Developing a Novel Fluorescent Reporter to Monitor Conformational Change of MutS, a Mismatch Repair Protein

by

Christopher W. Doucette
Class of 2011

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, 2011
Acknowledgements.

I would like to thank my family for their continual love, support, and late-night kitchen dance parties. Thank you 77 Home (you too, Jengo) for bearing with me during the stressful times, and more importantly for always finding new ways to make me smile. I would also like to thank the MB&B faculty and staff, and all members of the Hingorani lab (past and present) for their companionship, especially F. Noah Biro for patiently taking me from pipetting water to purifying proteins, and for providing me with both intellectual and personal support throughout my time as a young scientist. Thank you Mom and Dad for understanding when I can’t always call due to late-nights in the lab, and know that I feel incredibly lucky to have your complete love and support. Thank you Jonathan, it’s been incredibly exciting to grow and learn with you these past four years. Finally, Manju, thank you for the immense amount of time and energy you have invested into my education and growth as a researcher. My experience at Wesleyan would truly not have been the same without you!
Abstract.

DNA mismatch repair (MMR) is an evolutionarily conserved process that locates and repairs post-replicative errors in nascent DNA and is critical for maintaining the genome integrity. The first step in the MMR pathway is recognition of the base-base mismatch/insertion deletion loop (IDL) by MutS in prokaryotes and the MutS homologues, Msh2-Msh6 and Msh2-Msh3, in eukaryotes. While there is a significant body of literature concerning DNA-binding and ATPase activities of MutS, information on MutS conformational changes involved in the recognition of an array of non-Watson-Crick DNA structures is lacking. Here, we present a novel fluorescence reporter developed to investigate the kinetics of Thermus aquaticus MutS conformational changes upon mismatch binding from the perspective of the protein for the first time. M88 in the DNA binding domain was replaced by cysteine and labeled with a thiol reactive fluorophore that is sensitive to protein conformation. Our assay reports that MutS binds to a mismatch via a two-step binding event characterized by initial fast binding ($k_{on} = 5\times10^6$ M$^{-1}$sec$^{-1}$) step followed by a saturating step. Interestingly, our experiments show that the saturating phase is absent and the fast binding is slower ($k_{on} = 0.76\times10^6$ M$^{-1}$sec$^{-1}$) for a longer DNA substrate, suggesting that matched base-pairs affect MutS recognition of a mismatch. Additionally, we have begun characterizing the effects of nucleotides on the MutS recognition event. Having characterized the rate constants governing MutS interaction with mismatched DNA with and without nucleotides, we have gained novel mechanistic insights into MutS initiation of MMR and developed an effective reporter to monitor MutS action from the perspective of the protein.
Table of Contents.

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Vita</td>
<td>3</td>
</tr>
<tr>
<td>List of Figures</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>6</td>
</tr>
<tr>
<td>1: Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2: Materials and Methods</td>
<td>44</td>
</tr>
<tr>
<td>3: Results</td>
<td>63</td>
</tr>
<tr>
<td>4: Discussion</td>
<td>91</td>
</tr>
<tr>
<td>5: Future Directions</td>
<td>108</td>
</tr>
<tr>
<td>Appendices</td>
<td></td>
</tr>
<tr>
<td>A: DNA oligos</td>
<td>109</td>
</tr>
<tr>
<td>B: Abbreviations and Symbols</td>
<td>110</td>
</tr>
<tr>
<td>C: List of Equations</td>
<td>114</td>
</tr>
<tr>
<td>References</td>
<td>115</td>
</tr>
</tbody>
</table>
List of Figures.

<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Minimal model of prokaryotic mismatch repair</td>
<td>16</td>
</tr>
<tr>
<td>1.2</td>
<td>Minimal model of eukaryotic mismatch repair</td>
<td>21</td>
</tr>
<tr>
<td>1.3</td>
<td>Structure of the prokaryotic MutS protein</td>
<td>25</td>
</tr>
<tr>
<td>1.4</td>
<td>Structure of the MutL protein homologues</td>
<td>31</td>
</tr>
<tr>
<td>1.5</td>
<td>Structure of the prokaryotic MutH protein</td>
<td>34</td>
</tr>
<tr>
<td>1.6</td>
<td>Summary of mismatch repair models</td>
<td>39</td>
</tr>
<tr>
<td>MATERIALS &amp; METHODS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Labeling of MutS with IAANS at the M88C position</td>
<td>53</td>
</tr>
<tr>
<td>RESULTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>The M88 residues are in proximity when MutS is bound to a mismatch</td>
<td>64</td>
</tr>
<tr>
<td>3.2</td>
<td>MutS binds with high affinity to DNA containing a +T insertion as reported by different fluorescence-based assays</td>
<td>68</td>
</tr>
<tr>
<td>3.3</td>
<td>MutS exhibits two-step DNA binding kinetics</td>
<td>73</td>
</tr>
<tr>
<td>3.4</td>
<td>MutS\textsubscript{M88C-IAANS} binds tightly to a longer DNA substrate</td>
<td>79</td>
</tr>
<tr>
<td>3.5</td>
<td>MutS exhibits relatively slow 1-step binding kinetics on a longer DNA substrate</td>
<td>82</td>
</tr>
<tr>
<td>3.6</td>
<td>ATP binding is slower and hydrolysis is suppressed when MutS is bound to a mismatch</td>
<td>88</td>
</tr>
</tbody>
</table>
### DISCUSSION

4.1 Previously observed saturation event monitored by a 2-AP DNA substrate  

4.2 MutSM88C-IAANS binds at an apparently slower rate to a substrate with a higher concentration of matched base pairs  

4.3 Mismatched DNA inhibits rapid ATP hydrolysis
List of Tables.

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Germline mutations of MMR genes underlie many cancers</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Prokaryotic mismatch repair proteins and their eukaryotic homologues</td>
<td>12</td>
</tr>
<tr>
<td>RESULTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Rate constants for MutS interaction with nucleotides in the presence or absence of mismatched DNA</td>
<td>90</td>
</tr>
</tbody>
</table>
1: Introduction.

