cancers was revisited in 1962 when Henry Lynch documented the histories of several colorectal cancer-prone families [5, 84].

In 1993, the link between LS tumors and MMR deficiencies was first identified. A hallmark of nearly all LS tumors is that they display a form of genetic instability in microsatellite regions of the genome consisting of polynucleotide repeat sequences, known as microsatellite instability (MSI) [85, 86]. These tandem-repeat sequences are associated with strand slippage by replicative polymerases, giving rise to the incorporation of IDLs that are typically repaired by canonical MMR [14, 30,
87]. This instability of repeat regions has been used as a diagnostic tool to identify
tumors associated with the loss of MMR activity [14, 84]. In addition to the MSI
phenotype in cancer cell lines from LS patients, the cells show insufficient repair of
small IDLs and single base mismatches, further confirming that LS is a disease of
defective MMR [87]. A key feature of LS is the accumulation of downstream genetic
mutations, often manifesting in MSI, leading to accelerated carcinogenesis [30, 84].

From 1993 to 1999, several autosomal dominant heterozygous germline mutations
that are causative of LS were identified in the genes of MutS homologs, MSH2 and
MSH6, and MutL homologs, MLH1 and a small subset in PMS1 [83, 88-91]. These
discoveries provided crucial insight for determining the genetic basis for LS and
identifying the role of DNA repair and genomic stability in the development of LS-
associated cancers [5, 84].

The identification of mutations in four of the major MMR genes in LS
patients has brought about a new era of molecular-level LS diagnosis, leading to the
use of DNA sequencing for the purpose of personalized medicine. This is exemplified
by the development of the InSiGHT Variant Database (www.insight-
group.org/variants/databases/) curated by the International Society for
Gastrointestinal and Hereditary Tumors, which systematically characterizes all
known cancer-causing mutations in LS patients [5, 16, 92]. The idea behind the
development of the database is that intensive management of the cancers associated
with LS significantly reduces mortality rates [92]. As of 2014, the database was
comprised of 2,730 unique MMR gene variants of which 39% are in genes encoding
for MLH1, 36% in MSH2, 19% in MSH6, and 6% in PMS2 [92]. Of these, 43% were
identified as nonsense or frameshift mutations that lead to drastic alterations in the protein such as a truncation. An additional 32% of the variants are not obviously truncating, including missense substitutions and small insertion-deletion mutations that may contribute to more minor alterations in the overall structure of the protein [92, 93]. The pathogenic significance of these mutations is not obvious, which can hinder successful diagnosis and treatment by clinicians [5]. The remaining mutations categorized in the database are made up of intronic variants that, while not in the coding regions of the proteins, are still found to be associated with LS [92, 93].

The database recently adopted a 5-class system (Table 1.1), as recommended by the International Agency for Research on Cancer that classifies variants based on their pathogenicity. This system identifies variants in class 5 as highly pathogenic while class 1 as likely non-pathogenic. Certain variants have yet to be classified and are deemed Variants of Uncertain Significance (VUS) and are placed in class 3. Within the highly pathogenic class 5, variants are split between two subclasses, class 5a and 5b. Class 5a mutations are those that are assumed to be pathogenic based on their DNA sequence alone, known to cause nonsense or frameshift alterations that lead to large genomic deletions. On the other hand, class 5b mutations are those that have been experimentally evidenced to disrupt protein function. These variants have been identified in at least two tumors with the phenotype of increased MSI or loss of MMR protein expression [92].
<table>
<thead>
<tr>
<th>InSiGHT gene variant class</th>
<th>Predictive risk of Lynch syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>4</td>
<td>Likely pathogenic</td>
</tr>
<tr>
<td>3</td>
<td>Uncertain</td>
</tr>
<tr>
<td>2</td>
<td>Likely not pathogenic</td>
</tr>
<tr>
<td>1</td>
<td>Not pathogenic</td>
</tr>
</tbody>
</table>

Table 1.1 InSiGHT variant classification scheme for Lynch syndrome risk.

Adapted from Thompson, et al 2014 [92]. The InSiGHT classes are assigned to MMR variants based on the measure of causality and/or pathogenicity of the variants from statistical models of clinical data.

While many of the variants of MMR proteins categorized in the InSiGHT database obviously lead to a loss of function due to the presence of frameshift and nonsense mutations (class 5a variants), a large proportion of the variants categorized are the result of single point mutations (classes 4 and 5b). In MutS variants, many of these missense mutations are located within the DNA binding or ATPase active sites, as well on the critical MutS-MutL interface. Thus, the effect of these mutations on MMR is thought to result from the direct interference with the activity at these sites. However, several of the single point mutants found to be highly pathogenic in humans are located in residues far from the active sites and the MutL binding site. This poses the question, how are these point mutations altering MMR such that they are highly associated with LS in patients? It is possible that some of these mutations disrupt critical structures or dynamics that play a subtler role in the activity of the proteins. Another intriguing hypothesis is that some of these residues may be critical for the allosteric signaling mechanism of the proteins, thus disrupting the necessary
communication between the DNA binding and ATPase domains or within the two subunits. The molecular-level identification of variants of MMR genes associated with LS, as well as recent advances in biochemical assays and the reconstituted MMR pathway \textit{in vitro}, allows the possibility for studying the isolated effects of individual MMR variants that are thus far uncharacterized.

\textit{Impact of mutations on MutS structure and stability}

As described above, over half of the variants identified in the InSiGHT database are in the two genes that encode for MutS\textalpha, MSH2 and MSH6 [92]. When the crystal structure of human MutS\textalpha was solved, it allowed for the structural mapping of these cancer-linked mutations [50]. This revealed a broad distribution of the variants throughout the structure, not merely in or around the active sites. This is in line with the prediction that the mutations could be disrupting several aspects of the MutS mechanism, including an interference with the DNA binding or ATPase activities, a disruption of the allosteric communication pathways, altered protein-protein interactions or dimerization interfaces, or a general loss of structural stability [50]. In order to further investigate the specific effects of the mutations on these various MutS activities, several computational studies have been carried out.

Molecular dynamics (MD) simulations were conducted on the crystal structures of \textit{E. coli} and \textit{Taq} MutS and have identified several possible communication pathways that may be involved in the long-range signaling between the protein active sites. One study identified a mechanism in which subtle alterations of the structure allow for the efficient communication of information over large
distances [75]. An additional statistical coupling analysis identified a potential
network of MutS residues that were studied by MD simulations of alanine mutants at
these sites [46]. Consistent with the finding that MutS may use slight structural
perturbations as a means of communication, these simulations found groups of amino
acids that undergo correlated motions, adding to the idea that there may be an
underlying network of amino acid residues that are important for allosteric
communication [46]. These computational studies identified several amino acid
residues in MutS that likely play an important role in the communication pathways,
several of which are additionally identified in the InSiGHT database as being linked
with LS [46].

Another recent study utilized biophysical modeling to predict the effect of
single point mutations in MSH2 on the free energy of the protein [94]. Variants that
have significantly reduced structural stability compared to wildtype (WT) may
therefore be targeted for degradation. It has been shown that a structural
destabilization of only ~3 kcal/mol is sufficient to cause the protein to be degraded by
the ubiquitin-proteasome pathway [94]. The free energy of every possible point
mutation in the MSH2 structure was calculated using biophysical models to yield a
ΔΔG value, the change in the thermodynamic folding stability of the variants with
respect to WT. From these calculations, it was determined that LS-linked variants of
MSH2 are structurally destabilized by on average ~9 kcal/mol, whereas variants that
are not known to be cancer-causing are destabilized by only ~2 kcal/mol [94]. The
reduced stability of MutS mutations associated with LS is therefore presumed to play
an important role in their pathogenicity.
This recent computational work has allowed for better prediction of the impact of cancer-linked mutations on the mechanism of MutS, however little experimental investigation of the effects of these mutants, and thereby their function in MMR, has been carried out.

**Experimental objectives**

The aim of this study was to pursue the molecular-level investigation of several variants of MutS associated with Lynch syndrome. Specifically, variants identified by the InSiGHT database as being in classes 4 and 5b that are due to single point-mutations with non-obvious effects on the MutS function were selected. The impact of these mutations on the structure and function of MutS in MMR was analyzed by detailed mechanistic studies to identify any differences from the well-characterized WT protein. The goal was to survey several variants of MutS to identify any possible mechanistic alterations, which will be followed up on in the future with a full panel of biochemical assays. In particular, the effect of the MutS mutations on the critical conformational changes that the protein undergoes in the presence of DNA and ATP was investigated. The analysis of several mutants whose DNA binding and ATPase activity has been previously characterized have been followed up to pursue hypotheses about possible alterations in conformational changes to gain further insight into their mechanism of action. An additional subset of mutants that had been thus far untested were also surveyed to determine if and how the sequence alteration affects the function of the protein and potentially predict how MMR is being disrupted. This analysis involved equilibrium and transient monitoring of the kinetics
of the mutant proteins as they bind DNA, bind and hydrolyze ATP, as well as undergo related conformational changes. The activities of the mutants were compared with the WT protein to identify any differences. This study serves to further the proof-of-principle concept that the detailed biochemical characterization of variants of MMR proteins clinically associated with the disease phenotype may reveal critical information to further the understanding of the molecular basis of LS.

The *Taq* MutS protein was chosen as a model system for the human MutSα protein based on the close relation between the human and *Taq* MMR pathways. Although it is a prokaryote, MMR in *Taq* is methylation-independent and lacks a MutH protein, making it comparable to eukaryotic MMR [95]. Additionally, *Taq* MutS is both highly sequentially as well as structurally and functionally conserved with human MutSα [42]. It recognizes mismatches and short IDLs, and despite being a homodimer, its two subunits are nonequivalent and interact asymmetrically with DNA and ATP [42, 51]. Furthermore, *Taq* is a thermophile and overexpression of its proteins in *E. coli* cells often yields milligram quantities of pure, highly stable protein [95]. The *Taq* MutS protein also contains only two native cysteine residues while the hMutSα protein has 39. This allows for the more straightforward labeling of the *Taq* protein with a fluorophore by conjugation to cysteine residues for study of conformational dynamics. Additionally, the *Taq* MutS protein has been well characterized by the Hingorani lab and thus any findings can be compared with existing data to ensure consistency as well as allow for accurate interpretation.
Variants considered for the study

An initial set of sixteen LS-linked MutSα mutants were selected from the InSiGHT database to be considered for this study (Table 1.2). Only residues that are well conserved between the human homologs and the Taq MutS protein were selected. As previously mentioned, mutations located directly in either the DNA binding or ATPase active sites were not chosen as they likely have a more obvious effect on MutS actions in MMR. Similarly, mutations located at sites on the MutS-MutL interface were not selected based on their presumed interruption of the interaction between the two proteins that is necessary for MMR function. The mutations chosen were mapped onto the corresponding residue in Taq MutS based on their structural alignment and are listed in Table 1.2 and mapped onto the Taq MutS crystal structure in Figure 1.7. From the positions of the mutations, it is evident that these residues are located far from either DNA binding or ATPase active sites, with the closest residue approximately 20 Å away from the Phe-X-Glu DNA binding motif. Several of the mutations selected for this study are located at the critical three-domain junction between domains II, III, and V that is thought to play an important role in the allosteric signaling between DNA binding and ATPase domains [42].
<table>
<thead>
<tr>
<th>Taq MutS</th>
<th>Corresponding hMutSα mutation</th>
<th>Mutation class from InSiGHT</th>
<th>Location in Taq MutS crystal structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>G144R</td>
<td>MSH2 G162R</td>
<td>5b</td>
<td>Domain II</td>
</tr>
<tr>
<td>L145D</td>
<td>MSH2 V163D</td>
<td>5b</td>
<td>Domain II</td>
</tr>
<tr>
<td>A146E</td>
<td>MSH2 G164E</td>
<td>5b</td>
<td>Domain II</td>
</tr>
<tr>
<td>L166R</td>
<td>MSH2 L187R</td>
<td>5b</td>
<td>Domain II</td>
</tr>
<tr>
<td>G284E</td>
<td>MSH2 G338E</td>
<td>5b</td>
<td>Junction of Domains II, III, V</td>
</tr>
<tr>
<td>R305S</td>
<td>MSH2 R359S</td>
<td>5b</td>
<td>Junction of Domains II, III, V</td>
</tr>
<tr>
<td>I400G</td>
<td>MSH2 V470G</td>
<td>4</td>
<td>Loop between Domains III and IV</td>
</tr>
<tr>
<td>G434R</td>
<td>MSH2 G504R</td>
<td>4</td>
<td>Domain IV</td>
</tr>
<tr>
<td>D528G</td>
<td>MSH2 D603G</td>
<td>4</td>
<td>Junction of Domains II, III, V</td>
</tr>
<tr>
<td>L602R</td>
<td>MSH2 M688R</td>
<td>5b</td>
<td>Junction of Domains II, III, V</td>
</tr>
<tr>
<td>G606R</td>
<td>MSH2 G692R</td>
<td>4</td>
<td>Junction of Domains II, III, V</td>
</tr>
<tr>
<td>T113R</td>
<td>MSH6 C514R</td>
<td>4</td>
<td>Loop between Domains I and II</td>
</tr>
<tr>
<td>G222D</td>
<td>MSH6 G686D</td>
<td>4</td>
<td>Domain II</td>
</tr>
<tr>
<td>F243S</td>
<td>MSH6 F706S</td>
<td>4</td>
<td>Loop between Domains II and III</td>
</tr>
<tr>
<td>Y244A</td>
<td>MSH6 Y709A</td>
<td>-</td>
<td>Loop between Domains II and III</td>
</tr>
<tr>
<td>L533R</td>
<td>MSH6 L1063R</td>
<td>4</td>
<td>Junction of Domains II, III, and V</td>
</tr>
</tbody>
</table>

**Table 1.2.** Sixteen LS-linked variants of MutS were identified for this study. All of the mutants selected are highly conserved with the hMutSα mutations identified in the InSiGHT database. The mutants are characterized as class 4 or 5b single point mutations that are located in regions distant from the DNA binding and ATPase active sites. Although the mutants are shown to be highly associated with LS in patients, their direct effect on MutS function is not well understood.
Figure 1.7. Sixteen LS-linked variants of MutS were identified for this study and mapped onto the *Taq* MutS crystal structure. The point mutations were identified from the InSiGHT database as being class 4 or 5b mutations and are located in MSH2 or MSH6. These residues are highlighted in red on their corresponding subunit of the *Taq* MutS crystal structure. The mutations are distributed throughout the MutS structure and located on both subunits. The mutants are primarily located in domains II, III, and IV. Several of the mutations are located at the three-domain junction between domains II, III, and V, a region thought to be critical for the allosteric communication of the protein [42, 46].
Variants included in the study

From the sixteen proposed mutations, a subset of seven were selected for specific analysis in this study. These mutants were chosen based on their high yield during purification which allowed them to be fluorescently labeled, as well as their relative distribution throughout the protein. Some of the mutants selected have been previously analyzed for differences in their DNA binding and ATPase activity. These findings were followed up with additional mechanistic studies of the conformational changes that the mutants undergo with DNA and ATP. Though the mutations made are homologous to hMutSα variants associated with LS, it is noted that the homodimeric nature of Taq MutS results in the introduction of the mutations at both MutS subunits. Below is a discussion of the seven mutants selected for this study based on their location and structural interactions in the Taq MutS crystal structure (PDB ID: 1NNE) as well as the predicted effect of the specific mutation at these sites. An account of any previous findings as well as any predicted involvement of the mutant residues in the allosteric signaling network of MutS is given [46].
**T113R** – The mutation T113R is homologous to the LS-linked hMSH6 variant, S514R. This residue is located in a β-strand in domain I at the interface between domains I and II (Figure 1.8). Thr113 is involved in a hydrogen bonding interaction with residue Asp96 located on the adjacent β-strand. This hydrogen bond likely stabilizes the β-sheet and thus its mutation, which presumably disrupts the hydrogen bond, may result in the destabilization of the MutS structure. Furthermore, the LS-associated mutation introduces a positively charged arginine residue in proximity to the negatively charged aspartic acid residue, potentially resulting in the formation of a strong electrostatic interaction between the two charged residues and possibly altering the structure/dynamics of the β-sheet and domain. Although Thr113 is within the DNA binding domain I, it is located ~16 Å away from the mismatch binding site. Therefore, this mutation is not thought to directly alter the interaction with DNA and may play a more indirect role in the protein’s ability to recognize mismatches and undergo related conformational changes.
Figure 1.8. Thr113 is located in a $\beta$-sheet in domain I and the mutation to arginine likely disrupts a hydrogen bonding interaction. The Thr113 residue (shown in red) forms a hydrogen bonding interaction with Asp96 (denoted by the yellow dotted line), which may be altered by its mutation. The introduction of a positively charged residue near the negatively charged aspartic acid may lead to the formation of a salt bridge, which could affect the structural stability of the $\beta$-sheet or the flexibility of domain I.
The mutation G222D is homologous to the LS-linked hMSH6 variant, G686D. The residue Glu222 is located in an $\alpha$-helix in domain II at the interface between domains I and II (Figure 1.9). Despite the flexibility of glycine, this residue is positioned at the center of the $\alpha$-helix. In MD simulations of the Taq MutS structure that identified residues that may be important for the allosteric signaling between the DNA binding and ATPase sites, Gly222 was determined to form an important interaction with Arg221 [46]. Near Gly222, on the loop that connects the $\alpha$-helix to the neighboring $\beta$-sheet, is Glu210. Interestingly, this negatively-charged residue forms several hydrogen bonds with the neighboring Arg221 residue thought to be important for the allosteric communication. The mutation of Gly222 to a negatively-charged aspartic acid may introduce a repulsive electrostatic interaction with the Glu210. Thus, this mutation could perturb the structure of this $\alpha$-helix or the MutS protein at a gross level or potentially disrupt the communication between the two active sites of the protein.
Gly222 is located in an α-helix in domain II adjacent to several charged residues. The Gly222 residue is located adjacent to Glu210 and Arg221. An interaction between Gly222 and Arg221 has been found to be important for the allosteric signaling of the protein. The mutation G222D may therefore disrupt critical allosteric communication. Additionally, the introduction of a negative charge next to an existing negatively-charged residue could lead to a repulsive interaction that might alter the structure of the α-helix or the overall stability of the protein.
**F243S** – The mutation F243S is homologous to hMSH6 F706S. This residue is located in a small β-sheet in the loop that connects domains II and III (Figure 1.10). The Phe243 is positioned in close proximity to Phe241. Although the two residues are approximately 5.5 Å apart, the flexibility of the loop structure could allow these residues to come close enough to form a pi stacking interaction. Pi stacking is a phenomenon that results from the arrangement of electrons in aromatic molecules such that the π-electrons are free to become delocalized and circle around the plane above and below the face of the aromatic ring, creating a partial negative charge at the ring edge and a partial positive charge at the interior of the ring. When two aromatic rings are brought into close contact, they can arrange so that the partially positive region of one ring is aligned with the partially negative region of the other. Aromatic-aromatic interactions are defined by contacts between aromatic rings whose centers are located between 4.5 and 7 Å apart at a dihedral angle of between 30 and 90° [96]. By this definition, the two phenylalanine rings, Phe243 and Phe241 may be able to form a pi stacking interaction, which typically leads to stabilization by approximately -0.6 to -1.3 kcal/mol. Thus, the mutation to F243S could disrupt the aromatic interaction and reduce the overall stability of the protein. The introduction of a polar serine residue close to the nonpolar Phe241 may add to the destabilization effect. Previous preliminary analysis of this mutant showed that its ATPase and DNA binding kinetics were identical to that of WT MutS [97]. In this study, conformational changes concerted with the DNA binding and ATPase activities were further examined to determine if differences in the protein dynamics underlie its malfunction in MMR.
Figure 1.10. Phe243 is located in a small β-sheet in the loop between domains II and III. The Phe243 could form a pi stacking interaction with the conserved Phe241 residue, located 5.5 Å away, potentially making a contact important for the communication between the two distant active sites of the protein. The F243S mutation could alter this interaction and affect the allosteric signaling or possibly the overall stability of the protein.
**Y244A** – The mutation Y244A is homologous to the hMSH6 variant Y709A and is located in the loop that connects domains II and III (Figure 1.11). Although the remainder of the variants of MutS selected for this study were identified based on their association with LS, Y244A is an exception. This mutation was selected based on the MD simulations of Taq MutS in which a network of coevolving MutS residues were identified and examined for their relevance in MutS structure/dynamics and communication between the DNA binding and ATPase sites [46]. One residue pair identified as forming a contact potentially important for allosteric signaling was Ala146 and Tyr244. Although the residues are located ~7.1 Å apart, Tyr244 is in a flexible loop, which may allow them to approach with the dynamic protein motions. These residues are located along a contiguous pathway identified as potentially connecting the DNA binding and ATPase sites [46]. Despite the lack of clinical evidence suggesting that residue Tyr244 is linked with LS, the significance of this residue pair for allosteric signaling was investigated by the mutation of Tyr244 to alanine. Tyr244 is involved in a hydrogen bonding interaction between its terminal hydroxyl group and the Glu169 sidechain located on the adjacent α-helix, which may be important for stabilization of its position and its role in MutS structure/dynamics.
Figure 1.11. Tyr244 has been proposed to form a contact potentially important for allosteric signaling. The Y244A mutation likely disrupts a hydrogen bond that forms between the sidechains of Tyr244 and Glu169. Its disruption may lead to the destabilization of this residue, which could alter its role in MutS structure/dynamics and/or allosteric signaling.
The mutation I400G is homologous to the hMSH2 variant V470G and is located in a loop between the helices that connect domains III and IV (Figure 1.12). The mutation of this residue to a glycine may increase the flexibility of the loop and possibly alter dynamics of the clamp domain. This residue was also identified through MD simulations, which postulated that it is in a cluster with neighboring residues including Pro406, Arg412, Arg498, Phe503, and Glu499, potentially important for allosteric signaling [46]. Therefore, the mutation of this residue could lead to the loss of a critical hub of allosteric communication. Previous investigation of this mutant’s activity revealed that it dissociates from mismatch-containing DNA faster than WT MutS [97]. This may indicate that this variant does not remain bound to heteroduplex DNA for a sufficient amount of time to undergo the conformational changes that are necessary to enable the protein to interact with MutL. This finding is consistent with the hypothesis that the Ile400 residue is important for allosteric signaling and its disruption potentially prevents the communication of the presence of mismatched DNA to the remainder of the protein. In this study I400G was further investigated by monitoring its conformational changes in the presence of +T DNA and ATP, a combination known to result in the formation of the sliding clamp in WT MutS.
Figure 1.12. Ile400 is positioned in a loop between the α-helices that connect domains III and IV. This residue was identified as having a potential role in the allosteric signaling pathway of MutS, forming a cluster with several residues including Pro406, Arg412, Arg498, Phe503, and Glu499. The mutation at this site could result in the disruption of a large allosteric network that may alter the ability of the DNA binding site to communicate the presence of mismatched DNA.
**G434R** – The mutation G434R is conserved with the hMSH2 variant G504R. Gly434 is the helix-breaking residue of the long α-helix in domain IV (Figure 1.13). This residue is located approximately 9 Å from the DNA backbone. The mutation G434R introduces a positively-charged arginine residue in the vicinity of the DNA backbone. The mutation may alter the affinity of the protein for DNA by providing an additional attractive electrostatic contact. This residue is surrounded by four other arginine residues, Arg428, Arg430, Arg432, and Arg459. Although none of these residues are shown to directly contact the DNA backbone in the crystal structure, they still may be important for providing a favorable electrostatic environment for binding duplex DNA. Introduction of an additional positive charge may cause the mutant protein to interact more favorably with homoduplex DNA, making it less specific for mismatches and potentially hindering its ability to scan along duplex DNA. Additionally, like Ile400, Gly434 has been identified by MD simulations as being in a cluster with several residues including Glu429, Arg430, Thr433, and Ile435, that is potentially important for allosteric signaling [46]. The disruption of this residue could therefore lead to altered DNA binding as well as disrupted allosteric communication. An initial study of this mutant revealed that the mutant did, in fact, associate with +T DNA two-fold faster than WT, potentially indicating that it has a greater affinity for the DNA [97]. The conformational dynamics of G434R on +T DNA was followed up in this study. As noted above, the protein may also have an increased affinity for matched DNA, which remains to be investigated further.
Figure 1.13. Gly434 is located close to the DNA backbone at the α-helix in domain IV. The introduction of a positively charged arginine residue ~9 Å from the DNA backbone may increase the affinity of the mutant for the DNA and potentially prevent it from specifically recognizing a mismatch as opposed to binding non-specifically to homoduplex DNA. This residue was proposed to be of importance as part of a cluster of residues implicated in allosteric communication. Thus, the G434R mutant potentially disrupts important signaling pathways and may lead to altered DNA binding activity.
**L533R** – The L533R mutation is homologous to the hMSH6 L1063R variant and is located in the long lever-arm α-helix in domain III (Figure 1.14). This residue is positioned in the three-domain junction between domains II, III, and V, which has been proposed to be a critical site for allosteric signaling [42]. The location of this residue in a hydrophobic pocket at the junction and its mutation from a small, hydrophobic residue to a bulky, positively-charged arginine may alter the structure and stability of this pocket. The lever-arm is recognized as playing an important role in coupling DNA binding and ATPase active sites [44, 76]. This structural perturbation could therefore interrupt a region of the protein necessary for important communication activity and may affect the structure of the protein as a whole.
Figure 1.14. Leu533 is located in the long lever arm in domain III at the three-domain junction. This residue may play a role in the overall stability of this important region as well as in the ability of the lever arm to communicate information between the DNA binding and ATPase sites of the protein. Mutation of L533R potentially leads to gross structural perturbations and may alter the allosteric communication pathway.
Experimental design