Overview of Mismatch Repair

The double helix structure of DNA enables complete storage and transfer of genetic information through efficient and accurate replication by enzymes that utilize one of the strands as a ‘template’ with which to synthesize a new complementary ‘daughter’ strand by strict nucleotide pairing of adenine with thymine and guanine with cytosine. DNA polymerase, the enzyme charged with the task of pairing complementary bases during the extension of the daughter strand, maintains high fidelity by its high selectivity for Watson-Crick base pairs and its ability to carefully proofread each incorporated nucleotide pair and excise incorrect base pairs by its endonuclease activity; however, certain tautomeric forms of nucleotides can form mispaired bases that escape polymerase proofreading, and segments of the genome containing poly-nucleotide repeats are prone to polymerase slippage, which can lead to extrahelical insertions or deletions in the daughter strand. In addition to these biosynthetic errors, the duplex is also constantly bombarded with exogenous mutating agents including UV radiation, chemical modification, and strand breakage. What was previously thought to be a relatively static entity, the DNA helix structure has more recently been shown to be a highly dynamic structure, constantly being compromised and repaired by a variety of cellular mechanisms.

During a typical round of DNA replication, the replicase inserts a base-base mismatch or small insertion deletion loop (IDL) one in every $10^6$ nucleotides, an error rate that inevitably leads to genomic instability when one takes into consideration the vast size of the human genome, comprised of over six billion bases pairs (Kunkel and Bebenek 2000). An evolutionarily conserved process known as DNA mismatch repair
(MMR) initiates the excision and repair of these post-replicative errors and additionally has been shown to facilitate in a variety of other crucial cellular processes such as the recognition of damaged/oxidative bases, such as O\textsuperscript{6}-methylguanine, that results in apoptosis, or programmed cell death. The MMR pathway decreases biosynthetic mutational errors up to 1000-fold (Stojic, Brun et al. 2004; Hays, Hoffman et al. 2005) and the importance of this system is further emphasized by the fact that a variety of mutations in MMR genes, which render non-functioning proteins, underlie the Lynch Syndrome (Hereditary Non-Polyposis Colorectal Cancer; HNPCC), and are also associated with sporadic tumor growth (Peltomaki and Vasen 2004). Germline mutations in a number of MMR genes and their correlated cancer prognoses are listed in Table 1.1.

Information regarding the mechanism through which MMR proteins process base-base mismatches and IDLs first surfaced in early genetic experiments and increased exponentially once biologically active and purified forms of the MMR proteins were isolated and characterized. Today, crystal structures of these DNA binding proteins provide a view of the structures and even possible conformations these proteins must adopt in order to perform their crucial functions. However, these isolated snap shots provide little information of how these proteins function in solution and a detailed kinetic picture of how protein conformational changes and interactions relate to function is lacking. The dearth of knowledge limits our understanding of how MMR proteins work together to maintain the fidelity of the genome and safeguard against tumorigenesis.
<table>
<thead>
<tr>
<th>ORIGIN OF CANCER</th>
<th>MAMMALIAN MMR</th>
<th>PROKARYOTIC HOMOLOGUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Colorectum</strong></td>
<td>Msh2</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Msh6</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Mlh1</td>
<td>MutL</td>
</tr>
<tr>
<td><strong>Endometrium</strong></td>
<td>Msh2</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Mlh1</td>
<td>MutL</td>
</tr>
<tr>
<td><strong>Stomach</strong></td>
<td>Msh2</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Mlh1</td>
<td>MutL</td>
</tr>
<tr>
<td><strong>Gastric Adenoma</strong></td>
<td>Msh2</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Mlh1</td>
<td>MutL</td>
</tr>
<tr>
<td><strong>Ovary</strong></td>
<td>Msh2</td>
<td>MutS</td>
</tr>
<tr>
<td></td>
<td>Mlh1</td>
<td>MutL</td>
</tr>
</tbody>
</table>

Adapted from Peltomaki (2003).
First Evidence of a DNA Mismatch Repair System: Genetic Studies of

*Streptococcus pneumoniae*

The proposal of a cellular mechanism charged with the task of recognizing and subsequently repairing base-base mismatches in DNA first surfaced in 1966 when Ephrussi-Taylor’s laboratory noticed differences in transformation efficiencies of donor DNA in *Streptococcus pneumoniae*. Through further investigation, they found that these efficiencies correlated with differences in certain markers of the donor DNA (Ephrussi-Taylor et al. 1966). While the exact mechanism of genetic recombination, the mechanism through which successful transformation occurs, was not as well understood during this period, it was established that the uptake of double stranded DNA by the cell was immediately followed by degradation to a single strand intermediate in *S. pneumoniae* and this donor single strand would complement with a homologous portion of the recipient DNA through physical displacement of the recipient chromosomal duplex to form a heteroduplex intermediate species.