In order to assess the molecular-level effect of each of the variants on the MutS mechanism, the plan was to carry out a survey of the MutS-DNA complex formation, MutS conformational changes, and MutS ATPase activity to identify any key steps that might be altered as a result of the cancer-linked mutations. In the present study, we attempted to address these questions by an equilibrium and transient kinetics approach. In order to assess the activity of the Taq MutS mutants in the presence of DNA and ATP, fluorescent reporters were utilized to monitor the behavior of the protein and the DNA. The fluorescence intensity changes of these reporters serve as a measure of the changes in protein and DNA structure/dynamics and activity. In order to monitor the conformational dynamics of MutS, a fluorescent label on the domains I that is sensitive to changes resulting from MutS binding to mismatch-containing DNA and the formation of the sliding clamp of MutS in the presence of ATP was employed. These actions were also monitored with an on-site DNA fluorescent reporter that reflects the bending of mismatched DNA upon complex formation with MutS. Additionally, ATP hydrolysis by MutS in the presence and absence of DNA was monitored by a fluorescently labeled reporter protein.

First, the affinity of the MutS variants for error-containing DNA was assessed by measuring the dissociation constant ($K_D$) of the MutS-DNA complex. Any differences identified in the affinity of the MutS variants from the WT protein could signify alterations in the error-recognition activity of the protein. Due to the composite nature of the dissociation constant, combining the parameters from all steps of DNA binding and dissociation into one value, the individual steps were
measured to provide better resolution of the MutS DNA binding activity. Explicit transient measurement of the rate of association, $k_{\text{on}}$, serves to identify any differences in the binding rates and concentration dependence of the initial complex formation between MutS variants and DNA and the rate of dissociation, $k_{\text{off}}$, gives an idea of the lifetime of the complex. Pre-steady state analysis of the ATP hydrolysis activity of the MutS mutants in the presence and absence of DNA was used to identify changes in the ATPase mechanism, a key component of MutS mismatch recognition. Lastly, the conformational changes of the MutS variants were investigated as they bind mismatched DNA and form the ATP-induced sliding clamp, a step necessary for downstream signaling to MutL.

Each assay used for this study was designed to address key questions about individual steps in the kinetic mechanism of MutS mutant proteins during MMR and identify any alterations to the mismatch recognition and ATP-induced conformational changes in MutS involved in initiation of MMR. This set of experiments serves as an initial survey of the activities of the MutS mutants. The rates of DNA binding, ATPase activity, and conformational changes, together with future measurements of their effects on interactions with MutL and MutL activation, will be used to construct a mechanism of action for each mutant protein and to identify exactly how their function is compromised, leading to their pathogenicity. This study provides preliminary evidence that several cancer-linked mutants display differential activity that may disrupt MMR, opening a molecular-level window to the causes of LS.
II. MATERIALS AND METHODS

Nucleotides, DNA, and other reagents

For a list of primers and DNA sequences used see Appendices B and C, respectively. For a list of the contents of all buffers, see Appendix D. Buffer contents purchased from Sigma unless otherwise noted.

MutS sequence alignment

The sequence of the *T. Aquaticus* MutS protein was aligned with the sequences of *H. sapiens* MSH2 and MSH6. The sequences were obtained from the UniProt online database and aligned using the ClustalW and BoxShade software (ExPASy, Swiss Institute of Bioinformatics). The alignment is shown in Appendix A.

MutS mutagenesis

I would like to acknowledge and thank Helena Awad and Juan Liu for carrying out the mutagenesis of the T113R, G222D, F243S, I400G, G434R, and L533R plasmids used for this study. The Y244A mutagenesis was performed as follows:

*MutS C42A/M88C plasmid DNA purification* – Fresh plasmids of ampicillin resistant pET3a *Taq* MutS C42A/M88C (Integrated DNA Technologies, Inc.) were prepared in *E. coli* DH5α competent cells (Agilent). To the cells, 1 µL of the template DNA was added and incubated on ice for 30 minutes. The cells were then placed in a 42°C water bath for 2 minutes and then moved to ice for an additional 2 minutes. The cells were transferred to a culture tube containing 250 µL of sterile LB broth (Becton,
Dickinson and Co.) and shaken at 225 rpm and 37°C for 40 minutes. The cells were plated on LB-ampicillin plates (ampicillin purchased from IBI Scientific) and incubated at 37°C overnight (<16 hours).

One colony was selected from the plate and added to a culture tube containing 5 mL sterile LB broth and 100 µg/mL of ampicillin. The tubes were shaken overnight at 225 rpm and 37°C (<20 hours). The overnight culture was pelleted in a tabletop microcentrifuge at 10K rpm at room temperature. The plasmid DNA was purified from the cells using the QIAprep Spin Miniprep Kit (Qiagen). The pelleted DH5α cells were resuspended in 250 µL of Buffer P1 and transferred to a microcentrifuge tube. To each pellet, 250 µL of Buffer P2 was added and the mixture was inverted 4-6 times until the solution became clear. Then, 350 µL of Buffer N3 was added and immediately mixed by inversion of the tube 4-6 times. The lysate was then centrifuged for 10 minutes at 13K rpm in a tabletop microcentrifuge at room temperature. To a QIAprep 2.0 spin column, 800 µL of the supernatant was transferred and centrifuged for 1 minute at 13K rpm. The flow through was discarded and the column was washed by adding 750 µL Buffer PE and centrifugation for 1 minute at 13K rpm. The flow through was discarded and the spin column was centrifuged for an additional 1 minute to remove any residual wash buffer. The QIAprep column was placed in a clean microcentrifuge tube and 50 µL Buffer EB (10mM TrisCl, pH 8.5) was added to the center of the column and let stand for 1 minute. To elute the DNA plasmid, the spin column was centrifuged for 1 minute at 13K rpm. The purified DNA plasmids were quantified by UV absorption spectroscopy at 260 nm. Purified DNA plasmids were stored at -20°C.
Mutagenesis – The mutations used for this study were made to freshly purified MutS C42A/M88C plasmids using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.). The primer sequences used for mutagenesis are given in Appendix B. The reaction mixtures were assembled according the QuikChange Kit, with the 10X reaction buffer, 50 ng of the C42A/M88C MutS plasmids, 125 ng of the forward primer, 125 ng of the reverse primer, dNTPs, the QuikSolution Reagent, and the QuikChange Lightening Enzyme. The PCR reaction was performed by 1 cycle at 95°C for 3 minutes (denaturation); followed by 18 cycles at 95°C for 30 seconds, 60°C for 10 seconds, and 68°C for 4 minutes (annealing); and then 1 cycle at 68°C for 5 minutes (elongation). Following PCR, reactions were digested by addition of 2 µL of the QuikChange DpnI restriction enzyme to the reaction mixture. Mixtures were incubated at 37°C for 5 minutes.

2 µL of the DpnI-treated DNA was added to 45 µL of DH5α ultracompetent cells (Agilent) and swirled to mix. The mixture was incubated on ice for 30 minutes and then heat shocked in a water bath at 42°C for 90 seconds. Following heat shock, the cells were incubated on ice for 2 minutes. 250 µL of LB was added to each tube and the cells were incubated at 225 rpm and 37°C for 1 hour. The transformations were plated on LB-amp agar plates and incubated overnight. A single colony was selected from each plate and transferred to a culture tube containing 5 mL of sterile LB with 100 µg/mL of ampicillin. The tubes were shaken overnight at 225 rpm and 37°C (<20 hours). The mutant-containing MutS plasmids were purified using the QIAprep Spin Miniprep Kit (Qiagen) as described above in MutS C42A/M88C
plasmid DNA purification. Purified DNA plasmids were stored at -20°C. Mutations to the 8 kb MutS plasmids were confirmed with sequencing by GeneWiz, Inc. using the T7, N1, N2, and N3 primers to ensure the desired mutation was made and that no additional unwanted mutations were introduced.

MutS purification

I would like to acknowledge and thank Helena Awad and Juan Liu for carrying out the purification of the G222D, F243S, G434R, and L533R E. coli biomass used for this study. The purification of the WT, Y244A, I400G, and T113R was performed as follows:

Preparation of E. coli biomass – BL21 (DE3) E. coli competent cells

(Agilent) were used to overexpress the MutS proteins (WT, and mutants). 1 μL of the Taq MutS C42A/M88C plasmid, without mutation for WT and with the desired mutation for each mutant, was added to the BL21 (DE3) competent cells and the reaction mixture was incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 2 minutes and then placed on ice for 2 minutes. The cells were then transferred to a culture tube containing 250 μL of sterile LB and shaken at 225 rpm and 37°C for 40 minutes. The cell mixture was plated onto an LB-amp agar plate and incubated overnight at 37°C.

One colony was resuspended in 300 μL of sterile LB and 50 μL of the cell resuspension was added to 4 L flasks containing 2 L of sterile LB broth and 100 μg/mL of ampicillin, with the exception of T113R which was grown in 4 L of LB broth.
broth. The flasks were shaken at 225 rpm and 37°C for approximately 4 hours until the OD$_{600}$ reached between 0.6 and 0.8. The cells were induced by addition of IPTG to a final concentration of 0.5 mM and shaken at 225 rpm and 37°C for 3 hours. The cells were pelleted by centrifugation at 5K rpm and 4°C for 15 minutes. The pelleted cells were resuspended with 40 mL of cell resuspension buffer (see Appendix D) and transferred to a 50 mL falcon tube. The biomass was stored at -80°C.

Prior to starting the purification process, the biomass was thawed in a water bath at room temperature. The homogenizer was cooled, a water bath was set to 70°C, the 70 Ti rotor (Beckman Coulter, Inc.) was cooled to 4°C, and the L7 ultracentrifuge (Beckman Coulter, Inc.) was cooled to 4°C. Note: plastic was utilized whenever possible to minimize phosphate contamination from glass in the protein purification.

Preparation of MutS – The thawed *E. coli* biomass was added to a plastic beaker and NaCl (OmniPur) was added to a final concentration of 50 mM, DTT (Gold Biotechnology, Inc.) was added to a final concentration of 2 mM, and PMSF (G-Biosciences, Geno Technology, Inc.) was added to a final concentration of 1 mM; the mixture was stirred well. To the mixture, lyophilized lysozyme (Sigma-Aldrich, Inc.) was added at 3 mg per L of biomass grown and stirred at room temperature for 10 to 15 minutes. The cells were then lysed in an Emulsiflex-C5 homogenizer (Avestin, Inc.) with 13,000 to 17,000 psi of pressure. The cells were purged through the homogenizer four times until they were no longer opaque and returned to a falcon tube. The lysate was immediately placed in a heat bath at 70°C for 30 minutes and mixed by inversion every 10 minutes. The whole cell lysate was centrifuged at 26K
rpm and 4°C for 1 hour in an L7 Ultracentrifuge. The supernatant was decanted into a plastic beaker and transferred to the cold room.

0.2423 g of ammonium sulfate (VWR, Avantor) per mL of cleared cell lysate (43% AS) was added with fast stirring. Once dissolved, the speed of stirring was decreased and the solution was stirred in the cold room overnight. Several of the mutant proteins did not precipitate overnight and additional ammonium sulfate was needed to induce precipitation (Table 2.1).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Percent ammonium sulfate added</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>43%</td>
</tr>
<tr>
<td>T113R</td>
<td>66%</td>
</tr>
<tr>
<td>Y244A</td>
<td>88%</td>
</tr>
<tr>
<td>I400G</td>
<td>84%</td>
</tr>
</tbody>
</table>

Table 2.1. Percent ammonium sulfate needed to induce precipitation of MutS mutants.

The precipitated solution containing ammonium sulfate was spun down at 17K rpm and 4°C for 30 minutes. The supernatant was discarded and the pellet was resuspended with 2 mL of Buffer B₀ (see Appendix D). The resuspended solution was added to a falcon tube containing an additional 18 mL of B₀. The conductivity of the solution was measured to ensure the concentration of salt was below 100 mM.

A C 10/10 column (GE Healthcare) was rinsed and the tubing was purged with ddH₂O. The column was packed at room temperature with 2 mL of Q-Sepharose Fast Flow resin (GE Healthcare) and the ethanol storage buffer was rinsed by
washing with 10 column volumes (CV) of ddH$_2$O. The column was transferred to the cold room and stored until use. The Q-Sepharose column was equilibrated with 10 CV of buffer B$_{50}$ (see Appendix D) at 1 mL/min in the cold room. The tubing of the fraction collector was purged with ddH$_2$O and attached to the column. Once equilibrated, the resuspended ammonium sulfate pellet was loaded onto the column at 0.7 mL/min and the flow through was collected. The column was washed with 10 CV of B$_{50}$ at 1 mL/min and the wash was collected. The protein was eluted with 10 CV of a B$_{50}$-B$_{400}$ gradient at 0.7 mL/min collected in 1 mL fractions. The peak fractions were determined by mixing 5 µL of the elution fractions with 100 µL of Coomasie Plus (Thermo Fisher Scientific, Inc.) in a 96-well plate. The elution profile was confirmed by running the fractions around the peak on a 10% SDS-PAGE gel alongside the whole cell lysate, cleared cell lysate, ammonium sulfate solution, the column load, the flow through, and the column wash. The protein-containing fractions were pooled.

A 10 kDa MWCO Centrifugal filter (Pall Corporation) was rinsed by centrifugation of 10 mL of B$_{50}$ at 4K rpm for 10 minutes. The pooled protein was concentrated by centrifugation at 4K rpm and 4°C until the volume was reduced to less that 5 mL (30 minutes to 1 hour depending on the start volume). The concentrated sample was dialyzed in a 6-8 kDa MWCO dialysis membrane (Spectrum Laboratories, Inc.) in 500 mL of B$_{50}$ in the cold room overnight. The sample was transferred to a fresh 500 mL of B$_{50}$ and dialyzed for an additional 2 hours in the cold room. The concentration of the protein was determined using the Bradford assay by mixing 1 mL Coomassie Plus (Thermo Fisher Scientific, Inc.) with
1-6 μg of BSA (Thermo Fisher Scientific, Inc.) as a standard and 3 volumes of the protein. The absorbance of the samples was measured at $\lambda = 595$ nm and the concentration of the protein was determined by comparing its absorbance with the standards. The protein solutions were stored in 200 or 400 μL aliquots at -80°C.
**Figure 2.1. WT MutS and the seven mutant proteins were purified.** The molecular weight of the WT MutS monomer is 90,627.3 Da. The concentration of the purified proteins was quantified by the Bradford assay at $\lambda = 595$ nm. 5 µg of WT and each mutant as well as a BSA standard were run on a 10% SDS-PAGE gel. The purity of the preps is comparable and sufficient for preliminary experiments.
Labeling of MutS with IAANS

The labeling of MutS C42A/M88C with 2-[4’-(iodoacetamido)anilino]naphthalene-6-sulfonic acid (IAANS) dye (Invitrogen) was carried out under minimal light conditions. IAANS labeling of WT MutS and all of the mutants was performed as follows:

Conjugation of MutS with IAANS – The purified MutS protein intended for labeling was thawed on ice. The thawed protein was transferred to a 5 mL Eppendorf tube. The protein was reduced by addition of 10-fold molar excess of 20 mM tris(2-carboxyethyl)phosphine (TCEP; Invitrogen; prepared by dilution with ddH₂O) and rotated on a roller drum for 10 minutes. The protein was labeled with IAANS dye by the addition of 25-fold molar excess of 10 mM IAANS (prepared by dilution with dimethylformamide; DMF; J.T. Baker) to the protein in four parts. After each addition, the solution was rotated on the roller drum for 15 minutes at room temperature. After the final addition, the protein sample was moved to a roller drum in the cold room and the conjugation reaction was allowed to continue overnight. The IAANS conjugation reaction was stopped by the addition of 1 M DTT to a final concentration of 5 mM. After the addition, the sample was allowed to continue to rotate in the cold room for 10 minutes.

Removal of free dye with centrifugal filtration – A 15 mL 10 kDa MWCO Centrifugal filter (Pall Corporation) was rinsed by centrifugation of 4 mL of Buffer L (see Appendix D) at 4K rpm for 10 minutes. The excess free dye was removed from the sample by concentration at 5K rpm and 4°C for 10 minutes until the volume was
reduced by half. The sample was washed by the addition of 1 mL of Buffer L, followed by centrifugation at 5K rpm for 15 minutes. The wash step was repeated two times. Following the final wash, the sample was spun at 5K rpm until the sample volume was reduced to less than 500 µL, about one hour. The labeled protein retentate was transferred to a microcentrifuge tube. The concentrator sample reservoir was rinsed with 50 µL of Buffer L, which was added to the retentate tube.

_P-6 gel filtration_— 5 g of dry Bio-Gel P-6 media (BioRad) was added to a plastic beaker and hydrated with 60 mL degassed Buffer L for 4 hours to overnight. Half of the supernatant was decanted and the gel solution was transferred to a sidearm flask. The solution was degassed for 10 minutes with occasional swirling. The gel solution was returned to the beaker and 60 mL of Buffer L was added and swirled gently to resuspend. Once the resin had settled, the supernatant was decanted and the process was repeated three times to remove any fines. A 1.5 x 20 cm Econo-Column (BioRad) was rinsed and fit with a Teflon valve to close the column outlet. The column was packed by addition of 6 mL of Buffer L followed by the slow addition of the hydrated gel. Once the resin settled, the column was transferred to the cold room and equilibrated with 60 mL of Buffer L at 0.35 mL/min.

The equilibrated column was attached to a fraction collector and the buffer was drained to the level of the gel bed. The column outlet was closed, the column top was removed, and the concentrated sample was carefully layered onto the gel bed surface. The column outlet was opened and the sample was allowed to drain into the gel bed. Once the level of the sample reached the upper bed surface, the top of the
column was reattached to flow buffer. The sample was washed through the column with Buffer L at 0.35 mL/min and 80 fractions of 250 µL each were immediately collected. The peak fractions were determined by mixing 5 µL of the elution fractions with 100 µL of Coomassie Plus in a 96-well plate. The fractions were also analyzed using a fluorescence microplate reader (SpectaMax M5) by mixing 196 µL of Buffer L with 4 µL of each fraction in an opaque 96-well microtiter plate (Thermo Fisher, Inc.). The samples were excited at λ = 326 nm and the peak labeled fractions were determined by the emission readings at λ = 445 nm. The peak fractions were analyzed on a 10% SDS-PAGE gel along with the unlabeled protein, the labeled protein, the filtrate, and the retentate (column load). The gel was visualized by UV transillumination with a U:Genius Gel Imaging System (Syngene) to assess the labeling reaction (see Figure 2.2). The labeled protein-containing fractions were pooled.

**Quantification of labeling** – The concentration of the protein was determined using the Bradford assay with BSA as a standard at λ = 595 nm. The labeling efficiency was determined by measuring the absorbance of the samples by UV-visualization with a NanoDrop 2000c Spectrophotometer (Thermo Scientific, Inc.) at λ = 280 nm and λ = 326 nm. The degree of labeling was calculated using the expression,

\[
\frac{A_{\text{IAANS}}}{\varepsilon_{\text{IAANS}}} \times \frac{(\varepsilon_{\text{Monomer}})}{A_{280}} = \frac{[\text{IAANS}]}{[\text{MutS}]} \]
where $A_{\text{IAANS}}$ is the absorbance of IAANS at 326 nm; $\varepsilon_{\text{IAANS}}$ is the molar extinction coefficient of IAANS, 27,000 M$^{-1}$cm$^{-1}$; $A_{280}$ is the absorbance of the protein at 280 nm; and $\varepsilon_{\text{Monomer}}$ is the molar extinction coefficient of one monomer of MutS, 52,720 M$^{-1}$cm$^{-1}$. The equation yields the ratio of IAANS molecules per MutS monomer, which is multiplied by 100 to give the percent of MutS monomers labeled with IAANS. The labeling efficiency of all IAANS labeled proteins was between 70-95%.
Figure 2.2. WT MutS C42A/M88C was labeled with IAANS dye. (a) A 10% SDS-PAGE gel was run to visualize the peak fractions and the labeling efficiency (pU: unlabeled protein; pL: labeled protein; R: retentate; Fi: filtrate; 32-41: peak fraction numbers). The molecular weight of the MutS monomer is 90,627.3 Da. The concentration of the labeled protein was quantified by the Bradford assay at $\lambda = 595$ nm. (b) Prior to staining, the gel was visualized by UV transillumination to assess the labeling reaction and the removal of free dye.
**Preparation of DNA substrates**

I would like to acknowledge and thank Juan Liu for purifying the oligonucleotides used for this study. Single-stranded DNAs (Integrated DNA Technologies, Inc.) were purified by denaturation polyacrylamide gel electrophoresis (PAGE) and ethanol precipitation (see Appendix C for strand sequences).

A 1.5 mm 17% denaturing acrylamide urea gel (urea purchased from OmniPur) was prepared and prerun at 450 V for 30 minutes to warm. The 125-250 nmol ssDNA samples were prepared by dilution with 1X TBE buffer (see Appendix D), boiled for 3 minutes, and loaded onto the gel. The gel was run for 4-5 hours at 450 V until the oligonucleotide migrated one-half to three-fourths of the way through the gel. The oligonucleotide bands were visualized briefly by exposure to short-wave UV light and their location was marked. The oligo-containing bands were cut out with a razor blade and chopped into small pieces. The gel slices were added to a 15 mL falcon tube, immersed in 4 mL of DNA dialysis buffer (see Appendix D), and placed on a roller drum at room temperature to dialyze overnight. The overnight dialysis mixture was centrifuged at 13K rpm and 4°C for 5 minutes and the supernatant was collected using gel loading tips. The remaining acrylamide gel particles were removed by passing the supernatant through a 0.2 µm separating filter (Pall Corporation). The oligonucleotides were concentrated using a speed vacuum. The DNA was precipitated in non-low adhesion tubes by adding three times the sample volume of 100% ethanol (Pharmco-Aaper) and one-tenth the sample volume of 3 M sodium acetate (Spectrum Quality Products, Inc.), pH 5.2 and chilling at -80°C for at least 30 minutes. The oligonucleotide concentrate was pelleted by
centrifugation at 13K rpm for 20-30 minutes at 4°C. The pellet was washed with 80% ethanol and allowed to air dry. The dried pellet was resuspended in ddH₂O. The concentration of the oligonucleotide stock was determined by UV-absorption spectroscopy at λ = 260 nm.

Unlabeled +T DNA was annealed by mixing the purified +T 37mer single-stranded DNA with the purified +T 36mer single-stranded DNA (see Appendix C for strand sequences) in a 1:1 ratio in annealing buffer (see Appendix D) for a final concentration of 50 µM. 2-AP +T DNA was annealed by mixing the purified, 2-AP +T labeled 37mer single-stranded DNA with 2-AP +T unlabeled 36mer single-stranded DNA (see Appendix C for strand sequences) in a 1:1 ratio in annealing buffer for a final concentration of 50 µM. The annealing reaction was carried out by heating the samples to 95°C for 3 minutes and allowing them to cool to room temperature overnight. The duplex products were verified by running the annealed and single-stranded samples on a 16% non-denaturing PAGE gel and were confirmed to be > 90% pure.

Fluorescence-based assays to monitor MutS activity

Introduction to fluorescence theory – Fluorescence is a phenomenon that occurs in primarily aromatic molecules, known as fluorophores. Following the absorption of a photon of light, the electrons of fluorophores are excited from the singlet ground state (S₀) to a higher vibrational energy level (S₁ or S₂). Upon absorption of light, the electrons quickly relax from the highest (S₂) to the lowest (S₁) vibrational energy level in the excited state [98]. The excess energy can either be
given off in the form of heat as in thermal dissipation, radiative decay by reemission of a photon as in fluorescence, or as an energy transfer to another molecule as in Förster Resonance Energy Transfer (FRET). The rapid relaxation of electrons from high energy to lower energy excited states causes the emission spectra of fluorescent molecules to be independent of their absorption spectra since they are decaying from a lower energy excited state than that which they were excited to. Typically, the energy of emission is less than that of absorption and therefore fluorescence occurs at longer wavelengths, a phenomenon known as the Stokes’ Shift [98]. Fluorophores can be characterized by their fluorescence lifetime and quantum yield, measures of the average amount of time the molecule spends in the excited state before returning to the ground state and the number of photons emitted relative to the number absorbed, respectively [98].

The fluorescent properties of molecules are largely influenced by their environment due to the intrinsic polarization of molecules in their excited state. This polarization results in the stabilization of the excited state in a polar environment, shifting the emission to longer wavelengths (red shift). In a hydrophobic environment, the excited state is destabilized, shifting the emission to shorter wavelengths (blue shift). In addition to the environmental effects, fluorescence intensity can be decreased by a phenomenon known as quenching. Quenching can occur by dynamic quenching or static quenching. Dynamic, or collisional, quenching is a process by which the excited-state of a fluorophore is deactivated when it comes in contact with another molecule, known as the quencher, and is dependent on a molecular interaction between the pair. The process of static quenching is similar to
that of collisional quenching, however the fluorophore-quencher pair forms a nonfluorescent complex in which the quenching occurs in the ground state prior to excitation [98].

**Mechanism and use of 2-Aminopurine fluorescence** – 2-AP is a fluorescent adenine analog that forms a base pair with thymine with Watson-Crick geometry (Figure 2.3) [99]. When the fluorophore is excited between 310-320 nm, it emits fluorescence with a maximal emission at 370 nm [68]. 2-AP has been utilized extensively as a site-specific reporter of DNA structure because its fluorescent behavior is strongly affected by its environment. Incorporating 2-AP into a DNA duplex quenches its fluorescence due to its stacking with neighboring bases, reducing its quantum yield compared with the free nucleotide [99]. Additionally, proteins and DNA absorb minimally at 315 nm and therefore the signal-to-noise ratio of 2-AP is high. *In vacuo* calculations predict that 2-AP fluorescence is quenched both statically and dynamically based on its environment. When stacked with purines, 2-AP fluorescence is statically quenched due to the mixing of its ground state molecular orbitals with those of the adjacent bases, which results in electron delocalization of the ground state orbitals and a decreased intensity of the transition to the excited states. In contrast, when 2-AP is adjacent to pyrimidine bases it experiences dynamic quenching. When the fluorophore is excited, the electrons initially undergo transition from the ground state (S₀) to the second excited state (S₂). This excitation is followed by a rapid, nonradiative decay from the S₂ to the S₁ excited state, known as a “dark state,” resulting in a loss of fluorescence intensity and decrease in the lifetime of the
S₂ to S₀ emission. Due to its high sensitivity to base stacking interactions, the Hingorani lab has made extensive use of this fluorescent probe to monitor the interaction of Taq MutS Phe39 with a +T insertion, which is known to kink the DNA [42]. When 2-AP is incorporated into a DNA substrate adjacent to the +T insertion, this bending of the DNA at the mismatch site in the presence of MutS disrupts the base-stacking interactions of the 2-AP, relieving its quenching and resulting in an increase in fluorescence intensity by ~4-fold (Figure 2.3) [68].
Figure 2.3. A 2-AP fluorophore selectively reports MutS binding to +T DNA. (a) 2-Aminopurine is a fluorescent adenine analog. (b) The 2-AP fluorophore is positioned adjacent to the +T insertion and therefore, when MutS binds the +T DNA and inserts a phenylalanine residue as a part of specific mismatch recognition, the DNA is kinked. The quenching of 2-AP due to stacking with the duplex is relieved by increased solvent exposure when MutS binds.
Mechanism and use of IAANS fluorescence – A 2-[4’-(iodoacetamido)aniline] naphthalene-6-sulfonic acid (IAANS) fluorophore is used to assess the conformational dynamics of MutS during the initiation of MMR. When the IAANS fluorophore is excited at its peak wavelength of 326 nm, it emits maximally at 445 nm. IAANS is an environmentally sensitive probe—when exposed to solvent, its fluorescence intensity increases and when buried, its quantum yield is reduced, resulting in the quenching of its fluorescence [100]. While the exact mechanism of IAANS fluorescence quenching is not well characterized, the solvent effect may be a result of the dynamic quenching of the excited state of the fluorophore by the solvent [100].

IAANS has a sulfonyl group that can form a disulfide bond with the thiol group of a cysteine residue, allowing its simple conjugation to proteins. Each Taq MutS monomer contains only one cysteine residue, Cys42. This residue is located in the vicinity of the Phe39 residue important for mismatch recognition, and therefore the introduction of a bulky dye at this site could potentially disrupt the DNA binding activity of the protein. The mutation of Cys42 to alanine allows the selective introduction of a cysteine at the desired site. Structural analysis and MD simulations based on the Taq MutS crystal structure in complex with +T DNA revealed that the Met88 site is an ideal location for the conjugation of the IAANS dye. This residue is located in the DNA binding domain I at a distance of ~16 Å from the Phe39 residue, however previous studies showed that the protein labeled at this site does not alter the activity of MutS [48]. Additionally, this residue is non-conserved and it is thus presumed that its mutation will not significantly disrupt the activity of the protein.
The C42A/M88C MutS mutant was confirmed to have no effect on MutS function in MMR [48]. The C42A/M88C MutS protein was labeled with IAANS dye to selectively report the conformational changes of MutS domains I (Figure 2.4).

The location of the fluorophores at domains I makes this an ideal reporter for the conformational changes of MutS upon binding mismatched DNA; domains I are known to close together to form the specific contact with the error. A recent smFRET experiment with donor and acceptor dyes labeled at the same Cys88 position revealed that in Apo MutS, the distance between domains I fluctuates between 30 and 70 Å [55]. Based on MD simulations of the Taq MutS crystal structure as well as smFRET measurement, it is suggested that upon the formation of a complex between MutS and +T DNA, the fluorophores at the Cys88 position on each monomer are stabilized at a separation of ~20 Å [55]. Closure of the MutSIAANS domains I may reduce solvent accessibility of IAANS, resulting in fluorescence quenching [48, 56].
Figure 2.4. Each MutS monomer was labeled with an IAANS molecule at the Met88C position. (a) Structure of 2-[4’-(iodoacetamido)aniline]naphthalene-6-sulfonic acid (IAANS). (b) The crystal structure of Taq MutS with the Met88 residue highlighted in red shows the position of the IAANS labels. This fluorophore reports the closure of MutS domains I in the presence of mismatched DNA as well as the movement of domains I during MutS sliding clamp formation. When the two IAANS molecules are in close proximity, which likely results in reduction of solvent exposure, their fluorescence is quenched.
Mechanism and use of MDCC-PBP fluorescence – A 7-Diethylamino-3-(((2-Maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) fluorophore was used to label phosphate binding protein (PBP) in order to monitor the real-time release of phosphate (Pₐ) in solution. MDCC absorbs maximally at an excitation wavelength of 425 nm and emits at a peak wavelength of 465 nm. The mutation, A197C, was introduced in *E. coli* PBP to allow for the conjugation of the MDCC fluorophore near the Pᵢ binding site. PBP binds Pᵢ with a high affinity (Kᵦ ~ 0.1 µM) and rapidly at a rate of 1.4 × 10⁸ M⁻¹s⁻¹ and induces a rapid conformational change in the active site of PBP (317 s⁻¹) that closes the cleft of the binding site. This closure alters the MDCC fluorophore environment to be more hydrophobic, resulting in the increase in MDCC fluorescence intensity by ~7-fold [101, 102]. This fast rate of Pᵢ binding by MDCC-PBP as well as the excess of PBP used in the assay ensures that the rate of Pᵢ binding reported by MDCC-PBP accurately reflects the rate of phosphate release by MutS, which is on a significantly slower timescale [51, 65, 68]. Additionally, previous studies have shown that the rate of phosphate release and ATP hydrolysis by MutS occur at the same rate, therefore the rate of Pᵢ binding by MDCC-PBP also reports the rate of MutS ATP hydrolysis [51].

Introduction to steady state and pre-steady state kinetics – Steady state kinetic measurements monitor enzymatic activity over multiple turnovers of product formation. The term, “steady state” refers to conditions in which the substrate concentration is in far excess of the enzyme, allowing the relative concentration of the enzyme and enzyme-substrate complex to remain unchanged over multiple enzymatic
turnovers [103]. Steady state measurements of product formed over time are used to determine kinetic parameters as per the following minimal model:

\[
E + S \xrightleftharpoons[k_{-1}]{k_1} ES \to E + P
\]

where E is the enzyme, S is the substrate, and P is the product. The initial rate of P formation, \(v\), increases with S at low S concentration and approaches a maximum velocity, \(v_{\text{max}}\), at high concentrations where product formation is limited only by the rate of the conversion of the ES complex to E+P [103]. This is quantitatively described by the Michaelis-Menten equation,

\[
v = \frac{[E]_0[S]k_{\text{cat}}}{K_M + [S]}
\]

where \([E]_0\) represents the total enzyme concentration, \([S]\) is the substrate concentration at time t, \(k_{\text{cat}}\) is the rate of catalysis, and \(K_M\) is the Michaelis-Menten constant. The turnover number of the enzyme, \(k_{\text{cat}}\), describes the number of reactions an enzyme catalyzes in an amount of time and is therefore dependent on all steps the enzyme undergoes between substrate binding and product release [103]. The turnover rate is defined by the slowest step(s) in the reaction, masking the steps before and after the rate-limiting step(s). Measuring changes in \(v\) as a function of substrate concentration yields a Michaelis-Menten plot, from which the equilibrium constant, \(K_M\), can be determined. The \(K_M\) is a measure of the formation and decay of the ES complex. Although the ES complex decays to both E+S and E+P, under certain conditions where \(k_{-1} \gg k_2\), equilibrium can be assumed and \(K_M\) is used as an approximation for \(K_D\), the apparent dissociation constant that represents the decay of ES to E+S [104].
In contrast to steady state experiments where the focus is on multiple turnovers and measuring the turnover rate, experiments under pre-steady state conditions monitor the reaction in the first turnover, as the ES forms and decays into P. The measurements are performed at a higher concentration of the enzyme to obtain a good signal for the product formed in one turnover against a high concentration of substrate. Under these conditions, if product formation occurs prior to the rate determining step, burst activity will be observed, implying that a step after product formation is the rate-limiting step [103, 104].

*Equilibrium analysis of MutS binding +T DNA*—The affinity of MutS for a +T insertion was monitored under equilibrium conditions, in which MutS was titrated into 2-AP +T DNA and incubated until the reaction reached an equilibrium between MutS, the DNA substrate, and the MutS-DNA complex. The resulting fluorescence intensity change was used to obtain the dissociation constant, $K_D$, of the complex formation. The $K_D$, which represents the ratio of the rate of dissociation ($k_{off}$) to the rate of association ($k_{on}$), is a measure of the affinity of MutS for the +T insertion [103]. Titration of MutS into 2-AP +T DNA (see Appendix C for sequence) was monitored on a FluoroMax-3 spectrofluorometer (Jobin-Yvon Horiba Group) under equilibrium conditions. Samples of 2.5 mL of binding buffer (see Appendix D) alone and with 20 nM 2-AP +T DNA were prepared in 3 mL quartz cuvettes. MutS was titrated into both samples over a concentration range of 0-250 nM. After each addition, the samples were allowed to incubate at 40°C for 30 seconds while stirring and were excited at 315 nm with a 2 nm excitation slit width. Emission scans were
collected from 360-380 nm with an emission slit width of 5 nm. The emission scan recorded intensity data every 2 nm with an integration time of 2 seconds. The fluorescence intensity of the sample of MutS with buffer only was subtracted from the fluorescence intensity of the MutS with 2-AP +T DNA to correct for the background intrinsic fluorescence signal from the tryptophan and tyrosine residues. The experiment was performed three times for WT MutS and once for each mutant and the subtracted fluorescence intensity at 370 nm was plotted versus the concentration of MutS. The change in fluorescence intensity was normalized between 0 and 1 for all experiments.