Ephrussi-Taylor measured transformation efficiency of various strains of *S. pneumoniae*, and through these experiments, their laboratory ultimately delineated two classes of transformants: high efficiency (HE) markers, which had an efficiency ratio (defined as the relative amount of the DNA marker integrated to the amount of genomic DNA taken up by the cell) of approximately 1 and low efficiency (LE) markers which yielded an efficiency 5 – 20 fold lower than the HE markers (Tiraby 1973). Additionally, LE markers were defined as having a point mutation in the donor DNA which yielded a heteroduplex DNA intermediate with a base-base mismatch. Ephrussi-Taylor proposed that there may in fact be a ‘Destruction Choice
Mechanism in which cells would go through a revision process after heteroduplex formation and preferentially remove DNA sequences containing base-base mismatches, which appeared to explain the differences in the transformation efficiencies of these two markers (Ephrussi-Taylor et al. 1966).

In 1973, Tiraby and Fox noticed that while most of their stock strains of *S. pneumoniae* had this ability to distinguish between HE and LE markers (based on similarly observed transformation efficiencies found in Ephrussi-Taylor’s laboratory), some strains incorporated LE markers with the same high efficiency as HE markers, and they deemed these strains *Hex−*. It was proposed that the ability of *Hex+* strains to distinguish between HE and LE markers may be a result of a cellular system that could distinguish between matched and mismatched DNA, preferentially attacking and excising the donor strand containing the point mutation. This repair process would excise the strand containing the mismatch and finally resynthesize the strand, using the recipient strand as a complement (Tiraby 1973). Further evidence for this proposed repair mechanism surfaced when Guild and Shoemaker showed that strains transformed with LE markers that survived produced homozygosity of the original recipient genome (Guild et al. 1976). This observation lead to the conclusion that the donor strand containing a point mutation was excised and resynthesized. While the *Hex−* strains were deemed to elicit a mutator phenotype, the specific genetic and biochemical consequences of these mutants were not fully understood until the early 1980s.

The first set of Hex genes were cloned and characterized by Lacks in 1985 (Lacks et al. 1986). Two recombinant plasmids of the two Hex genes, the HexA and
HexB genes, were transferred into *B. subtilis* and found to recognize strand breaks in DNA and preferentially excise markers in the 5’ to 3’ direction. It was proposed that perhaps these two genes signified two distinct biological functions, one potentially being involved in recognition of the error in DNA and the other in strand excision. While the Hex system of *S. pneumoniae* provided initial evidence of a mechanism involved in maintaining post-replicative genome integrity, and yielded gene products that facilitate this process, a slightly different repair pathway was discovered in the gram negative bacteria, *E. coli*. This pathway would later be shown to be mismatch repair, now known to be conserved in prokaryotes and eukaryotes, including humans.

**The Mut System: The Model Prokaryotic Mismatch Repair Pathway**

Similar to the discovery of the Hex system, the repair pathway of *E. coli* was first observed when strains of the bacteria were transfected with bacteriophage λDNA, specially prepared to generate a mismatch by annealing of two homologous strands containing a point mutation. Genetic analysis of progeny phages of these experiments showed the same transformation efficiency pattern of certain markers as in the Hex system, and confirmed that a mismatch repair system also existed in *E. coli*. This system was named the Mut pathway (Wildenberg et al. 1975). This proposed pathway helped to explain previously unexplained genetic recombination phenomena such as gene conversion and localized negative interference.

By 1980, Glickman and Radam, through further genetic studies, had characterized the mutator genes in mismatch repair of the Mut system. Similar to the *Hex* strains of *S. pneumoniae*, they found a variety of mutants which were found to
have an increase in base-base substitutions and frameshift mutations leading to the conclusion that these mutants had all but lost their abilities to repair post-replicative errors in DNA. The mutations were mapped onto three known mutator loci: mutS, mutL, and mutH (Glickman et al. 1980). Additionally, these experiments showed that a dam− strain, which lacked the enzyme responsible for methylating newly replicated DNA at palandromic GATC sites had the same mutator phenotype as the mut− strains, suggesting that methylation was a key step for the mismatch repair pathway for this gram-negative bacteria.

By the late 1980s, the three Mut genes had been isolated and purified into their biologically active gene products, and the role one performed in mismatch repair system began to be elucidated, especially through the elegant in vitro assays developed by the Modrich laboratory. The first of these gene products formally purified and characterized was MutS, a protein that was found to bind to base-base mismatches in a newly developed in vitro mismatch correction assay (Modrich 1986).

The following year, the MutL and MutH gene products were purified and found to be involved in excision events. Specifically, MutL was found to be a crucial mediator between recognition of a mismatch by MutS and activation of MutH, a Mg2+- dependent endonuclease that cleaves at the 5’-end of the unmethylated strand in hemi-methylated DNA duplex, leading to a nick in the daughter strand containing the mismatch (Modrich et al. 1987; Modrich et al. 1989). The coordinated events of these three Mut proteins, in concert with UvrD helicase, DNA polymerase III, exonucleases, and DNA ligase were found to be sufficient for effectively excising and repairing the daughter strand.
Prokaryotic Mismatch Repair

The Mut system in the gram-negative bacteria, *E. coli*, has been well characterized and serves as a fundamental model for mismatch repair in higher organisms including humans. Thus, a brief overview of the prokaryotic system will serve as a useful introduction to MMR and be a helpful model for understanding the more complex eukaryotic systems (Modrich 1996; Kolodner 1996). In a prominent model of mismatch repair, MutS, a homodimer ATPase, scans the newly replicated DNA, presumably making non-specific contacts with the DNA backbone until it encounters a base-base mismatch or small insertion/deletion loop (IDL). Upon encountering such a non-Watson-Crick base pair, MutS binds it with high affinity, kinking the duplex approximately 60° to form a stable MutS-mismatch complex. The ATPase activity of MutS plays an important but as yet unclear role in the process. MutL, also a homodimer ATPase, interacts with MutS and DNA to form a ternary complex that activates MutH, a monomeric endonuclease. MutH preferentially nicks the error-containing strand at a GATC site, which allows UvrD helicase to unwind the duplex starting from the nick towards the mismatch. Subsequently exonucleases are recruited to the site and excise the strand containing the error. The resulting single strand is stabilized by single stranded binding protein (SSB), and serves as a template for DNA resynthesis by polymerase III. Finally, DNA ligase completes the repair reaction by ligating the corrected daughter stand.

Elegant *in vitro* mismatch repair assay systems have been reconstituted over the past few decades and have provided insights into how these multiple proteins coordinate their activities in order to recognize and specifically repair the error-
containing strand. MutS, the mismatch recognition protein, is primarily responsible for signaling downstream proteins to initiate repair of a mismatched base (Grilley, Welsh et al. 1989). In gram-negative bacteria, mismatch repair co-opt DNA methylation to determine which is the daughter DNA strand that must be corrected. Duplex DNA is methylated by dam methyltransferase on the N6 position of adenine in palindromic d(GATC) sites, and this marker is utilized by the organism to discriminate between native DNA and any foreign, unmethylated DNA that enters the cell. During DNA replication, there is a lag time (on the order of minutes) in between DNA synthesis by the polymerase and dam methylation, resulting in a newly formed duplex that is transiently hemi-methylated at the d(GATC) sites. Activated MutH preferentially breaks the phosphodiester backbone of unmethylated strand or hemi-methylated deuplex, and it is thus responsible for strand discrimination ensuring the daughter strand is targeted for mismatch repair for this pathway. (Welsh, Lu et al. 1987; Au, Welsh et al. 1992).
Figure 1.1: Minimal model of prokaryotic mismatch repair

1. **MutS** scans the DNA for mismatches.
2. Recognition of mismatch and recruitment of MutL.
3. Ternary complex activates MutH which nicks the daughter strand.
4. DNA polymerase III resynthesizes and the strand is ligated.
5. UvrD Helicase unwinds strand and exonucleases excise erroneous base.

**MutS:** Dimeric ATPase that recognizes mismatches

**MutL:** Dimeric ATPase required to activate MutH

**MutH:** Endonuclease, nicks daughter strand
Figure 1.1: Minimal model of prokaryotic mismatch repair. The MutS homodimer scans the newly replicated DNA and binds tightly to a mismatch or insertion deletion loop (IDL). Once bound, the protein recruits MutL and the ternary complex activates MutH, an endonuclease that preferentially nicks the unmethylated daughter strand. This nicking activity activates UvrD helicase to unwind the duplex and allow for exonucleases to excise the erroneous base approximately 100 base pairs past the mismatch. DNA polymerase, tethered to β-clamp, resynthesizes the daughter strand which is finally sealed by DNA ligase.

**Eukaryotic Mismatch Repair**

Evidence for a similar mismatch repair pathway in eukaryotes was first detected in *in vitro* experiments involving transformation of mismatched viral and plasmid DNA into mammalian cells (Folger, Thomas et al. 1985). The frequency of repair was comparable to *E. coli* systems, suggesting a similar pathway was at work in these higher organisms (Brown and Jirincy 1988). In fact, these post-replicative mismatch repair process has been found to be highly evolutionarily conserved and several homologues of the prokaryotic MMR proteins have been discovered in organisms ranging from archaic thermophiles to humans. Table 2.2 lists the *E. coli* mismatch repair proteins and their eukaryotic homologues.
### Table 2.2