The resulting binding isotherms of 16 points were fit to hyperbolic equations and the dissociation constants ($K_D$) were extracted as follows. The intensity at any given point can be related to the fraction of bound DNA by

$$ r = r_0 + (r_{\text{max}} - r_0)F $$

where $r$ is the fluorescence intensity at any point along the curve, $r_0$ is the initial fluorescence intensity, $(r_{\text{max}} - r_0)$ is the total intensity change, and $F$ is the fraction of DNA bound by MutS. The complex formation between MutS and the DNA can be represented by

$$ \frac{K_A}{K_D} $$

$$ \text{MutS} + \text{DNA} \rightleftharpoons \text{MutS} \cdot \text{DNA} $$

The fraction of DNA bound by MutS, $F$, can therefore be represented by

$$ F = \frac{[\text{MD}]}{[D]_t} = \frac{[\text{MD}]}{[D]_f + [\text{MD}]} $$
where \([MD]\) is the concentration of the MutS-DNA complex, \([D]\) is the total concentration of DNA, and \([D]_f\) is the concentration of the free DNA. The \(K_D\) is given by,

\[
K_D = \frac{[MD][D]_f}{[MD]}
\]

and the association constant \(K_A\) is

\[
K_A = \frac{[MD]}{[M]_t[D]_f}
\]

where \([M]_f\) is the concentration of free MutS. From these equations, it is evident that

\[
K_A = \frac{1}{K_D}.
\]

The concentration of the complex can be rewritten as

\[
[MD] = K_A[M]_f[D]_f
\]

and the fraction bound, \(F\), can therefore be represented by

\[
F = \frac{K_A[M]_f[D]_f}{[D]_f + K_A[D]_f[L]_f} = \frac{K_A[M]_f}{1 + K_A[M]_f} = \frac{[M]_f}{K_D + [M]_f}
\]

Substituting this equation for \(F\) into the intensity equation above, we get

\[
r = r_0 + \left(r_{\text{max}} - r_0\right) \frac{[M]_f}{K_D + [M]_f}
\]

This equation can be used to fit binding data when \([M]_t \gg [M]_f\). Therefore, \([M]_f\) can be substituted with \([M]_t\) to give the final equation,

\[
r = r_0 + \left(r_{\text{max}} - r_0\right) \frac{[M]_t}{K_D + [M]_t}
\]

where \((r_{\text{max}} - r_0)\) is the total amplitude change and \([M]_t\) is the concentration of MutS at each point. For preliminary analysis of all the MutS mutants tested in this study, the data were fit to the above hyperbolic function assuming \([M]_t \approx [M]_f\). This
equation was entered into KaleidaGraph (Version 4.5; Synergy Software Inc.) to fit the data and the $K_D$ was extracted from the fit.

**Transient kinetic analysis of MutS binding +T DNA** – The fast rate of the complex formation between MutS and the +T insertion was monitored under transient conditions in a KinTek SF-2001 stopped-flow instrument (KinTek Corp.) in the dark at 40°C. Samples were prepared in binding buffer (see Appendix D). The experiment was performed by mixing a 1:1 ratio of 2-AP +T DNA with an equal volume of MutS (final concentrations: 30nM 2-AP +T DNA; 0.2, 0.4, and 0.6 µM MutS). A xenon lamp was used to excite the 2-AP fluorophore at 315 nm and the emission at >350 nm was measured over time with a photomultiplier tube (PMT) set to 650 V and a long-pass cut-off filter (LG-350-F; Corion). The excitation slit width was set to 20 nm. The change in fluorescence signal was measured using a split time window collecting 400 points between 0-0.5 seconds and 600 points between 0.5-5 seconds. At least four traces were collected and at least two traces were averaged at each condition. The change in fluorescence intensity was normalized to the initial value and fit to a single exponential function using KaleidaGraph to determine the observed binding rate, $k_{obs}$, as described below.

The reaction of MutS with DNA is reversible, however under these conditions with a fast timescale and the relatively slow rate of MutS dissociation from +T DNA in the absence of nucleotides, the reaction is essentially irreversible and can be written as

\[
\frac{k_{on}}{k_{off}} \quad \text{MutS} + \text{DNA} \rightarrow \text{MutS} \cdot \text{DNA}
\]
Additionally, although the reaction is technically second order, where two molecules collide in order to form the product, the high concentration of MutS compared to the DNA implies that the reaction is under pseudo-first order conditions, meaning we can assume that the concentration of MutS is constant [103]. Therefore,

\[
\frac{d[\text{DNA}]}{dt} = -k_{\text{on}}[\text{DNA}]_t[\text{MutS}]_t
\]

where \([\text{DNA}]_t\) is the concentration of DNA at any time \(t\), \([\text{DNA}]_t\) is the total concentration of DNA and \([\text{MutS}]_t\) is the total concentration of MutS. The observed binding rate, \(k_{\text{obs}}\), is dependent on the excess component, in this case MutS, so

\[
k_{\text{obs}} = k_{\text{on}}[\text{MutS}]_t
\]

From this, \(k_{\text{on}}\), can be rewritten and substituted into the previous equation to give

\[
\frac{d[\text{DNA}]}{dt} = -k_{\text{obs}}[\text{DNA}]_t
\]

which can be rearranged to give

\[
\frac{d[\text{DNA}]}{[\text{DNA}]_t} = -k_{\text{obs}} dt
\]

Solving this equation by integration gives

\[
\ln[\text{DNA}]_t = -k_{\text{obs}} t
\]

and taking the exponential of this function gives

\[
\frac{[\text{DNA}]}{[\text{DNA}]_t} = e^{-k_{\text{obs}} t}
\]

Since \([\text{DNA}]_t = [\text{DNA}] + [\text{MutS} \cdot \text{DNA}]\), the above equation can be rewritten as

\[
\frac{[\text{MutS} \cdot \text{DNA}]}{[\text{DNA}]_t} = 1 - e^{-k_{\text{obs}} t}
\]
This equation serves as a measure of the fraction of DNA bound, F. In this assay, we are directly monitoring the formation of the MutS-DNA complex with the change in the 2-AP fluorescence intensity. Using the following equation,

\[ r_t = r_{\text{max}} + (r_0 - r_{\text{max}})F \]

where \( r_t \) is the fluorescence intensity at time \( t \), \( r_{\text{max}} \) is the maximum intensity, and \( r_0 - r_{\text{max}} \) is the amplitude change, we can replace \( F \) with the equation for fraction bound to get the final equation:

\[ r_t = r_{\text{max}} + r(1 - e^{-k_{\text{obs}}t}) \]

where \( r \) is the signal amplitude change [103, 104]. This equation was entered into KaleidaGraph and the \( k_{\text{obs}} \) was extracted from the fit. As described above, this rate is dependent on the concentration of the excess component, MutS, and therefore the \( k_{\text{obs}} \) at the three concentrations of MutS measured, 0.2, 0.4, and 0.6 \( \mu \)M, was plotted versus [MutS] and the bimolecular rate constant, \( k_{\text{on}} \), was determined by the slope of the linear fit.

**Kinetic analysis of MutS conformational changes with +T DNA** – The conformational changes of MutS binding mismatched DNA were monitored by the transient binding of MutS\textsubscript{IAANS} to +T DNA using a KinTek SF-2001 stopped-flow instrument in the dark at 40°C. Samples were prepared in binding buffer (see Appendix D). The experiment was performed by mixing a 1:1 ratio of MutS\textsubscript{IAANS} with an equal volume of +T DNA (final concentrations: 100 nM MutS\textsubscript{IAANS}; 1, 3, and 6 \( \mu \)M +T DNA). A xenon lamp was used to excite the IAANS fluorophore at 326 nm and the emission at >350 nm was measured over time with a photomultiplier tube.
(PMT) set to 600 V and a long-pass cut-off filter (LG-350-F; Corion). The excitation slit width was set to 10 nm. The change in fluorescence signal was measured using a split time window collecting 400 points between 0-0.5 seconds and 600 points between 0.5-5 seconds. At least four traces were collected and at least two traces were averaged at each condition. The change in fluorescence intensity was normalized to the initial value and fit to a double exponential function using KaleidaGraph to determine the observed rates of conformational changes, $k_i$ and $k_ii$, as described below.

The MutS IAANS mixing with +T DNA is under similar conditions as MutS binding to 2-AP +T DNA described above, however in this case the DNA is in excess over MutS. Using the same simple reaction scheme,

$$k_{on} \quad MutS + DNA \rightarrow MutS \cdot DNA$$

but instead making the assumption that $[DNA]_f = [DNA]_i$, we can say $k_{obs} = k_{on}[DNA]_t$. Following the same derivation as above gives,

$$\frac{[MutS \cdot DNA]}{[MutS]_t} = 1 - e^{-k_{obs}t}$$

the fraction of MutS bound, $F$ [103]. As before, this equation can be substituted in for the $F$ in the intensity equation to give,

$$r_t = r_{max} + r(1 - e^{-k_{obs}t})$$

However, examination of the $\chi^2$ value of the residuals of the single-exponential fit indicated a poor correlation. The data was therefore fit to a double-exponential function,

$$r_t = r_{max} + r(1 - e^{-k_i t}) + r(1 - e^{-k_{ii} t})$$
where $k_i$ represents the rate of the first phase and $k_{ii}$ represents the rate of the second phase. The $\chi^2$ value of the residuals of the double-exponential fit indicate a higher correlation. This equation was entered into KaleidaGraph and the two $k_{\text{obs}}$ values, $k_i$ and $k_{ii}$ values were extracted from the fit. The first rate, $k_i$, which is dependent on the concentration of the substrate, +T DNA, was therefore measured at three concentrations of +T DNA, 1, 3, and 6 µM, and plotted versus [+T DNA]. At low [+T DNA], $k_i$ shows linear dependence on DNA concentration [48], thus the bimolecular rate constant, $k_{\text{on}}$, was estimated at 1 µM +T DNA by

$$k_{\text{on}} = \frac{k_{\text{obs}}}{[\text{DNA}]_t}$$

At higher [+T DNA], $k_i$ approaches saturation and data were fit to a hyperbola using the equation

$$k_i = \frac{k_{\text{max}}[\text{DNA}]}{K_{1/2} + [\text{DNA}]}$$

and the $K_{1/2}$ and $k_{\text{max}}$ were extracted from the fit [48]. In contrast, $k_{ii}$ remains independent of [+T DNA].

**Kinetic analysis of the nucleotide effect on conformational changes of MutS** —

The conformational changes of MutS$_{IAANS}$ with ATP were monitored in the presence and absence of +T DNA using a KinTek SF-2001 stopped-flow instrument in the dark at 40°C. Samples were prepared in binding buffer (see Appendix D). The experiment in the absence of DNA was performed by mixing a 1:1 ratio of MutS$_{IAANS}$ with an equal volume of ATP (final concentrations: 100 nM MutS$_{IAANS}$, 100 µM ATP). The experiment in the presence of DNA was performed by mixing a 1:1 ratio of
MutSIAANS pre-incubated with +T DNA with an equal volume of ATP (final concentrations: 100 nM MutSIAANS, 1 µM +T DNA, 100 µM ATP). A xenon lamp was used to excite the IAANS fluorophore at 326 nm and the emission at >350 nm was measured over time with a photomultiplier tube (PMT) set to 600 V and a long-pass cut-off filter (LG-350-F; Corion). The excitation slit width was set to 10 nm. The change in fluorescence signal was measured by collecting 1,000 data points from 0 to 0.5 seconds. At least five traces were collected and at least four traces were averaged at each condition.

In the absence of DNA, the traces were biphasic and were fit to a double-exponential function in KaleidaGraph as described by

$$r_t = r_{\text{max}} + r(1 - e^{-k_{\text{up}}t}) + r(1 - e^{-k_{\text{down}}t})$$

where $k_{\text{up}}$ represents the rate of first phase, an increase in signal intensity, and $k_{\text{down}}$ represents the rate of the second phase, a decrease in signal intensity [104].

Previous studies of the MutSIAANS kinetics with ATP suggested the model,

$$\text{MutS + ATP} \rightleftharpoons \text{MutS} \cdot \text{ATP} \rightarrow \text{MutS} \cdot \text{ADP}$$

where $k_{\text{on}}$ represents the rate constant for ATP binding and is reported by $k_{\text{up}}$, $k_{\text{off}}$ represents the rate of ATP dissociation, and $k_{\text{hydrol}}$ represents the rate of ATP hydrolysis and is reported by $k_{\text{down}}$ [48]. The rate of ATP binding, $k_{\text{on}}$, is linearly dependent on the concentration of ATP and therefore,

$$k_{\text{on}} = \frac{k_{\text{up}}}{[\text{ATP}]}$$

where [ATP] is the concentration of ATP in the reaction [48].
In the presence of DNA, the traces were biphasic and fit to a double-exponential function in KaleidaGraph as described by

\[ n_t = r_{\text{max}} + r(1 - e^{-k_1 t}) + r(1 - e^{-k_{\text{ii}} t}) \]

where \( k_1 \) represents the rate of the first phase and \( k_{\text{ii}} \) represents the rate of the second phase. Previous analysis found that the kinetics of the MutS_{IAANS} +T DNA complex mixing with ATP follows the model,

\[
\begin{align*}
\text{MutS} \cdot \text{DNA} + \text{ATP} & \rightleftharpoons \text{MutS} \cdot \text{DNA} \cdot \text{ATP} \\
& \rightarrow \text{MutS}^* \cdot \text{DNA} \cdot \text{ATP} \\
& \rightarrow \text{MutS}^{**} \cdot \text{DNA} \cdot \text{ATP}
\end{align*}
\]

where \( k_1 \) and \( k_{\text{ii}} \) are concentration independent and represent the rates of conformational changes of MutS [48].

**Kinetic analysis of phosphate release by MutS** – I would like to acknowledge and thank Juan Liu for carrying out phosphate release experiments. As a follow-up to the MutS_{IAANS} assay in the presence of ATP, the rate of ATP hydrolysis by MutS WT, Y244A, and I400G was explicitly measured by monitoring the rate of phosphate (P_i) release in the absence and presence of +T DNA using a KinTek SF-2001 stopped-flow instrument in the dark at 40°C. Samples were prepared in P_i release buffer (see Appendix D). In the absence of DNA, phosphate release was monitored by mixing MutS with MDCC-PBP pre-incubated with ATP in a 1:1 ratio (final concentrations: 0.5 µM MutS, 500 µM ATP, 8 µM MDCC-PBP). In the presence of DNA, phosphate release was monitored by mixing MutS pre-incubated with +T DNA with MDCC-PBP pre-incubated with ATP in a 1:1 ratio (final concentrations: 0.5 µM MutS, 1 µM +T DNA, 500 µM ATP, 8 µM MDCC-PBP). The stopped-flow was “mopped” with a
coupled enzyme reaction containing 200 μM 7-methylguanosine (7-MEG) and 0.01 U/mL purine nucleoside phosphorylase (PNPase) for one hour prior to starting the experiment to sequester any phosphate (P\(_i\)) contamination. The P\(_i\) release assay was monitored by MDCC excitation at \(\lambda = 425\) nm and the emission at >450 nm was measured over time with a photomultiplier tube (PMT) set to 460 V and a long-pass cut-off filter (LG-450-F; Corion). The excitation slit width was set to 600 nm. A P\(_i\) calibration standard curve of four points (final concentrations of P\(_i\): 0, 1, 2, and 3 μM) was obtained prior to the experiment and used to convert the fluorescence intensity to [P\(_i\)]. At least four traces were collected and at least two traces averaged at each condition. The traces showing a burst were fit to a burst plus a linear phase equation as defined by,

\[
[P_i] = \frac{r_0(1 - e^{-k_{burst}t}) + k_{ss}t}{m_{P_i}}
\]

where \(r_0\) is the burst amplitude, \(k_{burst}\) is the rate of the burst phase, \(k_{ss}\) is the steady state rate of the linear phase (in units M s\(^{-1}\)), and \(m_{P_i}\) is the slope of the P\(_i\) standard curve [104]. The traces that showed only a linear phase were fit to a linear equation defined by \([P_i] = k_{ss}t\). The \(k_{cat}\) was obtained by dividing the steady state rate, \(k_{ss}\), by the concentration of the number of ATPase active sites; in this case, there are two ATPase sites per MutS dimer so

\[
k_{cat} = \frac{k_{ss}}{2[\text{MutS}]}
\]

where \([\text{MutS}]\) is the concentration of the MutS dimer in the reaction (in units M) and \(k_{cat}\) is the rate of the catalytic turnover (in units s\(^{-1}\)).
III. RESULTS AND DISCUSSION

This study served as a preliminary investigation into the activities of several Lynch syndrome-linked MutS mutants to identify any detrimental effects on their function. Using several fluorescent reporters, the kinetic properties of MutS-DNA complex formation, ATPase activity, and corresponding conformational dynamics of the MutS variants were analyzed. The assays were designed to monitor specific steps in the reaction and the results were used to construct a mechanism of action for each MutS mutant protein. The study provided preliminary evidence suggesting that several of the cancer-linked mutations alter MutS activities, which could potentially disrupt DNA repair, and thereby offering insights into the mechanistic underpinnings of LS.

Equilibrium analysis of MutS mutants binding +T DNA

In order to determine the affinity of the mutant proteins for error-containing DNA, the binding was monitored under equilibrium conditions. Previous studies have shown that MutS binds mismatched DNA with a relatively high affinity reflected by the dissociation constant, $K_D \sim 15-25$ nM [48, 68]. This tight binding is essential for the ability of MutS to specifically recognize a mismatch; therefore, differences in the $K_D$ values of the mutant proteins could indicate a compromised ability to recognize errors and ultimately to initiate MMR.

Binding was monitored with the 2-aminopurine (2-AP) fluorescent reporter placed adjacent to a +T insertion in the DNA substrate. As the DNA is bound by MutS, the fluorescence intensity of 2-AP increases (see Mechanism and use of 2-
Aminopurine fluorescence in Materials and Methods). Titration of MutS (0-250 nM) into the 2-AP +T DNA (20 nM), a 36mer DNA substrate with a central +T insertion, results in an increase in 2-AP fluorescence (Figure 3.1). With the addition of MutS to the solution, an increasing fraction of the DNA is bound, and the signal change saturates as an equilibrium is reached. The plot of fluorescence intensity change versus MutS concentration was fit with a hyperbolic function to obtain an estimate of the $K_D$ (see Equilibrium analysis of MutS binding +T DNA in Materials and Methods). The data for the WT binding curve are an average of three experiments, while the data for the MutS mutants are from one experiment. The binding isotherms for WT (Figure 3.1) and each mutant protein (Figures 3.2-3.3) are shown below and the $K_D$ values are listed in Table 3.1.
Figure 3.1. Binding of WT MutS to 2-AP +T DNA. MutS binding to a +T insertion was monitored by titration of 20 nM 2-AP +T DNA with 0-250 nM MutS under equilibrium conditions. The WT experiment was performed three times and the data points represent averaged data, while the error bars represent the standard error of the mean. The data were fit to a hyperbolic function, which yielded a dissociation constant, $K_D = 19.7 \pm 4.9$ nM.
Figure 3.2. MutS mutants T113R, G222D, F243S, Y244A, and L533R appear to bind 2-AP +T DNA with WT-like affinity. Binding of the MutS mutants to +T DNA was monitored by the titration of 20 nM 2-AP +T DNA with 0-250 nM MutS. One experiment was performed for each mutant protein and the data were fit to a hyperbolic equation to estimate the $K_D$. (a) T113R binds the +T insertion with $K_D = 19.9 \pm 3.1$ nM, (b) G222D binds with $K_D = 34.7 \pm 6.5$ nM, (c) F243S binds with $K_D = 14.6 \pm 2.6$ nM, (d) Y244A binds with $K_D = 32.5 \pm 5.5$ nM, and (e) L533R binds with $K_D = 32.9 \pm 6.7$ nM.
Figure 3.3. MutS mutants I400G and G434R bind 2-AP +T DNA with higher affinity than WT. Binding of the MutS mutants to +T DNA was monitored by the titration of 20 nM 2-AP +T DNA with 0-250 nM MutS. One experiment was performed for each mutant protein and the data were fit to a hyperbolic equation to estimate the $K_D$. (a) I400G binds the +T insertion with $K_D = 7.6 \pm 1.8$ nM and (b) G434R binds with $K_D = 5.8 \pm 1.0$ nM, 2-3-fold tighter than WT.
<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>19.7 ± 4.9</td>
</tr>
<tr>
<td>T113R</td>
<td>20.5 ± 4.4</td>
</tr>
<tr>
<td>G222D</td>
<td>34.7 ± 6.5</td>
</tr>
<tr>
<td>F243S</td>
<td>14.6 ± 2.6</td>
</tr>
<tr>
<td>Y244A</td>
<td>32.5 ± 5.5</td>
</tr>
<tr>
<td>I400G</td>
<td>7.6 ± 1.8</td>
</tr>
<tr>
<td>G434R</td>
<td>5.8 ± 1.0</td>
</tr>
<tr>
<td>L533R</td>
<td>32.9 ± 6.7</td>
</tr>
</tbody>
</table>

Table 3.1. The affinity of MutS mutants for +T insertion DNA was determined under equilibrium conditions. Binding isotherms for WT MutS and mutants with +T DNA were fit a hyperbolic equation to estimate $K_D$ (see *Equilibrium analysis of MutS binding to +T DNA* in Materials and Methods).
All of the MutS mutants bind the +T insertion with a low apparent dissociation constant of 5-33 nM, indicating a high affinity interaction. WT MutS binds the +T DNA tightly with $K_D = 19.7$ nM. This is within the range of previously reported dissociation constants of MutS for +T DNA of $K_D \sim 15-25$ nM [48, 68]. The mutants T113R and F243S bind the +T insertion with a similar affinity to WT MutS, while the mutants G222D, Y244A, and L533R exhibit slightly weaker affinity (Table 3.1). The mutants I400G and G434R exhibit ~2-3-fold tighter binding at $K_D = 7.6$ nM and 5.8 nM, respectively (Figure 3.3). Thus, overall the mutant proteins appear to bind +T DNA with high affinity like WT, and it remains to be determined if tighter binding by I400G and G434R mutants affects related MutS activities in MMR.