<table>
<thead>
<tr>
<th>E. coli Protein</th>
<th>Function in MMR</th>
<th>Homologues</th>
<th>Function in MMR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MutS</strong></td>
<td>Scans newly replicated DNA and binds tightly to a mismatch or insertion/deletion loop</td>
<td>MSH2-MSH6 (MutSα)</td>
<td>Binds to base-base mismatches and 1 - 2 IDLs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSH2-MSH3 (MutSβ)</td>
<td>Binds to single base IDLs and is predominantly involved in repair of IDLs &gt; 2 bases</td>
</tr>
<tr>
<td><strong>MutL</strong></td>
<td>&quot;Matchmaker&quot; that stabilizes the MutS-DNA complex and signals downstream repair events</td>
<td>MLH1-PMS2 (MutLα)</td>
<td>Coordinates recognition of mispairs by MutS to downstream repair and resynthesis of DNA and has latent endonucleolytic activity believed to nick the daughter strand</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLH1-MLH2 (hPMS1) (MutLβ)</td>
<td>Function unknown in humans, but has been shown to suppress IDL mutations in S. cerevisiae</td>
</tr>
<tr>
<td><strong>MutH</strong></td>
<td>Preferentially nicks unmethylated daughter strand at hemimethylated GATC sites</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td><strong>γ-δ Complex</strong></td>
<td>Loads β-clamp onto DNA</td>
<td>RFC Complex</td>
<td>Loads PCNA onto duplex and modulates base excision</td>
</tr>
<tr>
<td><strong>β-Clamp</strong></td>
<td>Enhances the processivity of DNA pol III and may recruit MutS to mismatches / replication fork</td>
<td>PCNA</td>
<td>Interacts with MutS and MutL homologues and recruits MMR proteins to a mismatch and participates in DNA repair synthesis</td>
</tr>
<tr>
<td><strong>Helicase II (UvrD)</strong></td>
<td>Loads onto DNA at the nicked site and unwinds DNA to allow excision of ssDNA</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td><strong>ExoI</strong></td>
<td>3’ to 5’ excision of ssDNA</td>
<td>EXO1 (Rth1)</td>
<td>Excision of ssDNA</td>
</tr>
</tbody>
</table>
**Eukaryotic Mismatch Repair**

Similar to the prokaryotic model, eukaryotic mismatch repair is initiated by the MutSα or MutSβ, both heterodimeric ATPase proteins made up of two MutS homologues, MSH2-MSH6 and MSH2-MSH3, respectively (Marti, Kunz et al. 2002). MutSα is primarily responsible for recognizing base-base mismatches and +1 insertion/deletion loops, while MutSβ has been suggested to primarily recognize insertion/deletion loops larger than 2 nucleotides (Marsischky, Filosi et al. 1996). Once bound to a mismatch, MutSα similarly kinks the duplex DNA and recruits MutLα, a heterodimeric ATPase made up of two MutL homologues, MLH1-PMS2 (Mlh1-Pms1 in yeast) to form a ternary complex (Hall and Kunkel 2001). Unlike the prokaryotic system, this ternary complex does not activate a MutH homologue, and in fact to date no eukaryotic MutH homologues have been found (Marti, Kunz et al. 2002). It has been shown that MutLα has latent endonuclease activity that can make nicks in DNA, and is thus thought to replace MutH in the repair reaction. Once MutLα forms a nick in a DNA strand, exonucleases excise the strand past the
mismatch. DNA polδ resynthesizes the strand, and ligase seals the nick to complete the reaction. (Figure 1.2; Reviewed in Kolodner and Marsischky 1999; Modrich 2006).

In the absence of a MutH homologue or a transiently hemi-methylated DNA duplex (for eukaryotes do not have the same d(GATC) methylation patterns as prokaryotes), researchers have had difficulty explaining how eukaryotic systems discriminate between the daughter and template strands. It has been previously shown that MutLα endonuclease activity is stimulated in the presence of a pre-nicked DNA substrate and prokaryotic systems with MutH knocked out can still perform successful MMR if the DNA substrate is pre-nicked. These observations have led to the hypothesis that newly replicated daughter strands, especially those in the lagging strand, have a higher instance of nicks which may serve as markers for eukaryotic MMR proteins to preferentially target the daughter strand (Guarné et al. 2010, Lahue, Su et al. 1987).
Figure 1.2: Minimal model of eukaryotic mismatch repair

1. **MutSα** scans the DNA for mismatches.
2. Recognition of mismatch and recruitment of **MutLα** which nicks daughter strand.
3. Ternary complex communicates with **PCNA** at nicked site.
4. **DNA polymerase δ** resynthesizes and the strand and is then ligated.
5. **Exonucleases** excise erroneous base.
**Figure 1.2: Minimal model for eukaryotic mismatch repair:** The MutSα heterodimer recognizes the mismatch and recruits MutLα which has latent endonuclease activity that cuts the daughter strand. The MutSα-MutLα-mismatch ternary complex signals exonucleases to excise the error and DNA polymerase δ resynthesizes the daughter strand, and finally the strand is sealed by DNA ligase. Note that in eukaryotes the DNA is not methylated at d(GATC) sites and the mode by which eukaryotic proteins discriminate between the daughter and template strand is poorly understood.

Recently, many of the MMR proteins have been crystalized both with and without the presence of DNA and have been crucial for researchers in gaining perspective on their structure-function relationship. A brief review of some of the salient structural features of the main players in MMR will assist in developing a fundamental understanding of how these proteins perform their crucial cellular function and highlight the questions these structures fail to elucidate. Additionally, these structural snap shots provide information with which rationally designed and targeted experiments can be implemented for future research projects.