**Transient kinetic analysis of MutS mutants binding +T DNA**

In order to gain more insight into the initial association event between MutS and DNA, binding of MutS to 2-AP +T DNA was monitored under transient conditions in the absence of nucleotides. Consistent with the equilibrium binding experiment, the 2-AP +T DNA fluorescence intensity increases rapidly on interaction with MutS. A previous study monitoring fluorescence of 2-[4'-(iodoacetamido) aniline]naphthalene-6-sulfonic acid (IAANS) dye-labeled MutS titrated with unlabeled +T DNA, or fluorescence anisotropy of 5-Carboxytetramethylrhodamine (TAMRA) dye-labeled +T DNA titrated with unlabeled MutS, found a linear increase in the observed binding rate to yield a bimolecular association rate, $k_{on} = 3 \times 10^6$ M$^{-1}$ s$^{-1}$, that saturates around 30 s$^{-1}$ for the interaction of WT MutS and +T DNA [48, 68]. Note that due to background signal from MutS tryptophan fluorescence, the 2-AP
DNA cannot be titrated to high enough concentration of MutS to determine the saturating rate; this step is monitored in the subsequent experiments with IAANS-labeled MutS (MutS\textsubscript{IAANS}) binding +T DNA. The bimolecular association rate, $k_{\text{on}}$, was estimated from a linear fit of the observed binding rate, $k_{\text{obs}}$, at three concentrations of MutS (see Kinetic analysis of MutS binding to +T DNA in Materials and Methods). Alterations to the on-rates of the mutant proteins may indicate differences in the mechanism of complex formation between the protein and a mismatch.

DNA binding kinetics were monitored by mixing equal volumes of MutS with 2-AP +T DNA in a stopped-flow instrument, as depicted in the mixing scheme in Figure 3.4a (final concentrations: 200, 400, and 600 nM MutS; 30 nM 2-AP +T DNA). For each experiment, the increase in fluorescence signal was fit to a single-exponential function and the $k_{\text{obs}}$ was extracted from the fit (see Kinetic analysis of MutS binding to +T DNA in Materials and Methods). A representative trace for WT (Figure 3.4b) and each mutant (Figures 3-3.6) at 200 nM MutS is shown below. The $k_{\text{obs}}$ rates were plotted versus the concentration of MutS and the data were fit to a line. The apparent bimolecular rate constant, $k_{\text{on}}$, was determined for each protein by the slope of the linear fit. The concentration dependence of $k_{\text{obs}}$ for WT is shown in Figure 3.4c and the rate dependence of each mutant protein is compared with WT in Figures 3.5e and 3.6d. The observed binding rates ($k_{\text{obs}}$) at 200, 400, and 600 nM of MutS and the apparent bimolecular rate constants ($k_{\text{on}}$) for WT and each mutant protein are given in Table 3.2.
Figure 3.4. Binding kinetics of WT MutS to 2-AP +T DNA. (a) The binding reaction was monitored by mixing equal volumes of MutS with 2-AP +T DNA (final concentrations: 200, 400, and 600 nM MutS; 30 nM 2-AP +T DNA). (b) A stopped-flow trace of 200 nM WT MutS mixed with 30 nM 2-AP +T DNA fit to a single-exponential function gave an apparent binding rate, \( k_{\text{obs}} = 1.4 \text{ s}^{-1} \). (c) The MutS concentration dependence of \( k_{\text{obs}} \) was plotted as a function of [MutS] to yield a linear increase with a \( k_{\text{on}} = 5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \).
Figure 3.5. Binding kinetics of MutS Y244A, I400G, G434R, and F243S to 2-AP +T DNA show WT-like rates of association. The binding reaction was monitored by mixing equal volumes of the MutS mutants with 2-AP +T DNA (final concentrations:
200, 400, and 600 nM MutS; 30 nM 2-AP +T DNA; see Figure 3.4a for mixing scheme). A stopped-flow trace of 200 nM MutS mixed with 30 nM 2-AP +T DNA fit to a single-exponential function yielded an apparent binding rate, (a) \( k_{\text{obs}} = 1.6 \text{ s}^{-1} \) for Y244A, (b) \( k_{\text{obs}} = 2.2 \text{ s}^{-1} \) for I400G, (c) \( k_{\text{obs}} = 1.5 \text{ s}^{-1} \) for G434R (d) and \( k_{\text{obs}} = 0.8 \text{ s}^{-1} \) for F243S. (e) The MutS concentration dependence of \( k_{\text{obs}} \) of Y244A, I400G, G434R, and F243S was plotted as a function of [MutS] to yield a linear increase with a \( k_{\text{on}} = 5.8 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \) for Y244A, \( 6.3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \) for I400G, \( 6.3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \) for G434R, and \( 3.3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \) for F243S. The apparent bimolecular rate constants for these mutants are comparable to that of WT MutS with \( k_{\text{on}} = 5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \), suggesting that they bind the +T insertion in the same manner as WT.
Figure 3.6. Binding kinetics of MutS T113R, G222D, and L533R to 2-AP +T DNA show faster rates of association than WT. The binding reaction was monitored by mixing equal volumes of the MutS mutants with 2-AP +T DNA (final concentrations: 200, 400, and 600 nM MutS; 30 nM 2-AP +T DNA; see Figure 3.4a for mixing scheme). (a) A stopped-flow trace of 200 nM MutS mixed with 30 nM 2-AP +T DNA fit to a single-exponential function yielded an apparent binding rate (a) \( k_{\text{obs}} = 1.2 \text{ s}^{-1} \) for T113R, (b) \( k_{\text{obs}} = 1.2 \text{ s}^{-1} \) for G222D, and (c) \( k_{\text{obs}} = 0.9 \text{ s}^{-1} \) for L533R. (d) The MutS concentration dependence of \( k_{\text{obs}} \) of T113R, G222D, and L533R was plotted as a function of [MutS] to yield a linear increase with a \( k_{\text{on}} = 13.8 \times 10^6 \text{ M}^{-1} \text{s}^{-1} \).
for T113R and G222D, and $k_{\text{off}} = 10.7 \times 10^6 \text{M}^{-1} \text{s}^{-1}$ for L533R. The apparent bimolecular rate constants for these mutants are 2-3-fold faster than that of WT MutS at $k_{\text{off}} = 5 \times 10^6 \text{M}^{-1} \text{s}^{-1}$, suggesting that their interaction with the mismatch may differ slightly compared to WT MutS.
Table 3.2. The MutS concentration dependence of $k_{\text{obs}}$ yields a bimolecular association rate for $+T$ DNA. The observed binding rates, $k_{\text{obs}}$, were monitored by mixing equal volumes of MutS with 2-AP $+T$ DNA for final concentrations of 200, 400, and 600 nM MutS and 30 nM 2-AP $+T$ DNA. The MutS concentration dependence of $k_{\text{obs}}$ was fit to a line to yield the apparent $k_{\text{on}}$ values for the WT and mutant proteins.
The binding rate constant for WT MutS at $5 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ measured in this study is in agreement with the previously reported association rate of $3 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ (Figure 3.4) [68]. This fast rate of complex formation indicates that MutS binds the +T insertion rapidly in the absence of nucleotides. The binding rate constants for Y244A, I400G, G434R, and F243S mutants indicate that they bind +T DNA at a similar rate to WT MutS (Table 3.2). Interestingly, the I400G and G434R mutants were found to have a tighter $K_D$ than WT in the equilibrium binding experiments, and the finding that they have similar $k_{on}$ rates suggests that the difference in affinity is a result of slower $k_{off}$ rates. This hypothesis needs to be followed up by explicit measurement of the off-rates with a dissociation experiment. In contrast, the mutants T113R, G222D, and L533R appear to bind the +T insertion slightly faster than WT (Table 3.2). From the equilibrium binding experiment, these mutants all were found to have a similar affinity for the +T DNA to WT, therefore this faster on-rate suggests faster off-rates for the MutS-mismatch recognition complex. Again, to specifically parse any alteration in the binding mechanism, the off-rates need to be measured as well.

**Kinetic analysis of MutS mutant conformational changes in the presence of +T DNA**

The transient conformational changes in MutS during complex formation with the mismatch were monitored with 2-[4’-(iodoacetamido)aniline]naphthalene-6-sulfonic acid (IAANS) dye-labeled MutS at the M88C residue of domains I. As MutS binds the +T insertion, the fluorescence intensity of the dye decreases (see *Mechanism and use of IAANS fluorescence* in Materials and Methods). The rate of
complex formation and the rate at which the binding saturates was measured with MutSIAANS binding to +T DNA in the absence of nucleotides. Upon the interaction with +T DNA, the IAANS fluorescence intensity rapidly decreases, which has been correlated with the insertion of the Phe residue into the minor groove and the kinking of the DNA at the mismatch [48]. Previous studies have shown biphasic kinetics for the binding of MutSIAANS to +T DNA; the first phase was found to be hyperbolically dependent on the concentration of +T DNA, reaching a maximum of 31 s\(^{-1}\) with a \(K_D\) of 2.3 \(\mu\)M [48]. The second phase was found to increase slightly with the concentration of +T DNA at a rate of 0.2 \(\mu\)M\(^{-1}\) s\(^{-1}\), however experiments with labeled DNA substrates indicate that this phase is likely due to a heterogeneity in the population of MutSIAANS, rather than a distinct step in the mechanism [48]. The hyperbolic dependence of the fast rate on the concentration of DNA indicates that the binding occurs in two steps—a fast association between MutS and the DNA is followed by a concentration-independent second step that involves a rate-limiting isomerization in MutS and the DNA [48]. The \(K_D\) of 2.3 \(\mu\)M obtained from the concentration dependence of the first rate indicates that the initial collision complex that forms between MutS and the +T DNA is \(~100\) -fold weaker than the specific recognition complex with a \(K_D\) of \(~20\) nM obtained from the equilibrium binding experiment (see Figure 3.1). This data suggests that the first step involves the formation of an initial weak complex followed by a second step entailing rate-limiting concerted conformational changes in MutS and the DNA to yield the high-affinity mismatch-specific complex [48]. The MutS and DNA isomerization event is thought
to be an important element for signaling the switch from MutS mismatch search to mismatch recognition [48].

The binding kinetics of WT and the mutant proteins were monitored by mixing equal volumes of MutS<sub>IAANS</sub> and +T DNA in a stopped-flow instrument as depicted in the mixing scheme in Figure 3.8a (final concentrations: 200 nM MutS<sub>IAANS</sub> and 1, 3, and 6 µM +T DNA). For each experiment, the decrease in the MutS<sub>IAANS</sub> fluorescence signal was fit to a double-exponential function and the observed rates of the two phases were extracted from the fit and are referred to as $k_i$ and $k_{ii}$ (see Kinetic analysis of MutS conformational changes with +T DNA in Materials and Methods). A representative trace for WT (Figure 3.8) and each mutant (Figures 3.9-3.11) at 1 µM +T DNA is shown below. The on-rate, $k_{on}$, was estimated for each mutant from the $k_i$ value at 1 µM +T DNA (see Kinetic analysis of MutS conformational changes with +T DNA in Materials and Methods). Due to the large number of mutants being studied and the complexity of the experiment, the initial titration was performed at only three different DNA concentrations; thus, the data must be considered to be in a preliminary stage. Based on the previous study showing that the $k_i$ rate of WT MutS increases hyperbolically with +T DNA and saturates around 6 µM +T DNA, the concentration dependence of $k_i$ for WT and each mutant was fit to a hyperbola in this study for preliminary analysis, despite the paucity of data points [48]. From the hyperbolic fit, the apparent $K_D$ and maximum $k_i$ ($k_{i(max)}$) were estimated. The rate dependence of $k_i$ and $k_{ii}$ for WT is shown in Figure 3.8 and for each mutant protein is shown in Figures 3.9-3.11. Also, the observed binding rates ($k_i$ and $k_{ii}$) at 1, 3, and 6 µM of +T DNA, the estimated bimolecular rate constants
(\(k_{on}\)), the apparent \(K_D\), and the apparent \(k_{i(max)}\) for WT and each mutant protein are listed in Table 3.3.
**Figure 3.7. Binding kinetics of WT MutS_{IAANS} to +T DNA.** (a) The binding reaction was monitored by mixing equal volumes of MutS_{IAANS} with +T DNA (final concentrations: 0.2 µM MutS_{IAANS}; 1, 3, and 6 µM +T DNA). (b) A stopped-flow trace of 0.2 µM WT MutS_{IAANS} binding to 1 µM +T DNA fit to a double-exponential function yielded apparent binding rates of $k_i = 3.4 \text{ s}^{-1}$ and $k_{ii} = 0.7 \text{ s}^{-1}$. From the rate of the first phase ($k_i$) at 1µM +T DNA, the bimolecular rate constant was estimated as $k_{on} = 3.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$. (c) The rate dependence on the concentration of +T DNA of $k_i$ was plotted as a function of the concentration of +T DNA and shows a hyperbolic increase with +T DNA to a maximum of 21 s\(^{-1}\) with an apparent $K_D$ of 2.5 µM; the second phase ($k_{ii}$) is concentration independent at 0.7-0.8 s\(^{-1}\).
Figure 3.8. Binding kinetics of MutS_{IAANS} Y244A, I400G, L533R, F243S to +T DNA show WT-like rates of complex formation. The binding reaction was monitored by mixing equal volumes of MutS_{IAANS} with +T DNA (final concentrations: 0.2 µM MutS_{IAANS}; 1, 3, and 6 µM +T DNA). (a) A stopped-flow trace of 0.2 µM MutS_{IAANS} Y244A binding to 1 µM of +T DNA fit to a double-exponential function yielded apparent binding rates of $k_i = 8.8$ s$^{-1}$ and $k_{ii} = 0.4$ s$^{-1}$. From the rate of the first phase ($k_i$) at 1 µM +T DNA, the bimolecular rate constant was estimated as $k_{on} = 8.8 \times 10^6$ M$^{-1}$ s$^{-1}$. (b) The rate dependence on the concentration of +T DNA of $k_i$ of Y244A was plotted as a function of the concentration of +T DNA and shows a hyperbolic increase with +T DNA to a maximum of 20 s$^{-1}$ with an apparent $K_D$ of 1.1 µM; the second phase ($k_{ii}$) is concentration independent at 0.4-0.7 s$^{-1}$. (c) Similar analysis of I400G yielded $k_i = 5.1$ s$^{-1}$, $k_{ii} = 0.5$ s$^{-1}$, and an estimated $k_{on} = 5.1 \times 10^6$ M$^{-1}$ s$^{-1}$ with (d) $k_{i(max)} = 14$ s$^{-1}$ and an apparent $K_D = 1.5$ µM; (e) L533R yielded $k_i = 5$ s$^{-1}$, $k_{ii} = 0.4$ s$^{-1}$, and an estimated $k_{on} = 5 \times 10^6$ M$^{-1}$ s$^{-1}$ with (f) $k_{i(max)} = 30$ s$^{-1}$ and an apparent $K_D = 3$ µM; and (g) F243S yielded $k_i = 18.5$ s$^{-1}$, $k_{ii} = 0.9$ s$^{-1}$, and an estimated $k_{on} = 18.5 \times 10^6$ M$^{-1}$ s$^{-1}$ with (h) $k_{i(max)} = 29$ s$^{-1}$ and an apparent $K_D = 1$ µM.
Figure 3.9. Binding kinetics of MutS\textsubscript{IAANS} T113R and G434R to +T DNA indicate a tighter initial collision complex. The binding reaction was monitored by mixing equal volumes of MutS\textsubscript{IAANS} with +T DNA (final concentrations: 0.2 µM MutS\textsubscript{IAANS}; 1, 3, and 6 µM +T DNA). (a) A stopped-flow trace of 0.2 µM MutS T113R binding to 1 µM +T DNA fit to a double-exponential function yielded apparent binding rates of $k_i = 8.2$ s$^{-1}$ and $k_{ii} = 0.8$ s$^{-1}$. From the rate of the first phase ($k_i$) at 1 µM +T DNA, the bimolecular rate constant was estimated as $k_{on} = 8.2 \times 10^6$ M$^{-1}$ s$^{-1}$, approximately 3-fold faster than that of WT MutS. (b) The rate dependence on the concentration of +T DNA of $k_i$ of T113R was plotted as a function of the concentration of +T DNA...
and shows a hyperbolic increase with +T DNA to a maximum of 12 s⁻¹ with an apparent $K_D$ of 0.5 µM, indicating an ~5-fold tighter initial association; the second phase ($k_{ii}$) is concentration independent at 0.7-0.8 s⁻¹. (c) Similar analysis of G434R yielded $k_i = 14.8$ s⁻¹ and $k_{ii} = 0.8$ s⁻¹ and an estimated $k_{on} = 14.8 \times 10^6$ M⁻¹ s⁻¹, which is ~5-fold faster than WT MutS, (d) and a $k_{i(\text{max})} = 18$ s⁻¹ with an apparent $K_D$ of 0.2 µM, indicating an ~10-fold tighter initial association.
Figure 3.10. Binding kinetics of MutS\textsubscript{IAANS} G222D to +T DNA. The binding reaction was monitored by mixing equal volumes of MutS\textsubscript{IAANS} with +T DNA (final concentrations: 0.2 µM MutS\textsubscript{IAANS}; 1, 3, and 6 µM +T DNA). (a) A stopped-flow trace of 0.2 µM MutS G222D binding to 1 µM +T DNA fit to a double-exponential function yielded apparent binding rates of $k_i = 8.5 \text{ s}^{-1}$ and $k_{ii} = 0.5 \text{ s}^{-1}$. From the rate of the first phase ($k_i$) at 1 µM +T DNA, the bimolecular rate constant was estimated as $k_{on} = 8.5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$, ~4-fold faster than that of WT MutS. (b) The rate dependence on the concentration of +T DNA of $k_i$ was plotted as a function of the concentration of +T DNA and shows a hyperbolic increase with +T DNA to a maximum of 43 s\(^{-1}\) with an apparent $K_D$ of 6.3 µM; the second phase ($k_{ii}$) is concentration independent at 0.4-0.7 s\(^{-1}\). The faster maximum rate and higher $K_D$ compared to WT MutS (see Figure 3.7) indicate that G222D may bind the +T DNA faster but form a weaker initial association complex than WT.
Table 3.3. The +T DNA concentration dependence of $k_i$ yields an apparent $K_D$

and $k_i^{(\text{max})}$ for the binding of MutS\textsubscript{IAANS} to +T DNA. The observed rates, $k_i$ and $k_{ii}$, were monitored by mixing equal volumes of MutS\textsubscript{IAANS} with +T DNA for final concentrations of 200 nM MutS and 1, 3, and 6 µM +T DNA. The observed rate $k_i$ at 1 µM +T DNA was used to estimate the $k_{on}$ of the interaction between MutS and +T DNA. The concentration dependence of $k_i$ on +T DNA was fit to a hyperbola to yield the apparent $K_D$ and $k_i^{(\text{max})}$ values for the WT and mutant proteins.
The binding kinetics of WT MutS\textsubscript{IAANS} to +T DNA yields a $k_{on}$ rate of $3.4 \times 10^6$ M\textsuperscript{-1} s\textsuperscript{-1} that is consistent with the on-rate found with the 2-AP experiment, as well as with previous findings (Table 3.3) [68]. The observed binding rates at three concentrations of +T DNA were fit to a hyperbola, to give an estimated $K_D$ of 2.5 µM with a rate of saturation of 21 s\textsuperscript{-1}, which is in agreement with previous findings that the initial collision complex between MutS and +T DNA has a $K_D$ of 2.3 µM and reaches a maximum of 31 s\textsuperscript{-1} [48]. These results confirm that WT MutS binds the +T insertion rapidly and weakly prior to undergoing the rate-limited concentration-independent isomerization in MutS and DNA that results in a tighter mismatch-specific complex with nanomolar affinity. The second rate, $k_{ii}$, appears to range from 0.4-0.9 s\textsuperscript{-1} for WT and all mutants at all concentrations, which is consistent with the hypothesis that this rate may be an artifact of heterogeneity in the population of IAANS-labeled MutS, rather than an actual step in the reaction [48].