**Structure and Function of MutS**

Crystal structures of the prokaryotic MutS protein alone and co-crystalized when bound to a mispaired DNA substrate have been crucial for gaining deeper insight into how this protein’s structure is related to its function. The *Thermus aquaticus* MutS crystal structure shows that this protein binds to a 21-base pair heteroduplex containing an unpaired thymidine (+T) nucleotide as a clamp-like
homodimer (Figure 1.3; Jirinec 2000; Obmolova, Ban et al. 2000; Yang, 2000). The protein has five structurally distinct domains that are structurally identical in each subunit. Domains I and IV (also known as the DNA binding and clamp domains, respectively) undergo the primary interactions with DNA, predominantly with the phosphate-sugar backbone. Domain IV is 40Å long and comprises mainly α-helices that encircle the DNA upon mismatch binding, while Domain I of subunit 1 donates mismatch-specific residues Phe39 and Glu41 that are highly conserved and crucial for bind the mismatch or insertion/deletion loop (IDL) with high affinity. The Phe39 forms aromatic stacking with the unpaired thymidine while the Glu41 donates a hydrogen bond to the nucleotide. It should be noted that while Subunit 2 also has these residues in the DNA binding domain, they do not make the same interactions with the mismatch, highlighting functional asymmetry between to the two monomers (Obmolova, Ban et al. 2000).

When stabilized on the mismatch, the protein domains and the DNA helix both undergo conformational changes that optimize the protein-DNA interactions. The MutS protein kinks the duplex helix approximately 60° towards the major groove to allow for a larger MutS-DNA interface, and in the process the protein forms a well-ordered clamp-like structure. It has been suggested that the mechanism through which MutS discriminates between matched and mismatched bases may be due to the lower energetic cost (~2 – 3 kcal/mol) associated with kinking an already slightly more flexible non-Watson-Crick base pair (Nag et al. 2007). Indeed, in the absence of DNA, Domain’s I and IV are highly disordered and mobile such that they cannot be defined in a 3Å resolution structure. It has been proposed that the flexibility of these
domains coupled with the asymmetric binding of the subunits accounts for the wide variety of substrates that this protein is capable of binding to and recognizing (Figure 1.3; Obmolova, Ban et al. 2000).
Figure 1.3: Crystal structure of *Thermus aquaticus* MutS

![Diagram of MutS crystal structure](image)

**Figure 1.3: Crystal structure of *Thermus aquaticus* MutS:** The crystal structure of *T. aquaticus* MutS unbound to a DNA substrate. Note the loss of ordered structures for Domain’s I and IV in the absence of a duplex. A front view of the MutS homodimer bound to a 21-bp heteroduplex DNA containing a +T insertion with subunit 1 labeled in blue and subunit 2 labeled in green. Note the kink of the DNA backbone. Domain IV encircles the duplex, while Domain I makes mismatch-specific contacts with the extrahelical nucleotide. Domain V is crucial for dimerization and contains the composite ATPase site. Subunit 1 of *T. aquaticus* MutS with the five structurally distinct domains.
The *T. aquaticus* MutS structure also shows that this protein is a member of
the ABC ATPase superfamily. The C-terminal Domain V region contains highly
conserved Walker A and Walker B motifs for ATP binding and hydrolysis,
respectively. Residues from both subunits contribute to these motifs, illustrating a
composite ATPase binding domain that is not only crucial for ATP hydrolysis, but
dimer formation as well. In fact, Domain V accounts for approximately 70% of the
dimer interphase through hydrophobic interactions between the anti-parallel α-helices
of each subunit, and without dimerization, the ATPase activity of the protein is
abolished. Previous reports have shown that ATP hydrolysis by MutS proteins is
severely lowered in the presence of a mismatch and concomitantly MutS mismatch
binding capability is weakened when the protein is bound to ATP. From these studies,
it has been proposed that there is a tightly coupled allosteric relationship between the
C-terminal composite ATPase domain and the DNA mismatch binding domain that is
crucial for the protein’s function.

The crystal structures of prokaryotic MutS bound to a mismatch have
supported this allosteric communication between the domains by showing that
mispair binding induces a much more rigid structure for Domains I and IV and induce
subtle alteration in the ATPase domain structure. There are highly conserved residues
at the junctions of domains II, III, and V that are shown to predispose patients to
cancer when mutated, suggesting that they may be important for transmitting
information between the two domains, which are separated by tens of angstroms.
These crystal structures, however, offer little insight into the conformational changes
needed to transduce a signal from the ATPase domain down to the N-terminal
binding domain, and vice versa, and in-solution experiments are necessary in order to characterize MutS dynamics during the repair reaction.

Recently, the eukaryotic human MutSα (MSH2-MSH6) has been co-crystalized with a heteroduplex DNA substrate containing a G:T, +T, or O6-methylguanine:thymidine mismatch (Beese et al. 2007). Similar to prokaryotic MutS, MutSα forms a heterodimeric oval clamp like structure around the DNA substrate and kinks it approximately 45° at the mismatch. While the eukaryotic homologue has the same five domains as the prokaryotic MutS, they differ slightly in length, sequence, and orientation. The mismatch binding domain (Domain I) has a slightly different protein-DNA interface. The MSH6 subunit, analogous to subunit 1 in prokaryotes, makes specific contacts with the mismatch and has an ordered N-terminal region with positively charged residues which assists in non-specific contacts with the DNA; however, the MSH2 subunit only makes one contact with the DNA backbone, in contrast to the subunit 2 in prokaryotes which makes extensive though non-specific contacts with DNA. MSH6 utilizes Phe432, Glu434, and Met454 to specifically interact with the mismatch. The clamp domain (Domain IV) similarly makes non-specific interactions with the duplex and is rearranged into an ordered clamp on binding to mismatched DNA.