The binding kinetics of the MutS\textsubscript{IAANS} mutants to +T DNA were monitored under the same conditions. MutS Y244A, I400G, F243S, and L533R appear to undergo similar interactions with +T DNA and form similar weak initial collision complexes with +T DNA to WT (Table 3.3). As discussed previously, I400G appeared to have an ~2-fold higher affinity for +T DNA in the equilibrium experiment, which could indicate that the subsequent tighter complex has different dissociation kinetics than WT. F243S has an ~5-fold faster apparent $k_{on}$ rate than WT MutS (Table 3.3). This faster on-rate could indicate that the mutant binds +T DNA faster than WT, however from the 2-AP binding data, it appeared to have a similar
on-rate to WT and therefore the discrepancy could be due to error in initial measurements and needs to be resolved with additional data.

The mutants T113R and G434R show an increased affinity for +T DNA. T113R binds the +T DNA ~2-fold faster than WT MutS (Table 3.3). This faster binding rate is consistent with the apparent $k_{on}$ determined for T113R with the 2-AP binding experiment which similarly appeared to be ~3-fold faster than WT (Table 3.2). From the concentration dependence of the apparent binding rate on +T DNA, the collision complex appears to have a $K_D$ of 0.5 µM and reaches a maximum rate of 12 s$^{-1}$ (Figure 3.9). These results indicate that this mutant has an ~5-fold higher affinity for initial complex with the +T insertion than WT and it converts to the tight, specific complex at an ~2-fold slower rate than WT. These data suggest that this mutant has a higher non-specific binding affinity and may bind matched DNA tighter than WT; that is, it has potentially lost some of its mismatch specificity. This hypothesis needs to be further explored, for example by monitoring the interaction of the protein with matched DNA.

Likewise, G434R appears to bind +T DNA ~5-fold faster than WT and the initial collision complex is ~10-fold tighter with a $K_D$ of 0.2 µM (Table 3.3). Although in the transient 2-AP binding experiment this mutant was found to bind +T DNA at the same rate as WT, it had an ~3-fold higher affinity for +T DNA (Tables 3.2 and 3.1). Taken together, these results indicate that the G434R mutant forms a higher affinity initial collision complex as well as a tighter MutS-mismatch bound complex than WT. The G434R mutation introduces a positively charged Arg residue in place of a Gly in the vicinity of the DNA backbone. This higher affinity is in line
with the hypothesis that the additional positively charged residue could provide a stabilizing contact. A follow-up dissociation experiment is needed to determine if there is any difference in the \( k_{\text{off}} \) rate compared with WT to determine the lifetime of the complex with +T DNA.

The mutant G222D also has an ~2-fold faster binding rate than WT (Table 3.3). Interestingly, unlike WT, the binding rate of G222D does not appear to saturate by 6 µM +T DNA. From a rough fit of the rate dependence, G222D reaches a \( k_{i(\text{max})} \) of 43 s\(^{-1}\), approximately 2-fold faster than WT and has an apparent \( K_D \) of 6.3 µM, which is ~3-fold weaker than WT. The difference in concentration dependence of the observed binding rates for this mutant may indicate that although the protein appears to bind the +T insertion faster than WT, the initial association forms a weaker complex than WT. This mutant was found to have a slightly weaker \( K_D \) at equilibrium (Table 3.1), which could indicate that it also has a faster off-rate than WT, a hypothesis that needs to be followed up in the future with dissociation experiments.

**Monitoring the kinetics of MutS mutant conformational changes with nucleotides in the absence of DNA**

The 2-AP and MutS\textsubscript{IAANS} binding experiments confirm that in the presence of a +T insertion, an initial collision complex is formed, followed by a conformational change to form a high affinity complex between MutS and the +T DNA. In addition to these conformational changes in the presence of mismatched DNA, the binding of nucleotides to domains V leads to rearrangement of domains I both in the presence and absence of DNA (see Figure 1.3 for the domain labeling scheme) [48, 55].
recent smFRET study found that in the absence of DNA, ATP binding to domains V triggers the closure of domains I and then rapid ATP hydrolysis returns MutS to an ADP-bound state with domains I open [55]. The dynamic movements of MutS domains I with ATP are thought to allow MutS to probe duplex DNA for errors and help convert it into a sliding clamp after mismatch recognition [55]. Recently, a study showed that mixing MutSIAANS with ATP in a stopped-flow results in an initial increase, followed by a decrease in MutSIAANS fluorescence [48]. A complementary experiment was performed in the presence of a non-hydrolyzable ATP analog, ATPγS, and revealed that only the initial increase in fluorescence intensity occurred. Moreover, in the presence of ADP, only a slight decrease in intensity was detected [48]. Notably, the rate of signal increase was found to be ATP concentration dependent at $0.8 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ and the rate of the decrease was hyperbolically dependent on ATP and limited to $11 \text{ s}^{-1}$, which correspond to previously measured rates of ATP binding and hydrolysis, respectively [48, 51]. These results as well as the smFRET study, indicate that the conformational change in MutS resulting in an increase in signal of MutSIAANS corresponds to the rate of ATP binding and the movement of MutS domains I proximal to each other and the subsequent drop in signal is representative of the rate of ATP hydrolysis at the high affinity site and the return of MutS domains I to a position distal to each other [48, 55]. This study further revealed that the domains I isomerization occurs only after ATP binds the weak ATPase site, so both the weak and the tight ATPase sites must be ATP bound to trigger the conformational change [48]. From these results, it was suggested that the existence of a slow step after ATP binding but prior to ATP hydrolysis may involve
MutS sampling intermediate states, possibly the sliding clamp state, prior to ATP hydrolysis. In the absence of a mismatch, the intermediate sliding clamp conformation is not stabilized and ATP is rapidly hydrolyzed to return MutS to the more stable ADP bound or nucleotide free state [48]. The catalytic cycle is able to repeat only after ADP dissociates from both ATPase sites [48].

In order to monitor the effect of ATP on the conformational dynamics of the MutS mutants during the mismatch search phase, the MutS interaction with ATP was investigated under transient conditions by mixing equal volumes of MutS\textsubscript{IAANS} and ATP (final concentrations: 100 nM MutS\textsubscript{IAANS} and 100 µM ATP). For WT and each mutant protein, the kinetics show two phases and the traces were fit to double exponential functions (see *Kinetic analysis of the nucleotide effect on conformational changes of MutS* in Materials and Methods). Based on the previous studies described above, the first phase, $k_{\text{up}}$, was interpreted as the conformational change in MutS corresponding with ATP binding and the second phase, $k_{\text{down}}$, was interpreted as the subsequent domains I rearrangement with ATP hydrolysis at the high affinity site [48]. The stopped-flow traces WT and each mutant protein in the presence of ATP are shown in Figures 3.12-3.14 and the rates of $k_{\text{up}}$, $k_{\text{down}}$, and the estimated bimolecular rate constant of ATP binding are listed in Table 3.4.
**Figure 3.11. Kinetics of WT MutS$_{IAANS}$ conformational changes driven by ATP binding and hydrolysis.** (a) The nucleotide effect on free MutS was monitored by mixing equal volumes of the protein with ATP (final concentrations: 0.1 µM of MutS$_{IAANS}$, 100 µM of ATP). (b) The stopped-flow trace of MutS$_{IAANS}$ mixed with ATP shows a biphasic increase and decrease in signal with ATP. The rate of the first phase, $k_{up}$, is 70 s$^{-1}$ and the rate of the second phase, $k_{down}$, is 6 s$^{-1}$. The concentration-dependent rate of ATP binding is estimated from $k_{up}$ to be $0.7 \times 10^6$ M$^{-1}$ s$^{-1}$. 
Figure 3.12. Kinetics of conformational changes of MutS mutants driven by ATP binding and hydrolysis. The nucleotide effect on free MutS was monitored by mixing equal volumes of the protein with ATP (final concentrations: 0.1 µM of MutSIAANS, 100 µM of ATP). (a) The stopped-flow traces of the mutants T113R, F243S, and G434R MutSIAANS mixed with ATP (100 µM) show a biphasic increase followed by a decrease in fluorescence intensity with ATP. The rate of the first phase, \( k_{up} \), of T113R and G434R is 63 s\(^{-1} \) and the rate of the first phase of F243S, is 71 s\(^{-1} \). These rates are consistent with that of WT MutS, indicating that these mutants undergo the conformational change in response to binding ATP at the same rate as WT. The rate of the second phase, \( k_{down} \), is 5 s\(^{-1} \) for all three mutants. This again is close to the rate for WT MutS and indicates that these mutant proteins also undergo the conformational change with ATP hydrolysis at the same rate as WT (see Figure 3.11). (b) The stopped flow traces of MutSIAANS G222D, I400G, and L533R mixed with ATP similarly have a biphasic increase followed by a decrease in MutSIAANS intensity. For G222D \( k_{up} = 57 \) s\(^{-1} \) and \( k_{down} = 5 \) s\(^{-1} \), for I400G \( k_{up} = 59 \) s\(^{-1} \) and \( k_{down} = 6 \) s\(^{-1} \).
s$^{-1}$, and for L533R $k_{\text{up}} = 61$ s$^{-1}$ and $k_{\text{down}} = 6$ s$^{-1}$. Again, these rates are comparable to those of WT and indicate that the mutants bind and hydrolyze ATP at a similar rate to WT. However, these mutants reach only half of the amplitude of the WT trace, which may indicate that a population of the mutants are not undergoing the conformational changes or the proteins are not undergoing the same extent of the conformational changes as the WT protein.
Figure 3.13. Kinetics of MutS Y244A conformational changes reveal differential activity in the presence of ATP. The nucleotide effect on free MutS Y244A was monitored by mixing equal volumes of the protein with ATP (final concentrations: 0.1 µM MutSIAANS, 100 µM of ATP). The stopped-flow trace of the MutSIAANS Y244A mixed with ATP shows a biphasic increase followed by a decrease in fluorescence intensity with ATP. The rate of the first phase, $k_{\text{up}}$, is consistent with that of WT MutS at 68 s$^{-1}$, however the second phase, $k_{\text{down}}$, is different. The rate of $k_{\text{down}}$ of Y244A is 2.4 s$^{-1}$, which is 2-3-fold slower than the WT rate. This slower decrease in intensity may represent a slower rate of ATP hydrolysis at the high affinity site or could indicate an uncoupling between ATP hydrolysis at domains V and conformational changes in domains I.
Table 3.4. Kinetic parameters of MutS conformational changes in the presence of ATP. For WT and each mutant protein, a fast increase in fluorescence intensity was followed by a slower decrease. The rate \( k_{up} \) is correlated with the rate of ATP binding and the rate \( k_{down} \) is concerted with the rate of ATP hydrolysis at the high affinity site. All of the mutants appear to bind ATP with a rate close to that of WT and undergo conformational changes to varying extents in response to ATP hydrolysis. The mutant, Y244A, appears to have a 2-3-fold slower \( k_{down} \) rate than WT.
Monitoring the conformational changes in MutS domains I in the presence of ATP revealed biphasic kinetics indicating that WT MutS undergoes two conformational changes. The first phase resulted in an increase in fluorescence intensity \((k_{\text{up}} = 70 \text{ s}^{-1})\) followed by a decrease in intensity \((k_{\text{down}} = 6 \text{ s}^{-1})\). This is consistent with previous data indicating that the first phase corresponds with the conformational change in domains I when ATP binds domains V and the second phase represents the domains I rearrangement with a burst of ATP hydrolysis at the high affinity site [48]. Based on the previous finding that the rate of ATP binding is linearly dependent on the concentration of ATP, the bimolecular rate of ATP binding was estimated to be \(0.7 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}\). This is consistent with the rates previously found with the IAANS-labeled MutS protein as well as from \(^{32}\text{P}\)-radiolabeled ATP binding [48, 51]. Thus, these results confirm that the first observed conformational change in domains I correlates with ATP binding to domains V. The rate of the second phase, \(6 \text{ s}^{-1}\), is in line with the ATP hydrolysis rate at \(~10 \text{ s}^{-1}\) (see phosphate release kinetics in Figure 3.17 below), confirming that the second conformational change in domains I correlates with ATP hydrolysis at domains V [48, 51].

The experiment was repeated for each mutant protein under the same conditions and revealed that all mutants undergo these conformational changes coupled to ATP binding and hydrolysis. T113R, F243S, and G434R appear to have WT-like activity with a fast conformational change at \(~70 \text{ s}^{-1}\), followed by slower conformational change at \(5 \text{ s}^{-1}\) (Figure 3.13, Table 3.4). Mutants G222D, I400G, and L533R also have similar domains I dynamics in response to ATP, however the amplitudes of the traces of these mutants appear to be about half that of WT. Because
this is a bulk experiment, the exact cause of the reduction in amplitude cannot be deduced. The lower amplitude could either imply that a population of these proteins are not undergoing the conformational change in response to ATP or that the variant is undergoing the conformational change to a lesser extent than the WT protein. One plausible explanation for the difference in amplitude could be that, based on previous findings, when MutS is pre-incubated with low concentrations of ADP, the ATP-induced increase in MutSIAANS fluorescence is blocked [48]. If any of the mutant protein preparations contain contaminating ADP, it could contribute to the reduction in observed amplitude.

For Y244A, the conformational dynamics in response to ATP appear to be altered compared to WT. Although Y244A has an apparent $k_{up}$ rate of 68 s$^{-1}$, indicating that it binds ATP at the same rate as WT, the rate of the second phase, $k_{down}$, is only 2.4 s$^{-1}$, ~3-fold slower than WT (Figure 3.14). This lower rate may indicate that this mutant hydrolyzes ATP slower than WT or it could hydrolyze ATP at the same rate as WT but undergo the conformational change slower, which would indicate an uncoupling between the ATPase activity and the corresponding protein dynamics. These hypotheses were followed up with explicit measurement of ATP hydrolysis and phosphate release by the mutant protein as described below (Figure 3.17). Additionally, a higher fraction of this mutant remains in the $k_{up}$ phase, indicating accumulation of the ATP-bound species that dissipates at the slower $k_{down}$ rate.
Monitoring the kinetics of MutS mutant conformational changes with nucleotides in the presence of +T DNA

Prior studies have revealed that in the presence of mismatched DNA, MutS ATP hydrolysis is suppressed and the protein stalls in an ATP-bound conformation in which it is able to diffuse away from the site of the mismatch [55, 58, 72, 78]. Previous study of MutSIAANS after binding a +T insertion with ATP revealed three distinct conformational changes during the formation of the sliding clamp. First a lag phase was observable at low concentrations of ATP, representing binding to both ATPase sites [48]. The following two phases were found to be concentration-independent increases in fluorescence intensity of MutSIAANS. The first rate, \( k_1 \), was 3.5 s\(^{-1}\), and the second, \( k_{ii} \), was 0.27 s\(^{-1}\) [48]. The MutSIAANS experiment alone was not sufficient to assign these rates to specific steps of the MutS mechanism, however other smFRET-based studies and experiments with multiple labels on the +T DNA revealed that the first conformational change is likely ATP-induced domains I rearrangement that allows for DNA unbending and the second, slower phase is the further movement of domains I to form the sliding clamp and fully release the mismatch [48, 55]. The second phase is consistent with the previously reported steady-state rate of ATP hydrolysis in the presence of +T DNA (\( k_{cat} = 0.4 \) s\(^{-1}\)) [51]. MutL was found to interact with MutS after the first ATP-induced conformational change during which the DNA unbends but prior to sliding clamp formation [74]. Therefore, the slow MutS conformational changes that stall MutS in the ATP-bound state are thought to be important for localizing MutL to the mismatch [74].
In order to investigate the effect of ATP on the sliding clamp formation of the MutS mutants, MutS was pre-incubated with +T DNA and the interaction of the complex with ATP was monitored on a stopped-flow instrument (final concentrations: 100 nM MutSIAANS, 1 µM +T DNA, 100 µM ATP). For WT and each mutant protein, the lag phase is undetectable at this high concentration of ATP and the kinetics show two phases, which were fit to double exponential functions (see *Kinetic analysis of the nucleotide effect on conformational changes of MutS* in Materials and Methods). The stopped-flow traces of the MutS-DNA complex mixed with ATP for WT and the mutants are given in Figures 3.15-3.17 and the rates $k_{i}$ and $k_{ii}$ are listed in Table 3.5. Alterations to the kinetics of the conformational changes of MutS domains I after mismatch recognition in the presence of ATP could indicate a loss of communication between the DNA binding and ATPase sites and possible loss of ability to form the sliding clamp and signal to downstream MutL proteins.
Figure 3.14. Kinetics of the conformational changes of WT MutS bound to +T DNA in the presence of ATP. (a) The nucleotide effect on the MutS-mismatch complex was monitored by mixing equal volumes of MutSIAANS pre-incubated with +T DNA with ATP (final concentrations: 0.1 µM MutSIAANS, 1 µM +T DNA, 100 µM ATP). (b) The stopped-flow trace of the MutSIAANS +T DNA complex mixed with ATP shows biphasic increase in fluorescence intensity. The rate of the first phase, $k_i$, is 4 s$^{-1}$ and the rate of the second phase, $k_{ii}$, is 0.27 s$^{-1}$. The conformational changes of MutS with ATP in the presence of +T DNA are significantly altered from those in the absence of DNA (Figure 3.12); MutS remains in a stable ATP-bound state and domains I move apart, rather than rapidly returning to a conformation with domains I proximal to each other as in the absence of DNA.
Figure 3.15. Kinetics of the conformational changes of MutS mutants bound to +T DNA in the presence of ATP. The nucleotide effect on MutS mutants complexed with +T DNA was monitored by mixing equal volumes of MutS and +T DNA with ATP (final concentrations: 0.1 µM MutS\textsubscript{IAANS}, 1 µM +T DNA, 100 µM ATP). (a) The stopped-flow traces of the MutS\textsubscript{IAANS} T113R, F243S, and G434R complexed with +T DNA and mixed with ATP are compared with the WT trace and all show biphasic increases in fluorescence intensity. T113R, F243S, and G434R have a $k_i$ of 5.7 s\(^{-1}\), 4.5 s\(^{-1}\), and 3.1 s\(^{-1}\), respectively, a $k_{ii}$ of 0.3 s\(^{-1}\), 0.29 s\(^{-1}\) and 0.26 s\(^{-1}\), respectively. The rates of these conformational changes are comparable to WT, indicating that these mutants undergo the conformational changes to unbend the DNA and form the MutS sliding clamp. (b) The stopped-flow traces of the MutS\textsubscript{IAANS} G222D and L533R complexed with +T DNA and mixed with ATP show similar rates of conformational changes with a $k_i$ of 5.8 s\(^{-1}\) and a $k_{ii}$ of 0.4 s\(^{-1}\) for G222D and a $k_i$ of 6 s\(^{-1}\) and a $k_{ii}$ of 0.27 s\(^{-1}\) for L533R, however, the amplitudes of the two phases are altered. These lower amplitudes may indicate that a fraction of each mutant protein is
not forming the sliding clamp or, alternately, the mutants may not undergo the same extent or type of conformational change as WT.
Figure 3.16. Kinetics of the conformational changes of Y244A and I400G bound to +T DNA with ATP show altered activity. The nucleotide effect on Y244A and I400G complexed with +T DNA was monitored by mixing equal volumes of MutS and +T DNA with ATP (final concentrations: 0.1 µM MutSIAANS, 1 µM +T DNA, 100 µM ATP). (a) The stopped-flow trace of the Y244A MutSIAANS +T DNA complex mixed with ATP is compared with the WT trace and shows a biphasic increase followed by a decrease in fluorescence intensity. The first rate, $k_i$, occurs at 3.7 s$^{-1}$ with an amplitude close to double that of the WT protein, while the second rate, $k_{ii}$, is a decrease in intensity at 2.2 s$^{-1}$. (b) The stopped-flow trace of the I400G MutSIAANS +T DNA complex mixed with ATP is compared with the WT trace and shows a biphasic increase followed by a decrease in fluorescence intensity. The increase in fluorescence intensity of the first phase occurs at 13 s$^{-1}$ and the subsequent decrease at 7 s$^{-1}$ indicates a difference in the conformational dynamics from WT.
<table>
<thead>
<tr>
<th>Protein</th>
<th>$k_1 \text{ (s}^{-1}\text{)}$</th>
<th>$k_2 \text{ (s}^{-1}\text{)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>4</td>
<td>0.27</td>
</tr>
<tr>
<td>T113R</td>
<td>5.7</td>
<td>0.3</td>
</tr>
<tr>
<td>G222D</td>
<td>5.8</td>
<td>0.37</td>
</tr>
<tr>
<td>F243S</td>
<td>4.5</td>
<td>0.29</td>
</tr>
<tr>
<td>Y244A</td>
<td>3.7</td>
<td>2.2*</td>
</tr>
<tr>
<td>I400G</td>
<td>13</td>
<td>7*</td>
</tr>
<tr>
<td>G434R</td>
<td>3.1</td>
<td>0.26</td>
</tr>
<tr>
<td>L533R</td>
<td>6</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*Corresponds to a decrease in intensity

Table 3.5. Conformational changes in MutS complexed with a +T insertion were monitored in the presence of ATP. For WT, T113R, G222D, F243S, G434R, and L533R, a biphasic increase in fluorescence intensity was observed. The conformational changes correlate with the MutS domains I dissociating from the DNA to release the mismatch and rearranging to form the ATP-bound sliding clamp [48]. Y244A and I400G show an increase followed by a decrease in fluorescence, indicating a difference in conformational dynamics, possibly the loss of sliding clamp formation.
After WT MutS forms a complex with the +T insertion, its conformational dynamics with ATP are altered such that it undergoes two slow conformational changes in domains I at 4 s\(^{-1}\) and 0.27 s\(^{-1}\). These rates are consistent with previous data showing that the initial conformational change is correlated with DNA unbending and the second is a further rearrangement to the MutS sliding clamp [48]. The ability of the mutant proteins to undergo these conformational changes was similarly monitored and compared with WT. T113R, F243S, and G434R show conformational changes comparable to WT with biphasic increases in fluorescence intensity (Figure 3.15a). The G222D and L533R mutants also undergo similar conformational changes to WT, however the traces appear to have slightly lower amplitudes (Figure 3.15b). These amplitude differences could represent a population of the proteins that are not undergoing the conformational change or that the mutants are not undergoing the same extent of conformational change as WT. Although this difference is subtle, it may hinder the ability to signal to the downstream MutL protein, which relies on the positioning of the MutS domains [70, 74]. This hypothesis needs to be followed up with future experiments of the ability of the MutS mutants to signal MutL activity.

The conformational dynamics of Y244A show significant differences. The first phase occurs at 3.7 s\(^{-1}\), which is close to WT at 4 s\(^{-1}\), however the amplitude is nearly 200% of that of WT (Figure 3.16a). Additionally, instead of undergoing the slow isomerization to the sliding clamp, the mutant shows a fluorescence intensity decreases at a rate of 2.2 s\(^{-1}\) in the second phase (Figure 3.16a). Interestingly, the rate of the second phase is consistent with the rate of the conformational change of
Y244A in the absence of DNA that corresponds with ATP hydrolysis, 2.4 s\(^{-1}\) (Table 3.4), potentially indicating that the mutant continues to hydrolyze ATP after binding +T DNA. This different response to ATP may indicate a decoupling of the DNA binding and the ATPase activities. Likewise, I400G also appears to have altered activity with +T DNA and ATP. The initial increase in fluorescence intensity occurs ~4-fold faster than WT (Figure 3.16b), potentially indicating that ATP binding triggers the release of I400G from +T DNA. The second conformational change results in a decrease in fluorescence intensity at 7 s\(^{-1}\) (Figure 3.16b), the same rate as found with ATP hydrolysis in the absence of DNA (Table 3.4). One plausible explanation for the differential activity of these mutants is that during the first conformational change in which the domains I dissociate from the DNA, the protein may release the DNA entirely. Then, in the absence of the mismatch-induced stabilization of the ATP-bound conformation, the proteins may hydrolyze ATP to quickly reset to an ADP-bound state that can rebind DNA. This indicates a compromised ability to successfully form the sliding clamp conformation necessary for the recruitment of downstream repair proteins. This hypothesis was followed up with measurement of ATP hydrolysis and P\(_i\) release by Y244A and I400G in the presence of +T DNA as described below (Figure 3.20).

**Kinetic analysis of ATP hydrolysis and phosphate release by MutS mutants**

As a follow-up to the results showing that Y244A and I400G may undergo altered conformational dynamics with ATP, the rate of ATP hydrolysis was measured with a phosphate release assay. This experiment makes use of a fluorescent reporter
protein, 7-Diethylamino-3-(((2-Maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) labeled phosphate binding protein (PBP). In this assay, MDCC-PBP binds free P$_i$ rapidly and with a high affinity, which induces a conformational change in the reporter that results in an increase in the fluorescence intensity of the MDCC dye by ~7-fold (see Mechanism and use of MDCC-PBP fluorescence in Materials and Methods) [101, 102]. Previous study of P$_i$ release by MutS found that in the absence of DNA, MutS rapidly hydrolyzes ATP at the high-affinity site, resulting in a burst of P$_i$ release at 10 s$^{-1}$, indicating that the rate-limiting step is a step after P$_i$ release [51]. This is followed by a slower steady state turnover rate ($k_{cat} = 0.3$ s$^{-1}$), which may be limited by the rate of ATP hydrolysis at the low-affinity site as well as ADP release [51, 63]. The P$_i$ release of the mutant proteins was monitored by mixing equal volumes of MutS with ATP and MDCC-PBP in a stopped-flow (final concentrations: 0.5 µM MutS, 500 µM ATP, 8 µM MDCC-PBP). The stopped-flow traces were fit to a burst plus a linear phase equation (see Kinetic analysis of phosphate release by MutS in Materials and Methods). Using a P$_i$ standard curve, the intensity of the signal was correlated to the concentration of P$_i$ and the concentration of the MutS dimer was used to plot P$_i$ released per MutS dimer. The traces for WT and each mutant protein and are shown in Figures 3.17 and 3.18, respectively.
Figure 3.17. Kinetics of WT MutS phosphate release. (a) The P_\text{i} release by WT MutS was monitored by mixing equal volumes of protein with ATP and MDCC-PBP (final concentrations: 0.5 \mu M MutS, 500 \mu M ATP, 8 \mu M MDCC-PBP). (b) The stopped-flow trace of MutS mixed with ATP and MDCC-PBP shows a burst of P_\text{i} release at 8 s^{-1} with an amplitude of 0.9 P_\text{i} released per MutS dimer, followed by a catalytic turnover rate $k_{\text{cat}}$ of 0.1 s^{-1}, indicating that ATP is hydrolyzed rapidly at one ATPase site followed by a steady state turnover, likely limited by ATP hydrolysis at the low-affinity site and/or ADP release.
Figure 3.18. Kinetics of phosphate release by MutS Y244A and I400G. The Pᵢ release by MutS Y244A and I400G was monitored by mixing equal volumes of protein with ATP and MDCC-PBP (final concentrations: 0.5 µM MutS, 500 µM ATP, 8 µM MDCC-PBP). Y244A shows a slower burst rate of 5 s⁻¹ with a lower burst amplitude of 0.6 Pᵢ per MutS dimer than WT with the same catalytic turnover rate at 0.2 s⁻¹. I400G shows the same burst rate ($k_{burst} = 7$ s⁻¹) and amplitude (1.0 Pᵢ per MutS dimer) as WT as well as the same catalytic turnover rate ($k_{cat} = 0.2$ s⁻¹).
Consistent with previous findings, the P$_i$ release of WT MutS in the absence of DNA shows a fast burst of ATP hydrolysis at one ATPase site at a rate of 8 s$^{-1}$, followed by a steady state turnover rate of 0.2 s$^{-1}$. The experiment with the mutant proteins reveals that the ATPase cycle of I400G without +T DNA is identical to that of WT (Figure 3.18). This is consistent with the data showing that the mutant undergoes the conformational changes in response to ATP binding and hydrolysis at the same rates as WT in the absence of DNA (Figure 3.12). Y244A, however, appears to undergo the burst of ATP hydrolysis ~2-fold slower than WT and reaches only half the amplitude at 0.6 P$_i$ released per MutS dimer (Figure 3.18). Since the protein cannot be releasing half a P$_i$, this sub-stoichiometric activity indicates a heterogeneity in the Y244A sample in which a fraction of the protein is inactive. The 2-fold reduction in the burst rate is consistent with the 2-3-fold decrease in the rate of the domains I conformational change with ATP hydrolysis (Figure 3.13).

**Kinetic analysis of phosphate release by MutS mutants in the presence of +T DNA**

In order to further analyze the differential conformational dynamics of the MutS+-T DNA complexes in response to ATP, the rate of phosphate release by Y244A and I400G was monitored in the presence of +T DNA. Previous studies found that the +T insertion suppresses the burst of ATP hydrolysis by MutS to give only a steady state rate ($k_{cat} = 0.4$ s$^{-1}$), indicating that the rate-limiting step becomes a step prior to ATP hydrolysis [51]. As described previously, after encountering a mismatch, ATP-bound MutS forms a sliding clamp that is essential for enabling interaction with MutL [58, 73, 74, 78]. The results of MutS$_{IAANS}$ and +T DNA mixed with ATP for
both Y244A and I400G indicate a loss of the formation of the sliding clamp. From this experiment, it was predicted that the mutants may hydrolyze ATP rapidly, even in the presence of a +T insertion, resulting in their dissociation from the DNA. To test this hypothesis, the ATP hydrolysis activity of Y244A and I400G in the presence of +T DNA was measured using MDCC-PBP as described above (see Kinetics of phosphate release by MutS in Materials and Methods). The Pi release was monitored by mixing equal volumes of MutS and +T DNA with ATP and MDCC-PBP in a stopped-flow (final concentrations: 0.5 µM MutS, 1 µM +T DNA, 500 µM ATP, 8 µM MDCC-PBP). As in the absence of DNA, a Pi standard curve was used to plot the signal intensity as a function of the concentration of Pi as well as Pi released per MutS dimer. The traces for WT and each mutant protein and are shown in Figures 3.19 and 3.20, respectively and all of the Pi release data is summarized in Table 3.6.
Figure 3.19. Kinetics of phosphate release by WT MutS bound to +T DNA. (a)
The P_i release by WT MutS bound to +T DNA was monitored by mixing equal volumes of protein and DNA with ATP and MDCC-PBP (final concentrations: 0.5 μM MutS, 1 μM +T DNA, 500 μM ATP, 8 μM MDCC-PBP). (b) The stopped flow trace of WT MutS in the presence of +T DNA shows a suppression of the burst phase and reveals only a steady state turnover rate, $k_{cat} = 0.2 \text{s}^{-1}$, indicating that the rate-limiting step is a step prior to P_i release.
Figure 3.20. Kinetics of phosphate release by MutS Y244A and I400G bound to +T DNA. The P_i release by MutS Y244A and I400G bound to +T DNA was monitored by mixing equal volumes of protein and DNA with ATP and MDCC-PBP (final concentrations: 0.5 µM MutS, 1 µM +T DNA, 500 µM ATP, 8 µM MDCC-PBP). (a) The stopped-flow trace of Y244A in the presence of +T DNA reveals only a slight suppression of the burst rate ($k_{\text{burst}} = 3 \text{ s}^{-1}$) with a burst amplitude of 0.5 P_i per MutS dimer followed by the same steady state turnover rate as seen without DNA ($k_{\text{cat}} = 0.1 \text{ s}^{-1}$), indicating that the mutant catalyzes rapid ATP hydrolysis even in the presence of mismatched +T DNA, and the rate-limiting step remains a step after P_i release. (b) The stopped-flow trace of +T DNA bound I400G similarly shows an ~2-fold suppression of the burst rate ($k_{\text{burst}} = 4 \text{ s}^{-1}$) with a burst amplitude of 0.7 P_i per MutS dimer, as well as an ~2-fold higher steady state rate of 0.3 s^{-1}, again indicating that the mutant catalyzes rapid ATP hydrolysis and the rate-limiting step remains after P_i release, even in the presence of +T DNA.
<table>
<thead>
<tr>
<th>Protein</th>
<th>( k_{\text{burst}} ) (s(^{-1}))</th>
<th>( k_{\text{cat}} ) (s(^{-1}))</th>
<th>( k_{\text{burst}} ) (s(^{-1}))</th>
<th>( k_{\text{cat}} ) (s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>8</td>
<td>0.9</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Y244A</td>
<td>5</td>
<td>0.6</td>
<td>0.2</td>
<td>3</td>
</tr>
<tr>
<td>I400G</td>
<td>7</td>
<td>1.0</td>
<td>0.2</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3.6. MutS phosphate release in the presence and absence of +T DNA. \( P_i \) release data with and without +T DNA shows only a partial suppression of the burst of ATP hydrolysis for Y244A and I400G.
The kinetics of WT $P_i$ release show a loss of the burst of ATP hydrolysis with +T DNA, consistent with previous findings that mismatch binding alters the ATPase cycle such that the slow step becomes a step prior to $P_i$ release. Interestingly, for Y244A and I400G, binding +T DNA does not inhibit the ATPase activity and the mutants continue to show a burst of $P_i$ release at 3 s$^{-1}$ and 4 s$^{-1}$, respectively (Figure 3.20). Both mutants also show slight reductions in burst amplitude to sub-stoichiometric levels, indicating that a population of the sample is hydrolyzing ATP at this fast rate while some portion may be inactive or only hydrolyze at the steady state turnover rate (Table 3.6). The lack of full suppression of the burst of ATP hydrolysis corroborates the findings with MutS$_{IAANS}$ Y244A and I400G where the proteins were found to exhibit altered conformational dynamics with ATP and +T DNA (see Figure 3.16). Additionally, these mutants were found to bind DNA like WT, and therefore the lack of change in ATPase activity with +T DNA cannot be attributed to an inability to bind DNA (see Tables 3.1, 3.2, and 3.3). Taken together, these data indicate that Y244A and I400G are not responding to the +T insertion like WT—instead, they hydrolyze ATP rapidly, likely unbinding the +T insertion, and stimulating the release of ADP, at which point they may rebind ATP and then re-associate with the DNA. Rather than remaining ATP-bound and undergoing the conformational change to the sliding clamp, these mutants could be cycling on and off the DNA, losing their ability to signal to MutL proteins to carry out the downstream repair. Given how precisely the ATPase activity is coupled to MutS actions on DNA and corresponding conformational changes, these differences could have a significant impact on MutS activity in MMR.
Discussion of Results

The above results can be used to better understand the DNA binding activity, conformational dynamics, and the ATPase cycle of the MutS mutants. A comprehensive review of the data offers preliminary interpretation of the behavior of each mutant, as described below. The locations of the mutant residues shown previously in Figures 1.8 to 1.14 are shown again in Figure 3.21 to aid in the description.
Figure 3.21. Seven LS-linked variants of MutS were included in this study. (a) The point mutations selected are mapped onto the Taq MutS crystal structure (PDB ID: 1NNE). The MutS mutants are (b) T113R, (c) G222D, (d) F243S, (e) Y244A, (f) I400G, (g) G434R, and (h) L533R.
**T113R** – From the experiments conducted thus far, T113R shows altered DNA binding activity compared to WT. From the 2-AP +T DNA and MutS\textsubscript{IAANS} binding experiments T113R appears to bind a +T insertion ~3-fold faster than WT (see Figures 3.6 and 3.9). Additionally, the MutS\textsubscript{IAANS} experiment revealed that the initial collision complex has a 5-fold higher affinity than that of WT ($K_D = 0.5 \ \mu\text{M}$ compared with $K_D = 2.5 \ \mu\text{M}$; Table 3.3) and undergoes a slightly slower rate-limiting conformational change to the mismatch specific complex ($k_{i(\text{max})} = 12 \ \text{s}^{-1}$ compared with $k_{i(\text{max})} = 21 \ \text{s}^{-1}$; Table 3.3). These differences in DNA binding activity indicate faster binding to DNA and the formation of a tighter initial collision complex, however T133R appears to undergo the conformational change to the mismatch-specific complex slower than WT. The higher affinity non-specific complex and slower rate of formation of the mismatch specific complex predicts that the mutant may have an increased affinity for homoduplex DNA, potentially resulting in a loss of mismatch specificity. Once the mutant forms the stable complex with the +T insertion, it appears to be fully functional, hydrolyzing ATP and undergoing the correlated conformational changes like WT. The altered DNA binding activity could be due to the introduction of an arginine residue adjacent to Asp96, which possibly results in the formation of an attractive interaction that potentially alters the structure or dynamics of the protein so as to change its ability to bind DNA. The activity of this mutant on matched DNA needs to be monitored in the future in order to determine if it has altered mismatch searching behavior.
**G222D** – Thus far, the mutant G222D appears to have only slight alterations in its mechanism, however these could have a compounding effect leading to impaired DNA repair activity. The two DNA binding experiments showed that G222D binds the +T insertion 2-3-fold faster than WT (Figures 3.6 and 3.10). Additionally, the assay with MutS\textsubscript{IAANS} G222D revealed differences in the MutS-mismatch complex formation where the mutant showed an ~3-fold weaker affinity initial collision complex ($K_D = 6.3 \mu$M compared with $K_D = 2.5 \mu$M; Table 3.3) and appears to undergo the conformational change to the tight mismatch-bound complex faster than WT ($k_{i{\text{max}}} = 43 \text{ s}^{-1}$ compared with $k_{i{\text{max}}} = 21 \text{ s}^{-1}$; Table 3.3). This difference in behavior may indicate that G222D binds the +T insertion quickly and then rapidly undergoes the conformational change to the specific mismatch-bound complex at a rate faster than WT. Additionally, although G222D showed similar rates of the conformational changes in response to ATP, the amplitudes of the traces were lower, which may indicate that a population of the proteins are compromised in their ability to undergo these conformational changes or the mutation may be altering the extent of the conformational changes. In either case, the mutant may have slightly altered conformational dynamics compared with WT, which could disrupt its ability to interact with downstream MutL. These findings are in line with the hypothesis that the loss of a flexible glycine residue and its replacement with a negatively charged aspartic acid residue may create a repulsive interaction with the nearby Glu210 residue and potentially destabilize this region. Gly222 was also identified by MD simulations as potentially interacting with Arg221 for allostERIC signaling [46]. Together, these data indicate that this mutation may be hindering the conformational
flexibility of the domain or the protein on a gross level and/or altering the signaling pathways between domains.

**F243S** – With the assays tested thus far, F243S does not appear to have any significant differences from WT. From the DNA binding experiments, F243S was found to have WT-like activity (see Tables 3.1, 3.2, and 3.3). Similarly, the rates of the conformational changes with ATP appear to be identical to WT (see Figure 3.12 and 3.15). Previous investigation into the activity of this mutant found that the rates of dissociation from the +T insertion both in the presence and absence of ATP as well as the rates of phosphate release before and after binding +T DNA are indistinguishable from WT [97]. Although this variant is associated with increased carcinogenesis in patients, the mutation appears to not compromise the steps of the MutS mechanism monitored thus far. Future experiments are needed to test the ability of F243S to activate downstream MutL activity, which could be responsible for its hindered DNA repair capability.

**Y244A** – Some notable differences in the mechanism of action of Y244A have emerged in this study. The DNA binding activity indicates that Y244A is fully capable of recognizing and binding a +T insertion and does so at rates and with affinities comparable to WT (see Tables 3.1, 3.2, and 3.3). The differences in activity appear when ATP is introduced. In the absence of DNA, the mutant binds ATP and undergoes the same associated conformational change as WT, however, it hydrolyzes ATP 2-fold slower than WT and the concerted conformational change occurs at the
same reduced rate (see Figures 3.13 and 3.18). This indicates an altered ATPase cycle of the mutant, even prior to the introduction of DNA. After Y244A forms a complex with +T DNA, its response to ATP is significantly different from WT. The Pi release assay shows that the presence of +T DNA does not diminish the burst of ATP hydrolysis at the high-affinity ATPase site (see Figure 3.20). These results are recapitulated in the assay monitoring the associated conformational changes and show that domains I release the DNA at the same rate as WT, but after releasing the DNA it undergoes a conformational change that is not consistent with the formation of the sliding clamp (see Figure 3.16). From these data, it appears that Y244A continues to hydrolyze ATP after recognizing a mismatch, which causes it to dissociate from the DNA. These results are in line with the initial hypothesis that this residue contributes to allosteric communication between the DNA binding and ATPase sites and its mutation may disrupt signals that are critical for MutS function [46]. Due to the apparent inability of Y244A to form the sliding clamp, it is predicted that this mutant is compromised in its ability to activate MutL nicking activity, which will be experimentally tested in the future (see Future Directions below).