The eukaryotic ABC-ATPase domain (Domain V) is the most highly conserved region of the protein, being 48% identical to the sequence of the *E. coli* domain V. Like its prokaryotic homologue, the ATPase domain consists of two composite Walker A and B motifs that are essential for ATP hydrolysis, and additionally contains a conserved helix-turn-helix motif that stabilizes the dimer
interface between the two subunits in the absence of nucleotides. A conserved loop in the MSH6 lever domain (Domain III) may be involved in an allosteric or signal transduction interaction between the ATPase and DNA binding domain similar to the prokaryotic MutS. The eukaryotic and prokaryotic crystal structures show a highly evolutionarily conserved ternary structure between the MutS homologues.

Structure and Function of MutL

The primary mechanism through which MutL performs its mediation between MutS recognition of a mismatch to signaling of the error’s excision has remained elusive in large part because of the lack of structural information regarding the entire protein complex (Ban and Yang 1998; Yang 2000; Guarne, Junop et al. 2001; Guarne, Ramon-Maiques et al. 2004; Kosinski, Steindorf et al. 2005). Fragments of the N- and C-terminal domains of E. coli MutL have been separately crystallized; however, due to highly disordered and dynamic linker arms between these domains, the entire complex has not been successfully resolved. A theoretical protein complex has been proposed which show a conserved N-terminal ATPase site and a positively charged groove inside the saddle-shaped dimer that can interact with the DNA backbone (Figure 1.4A; Guarne, Ramon-Maiques et al. 2004). It has been shown that the N-terminal fragment can bind to DNA; however, to a much lesser extent then with the C-terminal domain present.

The sequence homology of the C-terminal domain (CTD) is much less conserved across MutL homologues compared with the N-terminus, and this may be in part due to the different roles this domain plays in gram-negative prokaryotes
versus eukaryotic homologues. As previously mentioned, the eukaryotes do not contain a MutH homologue and the strand nicking is apparently performed by the latent endonuclease activity of MutL. The CTD of *B. subtilis* has recently been solved and has provided more insight into how MutL nicking activity is performed and regulated (Guarné et al. 2010). The CTDs domain, which is conserved in 3D organization to other MutL homologues, exists as a dimer with regulatory, dimerization, and lever subdomains. The lever subdomain contains the conserved endonuclease catalytic site which has weak Mg$^{2+}$-dependent activity that is required for mismatch repair (point mutations within the Mg$^{2+}$ binding domain results in loss of MMR *in vivo*). The metal cation is believed to play a crucial role in MutL activity by ordering the dimerization and regulatory domains, which are highly flexible and mobile without metal binding, and orienting the protein into the proper position for its catalytic activity on DNA.

Interestingly, the regulatory subdomain is coated with a group of negatively charged residues that appear to inhibit non-specific nicking activity by the protein. It has been proposed that a large conformational change must occur in order to prevent electrostatic repulsions from this domain and the DNA backbone which may require the binding of a cofactor such as ATP at the N-terminal or perhaps another repair protein. It has previously been shown that the human MutL$\alpha$ homologue interacts with MutS$\alpha$ through its ATPase domain, and the *B. subtilis* MutL CTD crystal structure reveals a clamp binding motif that may interact with $\beta$-clamp, a homologue of eukaryotic PCNA, and also regulate MutL endonuclease activity (Guarné et al. 2010). The limited structural information of MutL has hindered detailed
characterization of this protein’s function; however, current models suggest that it may encircle the duplex in an ATP bound form and nick the strand when activated by MutS. Further in vitro assays will need to be performed in order to fully articulate the coordinated activities of this protein with other mismatch repair proteins.
Figure 1.4: Crystal structure of MutL homologues

A) A proposed structure of the *E. coli* MutL. B) CTD dimer of *Bacillus subtilis* with domain 1 labeled red and domain 2 labeled orange. The domain 1 of the MutL CTD. The three main subdomains of this monomer are its dimerization (shown in red), regulatory (shown in yellow) and lever (shown in orange) subdomains. Notice the lever domain has the highly conserved endonuclease active site motif (Guarné et al. 2010). C) When MutL binds ATP it is believed to be in a closed or clamp-like structure.

Figure 2.4: Crystal Structure of MutL homologues: A) A proposed structure of the *E. coli* MutL. B) CTD dimer of *Bacillus subtilis* with domain 1 labeled red and domain 2 labeled orange. The domain 1 of the MutL CTD. The three main subdomains of this monomer are its dimerization (shown in red), regulatory (shown in yellow) and lever (shown in orange) subdomains. Notice the lever domain has the highly conserved endonuclease active site motif (Guarné et al. 2010). C) When MutL binds ATP it is believed to be in a closed or clamp-like structure.
Structure and Function of MutH

As previously mentioned, MutH is a Mg\(^{2+}\)-dependent endonuclease that preferentially cleaves single stranded DNA strands 5’ to a hemimethylated d(GATC) site in gram-negative bacteria upon activation by MutS and MutL. The single stranded nick formed by MutH allows exonucleases to excise the daughter strand containing the erroneous nucleotide and for DNA polymerase III to resynthesize the daughter strand, which can be ligated by DNA ligase. From the crystal structure, it appears that the protein is active as a monomer and forms a clamp-like structure with a cleft separating two subdomains, the N-arm and the C-arm (Figure 1.5; Ban and Yang 1998). The large cleft running down the middle of the protein contains the DNA binding active site, located towards the bottom of the groove and at the hydrophobic interface of the two subdomains where positively charged loops and aromatic residues interact with the DNA backbone and nucleotides, respectively. As expected, the sequence of the active site is related to Type II restriction enzymes; however, there are some overall structural differences between MutH and typical restriction enzymes that may help to explain how this protein performs its highly specific function.