**I400G** – MutS I400G also shows altered conformational dynamics and ATPase activity when the mismatch-bound protein encounters ATP. The DNA binding experiments reveal that I400G binds the +T insertion at the same rate as WT and forms similar initial collision and mismatch-specific complexes (see Tables 3.1, 3.2, and 3.3). Additionally, in the presence of ATP, I400G shows WT-like rates of ATP hydrolysis and associated conformational changes (see Figure 3.18 and 3.12).
and a previous study found that I400G dissociates from the +T insertion at the same rate as WT when no ATP is present [97]. After +T DNA is introduced, however, the mutant shows only a partial suppression of the burst of ATP hydrolysis, indicating that the high-affinity site continues to rapidly hydrolyze ATP, even after binding the +T insertion (Figure 3.20). This finding is consistent with the results of the assay monitoring the conformational changes of I400G where in the presence of +T DNA and ATP, the mutant seems unable to undergo the conformational change to the sliding clamp necessary for downstream signaling to MutL, as well as a previous study that found that I400G dissociates from the +T insertion ~7-fold faster than WT in the presence of ATP. Like with Y244A, these results point to an altered mechanism of action of I400G where instead of undergoing the slow conformational change to the sliding clamp, it appears to continue to hydrolyze ATP and rapidly dissociate from the mismatch. These results are consistent with the hypothesis that the Ile400 residue contributes to allosteric communication between the mismatch binding and ATPase sites and its mutation may disrupt signals that are critical for MutS function [46]. As with Y244A, I400G is predicted to have a reduced ability to activate MutL, which again will be tested in future studies (see Future Directions below).

**G434R** – Thus far, G434R appears to be altered only in its DNA binding activity. Although the bimolecular association rate found with the 2-AP +T DNA binding experiment is close to WT, the MutS\textsubscript{IAANS} binding assay indicates that the estimated on-rate is nearly 4-fold faster (see Tables 3.2 and 3.3). This faster rate of association results in a much tighter initial collision complex (apparent $K_D = 0.2 \ \mu M$
compared with 2.5 µM; Figure 3.9) as well as a higher affinity mismatch-specific complex ($K_D = 5.8$ nM compared with 19.7 nM; Figure 3.3). These results align with the prediction that introduction of a positively charged arginine residue in close proximity to the DNA backbone may allow for the formation of an additional stabilizing electrostatic contact with the DNA. This increase in affinity of the initial collision complex could indicate a greater affinity for homoduplex DNA as well, which may lead to a decrease in the ability of the G434R mutant to discriminate between matched DNA and mismatched errors. As with T113R, the activity of this mutant on matched DNA needs to be investigated in order to determine if its scanning behavior is altered.

**L533R** – From the assays conducted in this study, L533R does not appear to have any major alterations to its mechanism of action. Although the 2-AP +T DNA binding experiment showed a two-fold faster bimolecular association rate ($k_{on} = 10.7 \times 10^6$ M$^{-1}$ s$^{-1}$ compared with $5 \times 10^6$ M$^{-1}$ s$^{-1}$; Figure 3.6), this slightly faster rate of association was not seen in the binding experiment with the IAANS-labeled protein, and therefore more data is needed to determine if the differences in DNA binding are significant. In the presence of ATP, the mutant appears to undergo conformational changes at rates consistent with WT, however with lower amplitudes. This trend holds with the addition of +T DNA. The loss of signal change for the mutant protein may either indicate subtle differences in the conformational dynamics of the protein or the loss of activity in a subset of the population. If these slight alterations do correspond with differences in the conformations of the protein, it may lead to a
reduced ability of L533R to interact with MutL. Future experiments of the MutL activation activity by L533R will test this hypothesis.
IV. CONCLUSIONS

The results of this study serve to identify alterations in the MutS mechanism of action of several Lynch syndrome-linked MutS variants. The most notable finding is the change in the ATPase activity and accompanying conformational dynamics of the Y244A and I400G mutants. Both proteins appear to respond to the presence of +T DNA differently than WT MutS—they do not form the stable ATP-bound sliding clamp but rather continue to rapidly hydrolyze ATP and release the DNA, potentially indicating a loss of the ability to signal downstream MMR after identifying a mismatch. T113R shows faster and tighter initial association with the DNA, possibly due to a higher nonspecific affinity for DNA. Similarly, G434R appears to bind the +T insertion faster and forms a higher affinity mismatch-specific complex, which may result in altered error recognition ability. G222D binds faster and shows slightly different conformational dynamics in response to +T DNA, possibly hindering its ability to signal to MutL. With the assays performed thus far, F243S and L533R behave nearly identically to WT, and therefore their activities beyond mismatch recognition, ATP hydrolysis, and the associated conformational dynamics must be examined for potential disruptions. In summary, this study provides the first experimental demonstration of how point mutations in MutS that are distant from the two active sites may disrupt its function; it offers mechanistic explanations for the linkage between these mutations and Lynch cancer phenotype; it generates specific testable hypotheses for future analysis of these mutations; and finally, it helps create a blueprint for detailed mechanistic analysis of other LS-linked mutants.
V. **Future Directions**

The aim of this study was to employ assays developed by the Hingorani lab to investigate the activities of several cancer-linked variants of MutS during initiation of MMR as a means to identify changes in their mechanism that may underlie their association with the cancer phenotype. While this study allowed for a comprehensive survey of the activities of seven MutS mutants, additional tests are needed to construct the full mechanism of action of each mutant protein. The future goals of this project include obtaining a full set of data for the DNA binding, conformational dynamics, and ATPase activities of the sixteen mutant proteins initially identified for this study. The analysis would include conducting the assays carried out in this study, and performing additional measurements such as the rate of MutS dissociation from a +T insertion with and without ATP and the P\(_i\) release kinetics of all mutants.

Thus far we have identified a few mutant proteins, particularly Y244A and I400G, that are worth pursuing with a full panel of assays. This includes repeating the experiments carried out in this study under more rigorous conditions suited to the specific activity of each protein, as well as full titrations for more accurate measures of kinetic and thermodynamic parameters. In addition to acquiring a more complete data set for these mutants, several additional experiments should be carried out on all of the mutant proteins to monitor their downstream activities and actions in MMR. For example, an assay has been recently developed by the Hingorani lab to test the MutS-dependent activation of MutL. Briefly, a 3kb plasmid has been engineered to contain a single +T insertion, which can be used for gel-based analysis of MutL endonuclease activity at 70°C, the physiologically relevant temperature for *Taq*
proteins. This assay will allow us to test the ability of the MutS mutants to stimulate MutL activity and determine whether there is any defect, either from disruption in ternary complex formation and/or communication between the two proteins to stimulate nicking of DNA by MutL.

Additional structural analysis of the MutS mutants is also warranted. For example, the integrity and stability of MutS can be assessed by a circular dichroism (CD) melt, which would allow the monitoring of global destabilization effects of the mutations on the structure. Future in vivo studies could also be used to test if the altered activity of these mutant proteins gives rise to known genomic markers of LS such as microsatellite instability. Overall, these experiments can lead to a more complete picture of how alterations in the mechanism of LS-linked mutants of MutS results in defective MMR function and moreover, how they contribute to disease phenotypes.
### APPENDICES

*Appendix A. MutS sequence alignment*

<table>
<thead>
<tr>
<th>TagMutS</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HsaMSH6</td>
<td>1</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>1</td>
</tr>
<tr>
<td>consensus</td>
<td>1</td>
</tr>
<tr>
<td>TagMutS</td>
<td>81</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>81</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>81</td>
</tr>
<tr>
<td>consensus</td>
<td>81</td>
</tr>
<tr>
<td>TagMutS</td>
<td>161</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>161</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>161</td>
</tr>
<tr>
<td>consensus</td>
<td>161</td>
</tr>
<tr>
<td>TagMutS</td>
<td>241</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>241</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>241</td>
</tr>
<tr>
<td>consensus</td>
<td>241</td>
</tr>
<tr>
<td>TagMutS</td>
<td>321</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>321</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>321</td>
</tr>
<tr>
<td>consensus</td>
<td>321</td>
</tr>
<tr>
<td>TagMutS</td>
<td>401</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>401</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>401</td>
</tr>
<tr>
<td>consensus</td>
<td>401</td>
</tr>
<tr>
<td>TagMutS</td>
<td>471</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>471</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>471</td>
</tr>
<tr>
<td>consensus</td>
<td>471</td>
</tr>
<tr>
<td>TagMutS</td>
<td>551</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>551</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>551</td>
</tr>
<tr>
<td>consensus</td>
<td>551</td>
</tr>
<tr>
<td>TagMutS</td>
<td>631</td>
</tr>
<tr>
<td>HsaMSH6</td>
<td>631</td>
</tr>
<tr>
<td>HsaMSH2</td>
<td>631</td>
</tr>
<tr>
<td>consensus</td>
<td>631</td>
</tr>
</tbody>
</table>

---

161
### Appendix B. MutS primer sequences

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T113R</td>
<td>5’ – GGG TGA GGA GCT GCC TCA CCT CCC GCC T – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – CCC ACT CCT CGA CGG AGT GGA GGG CGG A – 5’</td>
</tr>
<tr>
<td>G222D</td>
<td>5’ – GGC CAG AAG GGC ATC CCG GGC CCC CC – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – CCG GTC TTC CCG TAG GGC CCG GCC GG – 5’</td>
</tr>
<tr>
<td>F243S</td>
<td>5’ – CCC GGG GTC GTA GCT GCG GAA AGG CTG G – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – GGG CCC CAG CAT CGA CGC CTT TCC GAC C – 5’</td>
</tr>
<tr>
<td>Y244A</td>
<td>5’ – GCC TTT CCG CTT CGC CGA CCC CGG GCC GC – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – GGC CCC GGG GTG GCC GGC GAA GCG GAA AGG C – 5’</td>
</tr>
<tr>
<td>I400G</td>
<td>5’ – GTA CCC TTC CCG GCC GAG CCC CCC CTC G – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – CAT GGG AAG GGC CGG CTC GGG GGG GAG C – 5’</td>
</tr>
<tr>
<td>G434R</td>
<td>5’ – GGG TGG GGA TGC GGC TCC TCT CTC T – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – CCC ACC CCT ACG CGC AGG AGA GAG A – 5’</td>
</tr>
<tr>
<td>L533R</td>
<td>5’ – GGC CAC CTC GGC CCT GGC GCC GTA GAC G – 3’</td>
</tr>
<tr>
<td></td>
<td>3’ – CCG GTG GAG CCG GGA CCC GGC CAT CTG C – 5’</td>
</tr>
</tbody>
</table>
Appendix C. DNA substrate sequences

<table>
<thead>
<tr>
<th>Duplex Name</th>
<th>Strand Name</th>
<th>Strand Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>+T</td>
<td>36-mer</td>
<td>3’ – TAC ACT TAG TCA TAC CAT TAT AGA CGA CTT CCT TTA – 5’</td>
</tr>
<tr>
<td></td>
<td>37-mer</td>
<td>5’ – ATG TGA ATC AGT ATG GTA \textbf{T} ATA TCT GCT GAA GGA AAT – 3’$^*$</td>
</tr>
<tr>
<td>2-AP +T</td>
<td>36-mer</td>
<td>3’ – TAC ACT TAG TCA TAC CAT TAT AGA CGA CTT CCT TTA – 5’</td>
</tr>
<tr>
<td></td>
<td>37-mer</td>
<td>5’ – ATG TGA ATC AGT ATG GTA \textbf{T} \textbf{T} TA TCT GCT GAA GGA AAT – 3’$^{*†}$</td>
</tr>
</tbody>
</table>

$^*$ \textbf{T} denotes +T insertion

$^{†}$ \textbf{T} denotes 2-aminopurine nucleotide
### Appendix D. List of buffers

<table>
<thead>
<tr>
<th>Buffer Name</th>
<th>Buffer Contents</th>
</tr>
</thead>
</table>
| Cell resuspension buffer | 10 mM Tris-HCl, pH 7.5  
                          | 10% sucrose                                               |
|                      | 50 mM NaCl                                                |
| B₀                   | 50 mM Tris-HCl, pH 7.5  
                          | 1 mM EDTA                                                 |
|                      | 5% glycerol                                               |
| B₅₀                  | 50 mM Tris-HCl, pH 7.5  
                          | 1 mM EDTA                                                 |
|                      | 5% glycerol                                               |
|                      | 50 mM NaCl                                                |
| B₄₀₀                 | 50 mM Tris-HCl, pH 7.5  
                          | 1 mM EDTA                                                 |
|                      | 5% glycerol                                               |
|                      | 400 mM NaCl                                               |
| Buffer L             | 50 mM Tris-HCl, pH 7.5  
                          | 50 mM NaCl                                                |
|                      | 5 mM EDTA                                                 |
|                      | 5% glycerol                                               |
| 1X TBE buffer        | 90 mM Tris-HCl, pH 7.5  
                          | 90 mM Boric acid                                          |
|                      | 2 mM EDTA                                                 |
|                      | 45% Formamide                                             |
| DNA dialysis buffer  | 50 mM Tris-HCl, pH 7.5  
                          | 50 mM NaCl                                                |
|                      | 1 mM EDTA                                                 |
| Annealing buffer     | 20 mM Tris-HCl, pH 8.8  
                          | 100 mM NaCl                                               |
| Binding buffer       | 20 mM Hepes-NaOH, pH 7.5  
                          | 50 mM NaCl                                                |
|                      | 5 mM MgCl₂                                                |
| P₁ release buffer    | 30 mM Hepes-NaOH, pH 7.5  
                          | 100 mM NaCl                                               |
|                      | 10 mM MgCl₂                                               |
|                      | 200 µM 7-MEG                                              |
|                      | 0.01 U/mL PNPase                                           |
Appendix E. Abbreviations and symbols

° degree(s)
2-AP 2-aminopurine
1D one-dimensional
3D three-dimensional
7-MEG 7-methylguanosine
Å Ångstrom(s)
ΔΔG change in Gibb’s free energy
ε extinction coefficient
λ wavelength
μg microgram
μL microliter(s)
μm micrometer
μM micromolar
χ² chi squared
A adenine
A alanine
ABC ATP binding cassette
ADP adenosine 5’-diphosphate
ADP-BeF₃ adenosine 5’-diphosphate beryllium fluoride
ADP-PNP adenylyl-imidodiphosphate
apo apoprotein
Ala alanine
Amp ampicillin
Arg arginine
Asp aspartatic acid
ATP adenosine 5’-triphosphate
ATPγS adenosine 5’-(3-thiotriphosphate)
BSA bovine serum albumin
C carboxyl
C Celsius
C cysteine
C cytosine
CD circular dichroism
cm centimeter
cys cysteine
CV column volume(s)
D aspartate
Da dalton(s)
ddH₂O  double-distilled water
DMF    dimethylformamide
DNA    deoxyribonucleic acid
dNTP   deoxynucleotide
DTT    dithiothreitol
E      enzyme
E      glutamate
E. coli Escherichia coli
EDTA   ethylenediaminetetraacetic acid
F      phenylalanine
Fi     filtrate
FRET   Förster resonance energy transfer
g     gram(s)
G      glycine
G      guanine
Glu    glutamic acid
Gly    glycine
h      human
HCl    hydrochloric acid
Hepes  4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HNPCC  hereditary nonpolyposis colorectal cancer
I      isoleucine
IAANS  2-[4′-(iodoacetamido)anilino]naphthalene-6-sulfonic acid
ID     identifier
IDL    insertion-deletion loop
Ile    isoleucine
IPTG   isopropyl β-D-1-thiogalactopyranoside
InSiGHT International Society for Gastrointestinal Hereditary Tumours
K      thousand
kb     kilobase
k₁₅     burst rate
kcal   kilocalorie
kcat    rate of catalytic turnover
K_D    dissociation constant
kDa    kilodalton(s)
k_{down} rate of down phase
ki     rate of first phase
k_{ii}  rate of second phase
K_M    Michaelis constant
k_{obs} observed rate of binding
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{\text{off}}$</td>
<td>rate of dissociation</td>
</tr>
<tr>
<td>$k_{\text{on}}$</td>
<td>rate of binding</td>
</tr>
<tr>
<td>$k_{\text{ss}}$</td>
<td>steady state rate</td>
</tr>
<tr>
<td>$k_{\text{up}}$</td>
<td>rate of up phase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton(s)</td>
</tr>
<tr>
<td>L</td>
<td>leucine</td>
</tr>
<tr>
<td>L</td>
<td>liter(s)</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>Leu</td>
<td>leucine</td>
</tr>
<tr>
<td>LS</td>
<td>Lynch syndrome</td>
</tr>
<tr>
<td>M</td>
<td>methionine</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>MD</td>
<td>molecular dynamics</td>
</tr>
<tr>
<td>MDCC</td>
<td>7-diethylamino-3-(((2-maleimidyl)ethyl)amino)carbonyl)coumarin</td>
</tr>
<tr>
<td>Met</td>
<td>methionine</td>
</tr>
<tr>
<td>mg</td>
<td>milligram(s)</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>mL</td>
<td>milliliter</td>
</tr>
<tr>
<td>MLH</td>
<td>MutL homolog</td>
</tr>
<tr>
<td>mm</td>
<td>millimeter</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MMR</td>
<td>mismatch repair</td>
</tr>
<tr>
<td>MSH</td>
<td>MutS homolog</td>
</tr>
<tr>
<td>MSI</td>
<td>microsatellite instability</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut-off</td>
</tr>
<tr>
<td>N</td>
<td>amino</td>
</tr>
<tr>
<td>N</td>
<td>nitrogen</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram(s)</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer(s)</td>
</tr>
<tr>
<td>nM</td>
<td>nanomolar</td>
</tr>
<tr>
<td>nmol</td>
<td>nanomole(s)</td>
</tr>
<tr>
<td>OD$_{600}$</td>
<td>optical density at 600 nm</td>
</tr>
<tr>
<td>oligo</td>
<td>oligonucleotide</td>
</tr>
<tr>
<td>P</td>
<td>product</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBP</td>
<td>phosphate binding protein</td>
</tr>
</tbody>
</table>
PCNA  proliferating cell nuclear antigen
PCR  polymerase chain reaction
PDB  Protein Data Bank
Phe  phenylalanine
$P_i$  inorganic phosphate
$pL$  labeled protein
PMS  postmeiotic segregation increased
PMSF  phenylmethanesulfonyl fluoride
PMT  photo-multiplier tube
PNPase  purine nucleoside phosphorylase
Pro  proline
psi  pound(s) per square inch
$PU$  unlabeled protein
R  arginine
R  retentate
RNA  ribonucleic acid
RPA  Replication Protein A
rpm  revolution(s) per minute
s  second(s)
S  serine
S  substrate
$S_0$  ground state
$S_1$  first excited state
$S_2$  second excited state
S. cerevisiae  *Saccharomyces cerevisiae*
smFRET  single molecule Förster resonance energy transfer
SDS  sodium dodecyl sulfate
SSB  single-stranded DNA binding protein
T  threonine
T  thymine
TAMRA  5-Carboxytetramethylrhodamine
Taq  *Thermus aquaticus*
TBE  1,1,2,2-tetrabromoethane
TCEP  tris(2-carboxyethyl)phosphine
Thr  threonine
Tris  2-amino-2-(hydroxymethyl)-1,3-propanediol
Trp  tryptophan
Tyr  tyrosine
U  units
UV  ultraviolet
V  valine
V  volt(s)
Val  valine
VUS  variants of uncertain significance
W  tryptophan
WT  wildtype
Y  tyrosine
REFERENCES


al. MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto 
DNA. Elife. 2015;4:e06744.

2016;539:583-7.

imaging reveals target-search mechanisms during DNA mismatch repair. Proc Natl 

between two fundamentally distinct clamps during mismatch repair. Nature Structural 


Signaling in MutS and MSH Homologs via Switching of Dynamic Communication 

and biochemical analysis of the MutS.ADP.beryllium fluoride complex suggests a 
2003;278:16088-94.

mediates heteroduplex loop formation by a translocation mechanism. Embo j. 
1997;16:4467-76.

hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. 

[79] Pluciennik A, Modrich P. Protein roadblocks and helix discontinuities are 
barriers to the initiation of mismatch repair. Proceedings of the National Academy of 
Sciences. 2007;104:12709.

Structure of the MutL C-terminal domain: a model of intact MutL and its roles in 