Most notably, typical Type II restriction enzymes, such as PvuII and EcoRV (which have little sequence homology, yet are very similar in structure to MutH) exist as dimers in solution and cut both strands of a duplex. MutH exists as a monomer in solution, which may explain its preferential single stranded nicking activity.

Secondly, regulation of MutH endonuclease activity is believed to be governed by the relative pivoting motions between the two arm subdomains, which are capable of
exposing or blocking off the DNA binding site depending on their relative positions. This type of regulation is different from typical restriction enzymes whose activity is regulated via movement between dimers. Finally, a C-terminal hydrophobic tail located in the C-arm subdomain of MutH potentially interacts with MutS/MutL MMR proteins as a lever, which could modulate the pivoting of the two subdomains and ultimately the catalytic activity of this protein (Ban and Yang 1998).
Figure 1.5: Crystal structure of prokaryotic MutH

The structure of the gram-negative MutH protein shows a clamp-like structure with a DNA binding cleft separating the two halves. The protein is split into two subdomains: A N-arm subdomain (in dark blue) and a C-arm subdomain (in magenta) that pivot relative to each other to regulate catalytic activity and are connected by three linker regions shown in tan. The C-terminal lever arm (in yellow) is proposed to regulate the relative movement of the subdomains and may be the site of communication between MutH and MutS/MutL. Five active site residues in the DNA binding domain are shown: Glu56, Asp70, Glu77, Lys79, Lys116. Oxygen and nitrogen atoms on these residues are represented by blue and red spheres, respectively (From Ban and Yang 1998).

Figure 2.5: Crystal structure of prokaryotic MutH: The structure of the gram-negative MutH protein shows a clamp-like structure with a DNA binding cleft separating the two halves. The protein is split into two subdomains: A N-arm subdomain (in dark blue) and a C-arm subdomain (in magenta) that pivot relative to each other to regulate catalytic activity and are connected by three linker regions shown in tan. The C-terminal lever arm (in yellow) is proposed to regulate the relative movement of the subdomains and may be the site of communication between MutH and MutS/MutL. Five active site residues in the DNA binding domain are shown: Glu56, Asp70, Glu77, Lys79, Lys116. Oxygen and nitrogen atoms on these residues are represented by blue and red spheres, respectively (From Ban and Yang 1998).
Current Models for Mismatch Repair

While the many elegant genetic, in vivo, reconstituted mismatch repair experiments and protein crystal structures have elucidated many key elements of the MMR pathway, one fundamental mechanistic issue has eluded researchers: namely, how do mismatch repair proteins communicate information from a mismatch to a distant nicked site (which can be up to 1 kilobase away) to ultimately excise the error? Three current proposed models have evolved over the past twenty years in an attempt to reconcile this poorly understood phenomenon, and a comparative analysis on the strengths and shortcomings of each scheme will help in gaining deeper insight into what is currently held to be true regarding mismatch repair, and what questions still remain to be answered.

The Translocation Model

The translocation model proposes that MutS scans the newly replicated DNA in a nucleotide-free state until it rapidly binds and hydrolyzes ATP upon recognition of a mismatch. The protein’s ATPase activity induces a conformational change that weakens its affinity for the mismatch while also providing energy for the protein to physically move away from the error. Subsequent rounds of ATP hydrolysis provide additional energy for the protein to thread both flanking strands of DNA through its clamp domain, simultaneously sequestering the mismatch within a looped-structure and bringing the distant nicked site into close proximity to MutS. Proponents of this model suggest it is the ATP-hydrolysis dependent bidirectional movement of the MutS protein that ultimately brings the nick and mismatch site into close proximity,
thus effectively directing downstream players to selectively excise the erroneous base (Grilley, Griffith et al. 1993; Modrich et al. 1998).

The first proposal of a translocation model for mismatch repair arose when researchers observed MutS-DNA complex using electron microscopy and found that heteroduplex substrates containing a G:T mismatch formed $\alpha$-shaped loop structures with a MutS protein dimer at the base of the structure. (Allen, et al. 1997). Interestingly, the mispair was found to be located in the middle of the loop, or at a site distant to the MutS protein, suggesting that the protein had physically moved away from the mismatch after recognition. When MutS was mixed with a homoduplex substrate or in the absence of nucleotides, the number of looped-structures was greatly reduced suggesting that structure-formation was mismatch and ATP dependent. Moreover, when the reaction was assayed in the presence of ATP$_{\gamma}$S, the number of $\alpha$-loops observed was reduced nearly 90% and when AMP-PNP (a non-hydrolyzable ATP analogue) was added to the reaction containing ATP, loop growth was inhibited. These results, coupled with the observation that loop-growth had a near-linear dependence on ATP concentration, lead researchers to the conclusion that the movement of MutS along the duplex required ATP hydrolysis and ultimately led to the proposal of a translocation model.

**Molecular Switch Model**

The molecular switch model suggests that MutS scans newly replicated DNA in an ADP-bound form and, upon stable binding to a mismatch, exchanges ADP for ATP. The MutS-ATP complex subsequently leaves the mismatch in a hydrolysis-
independent fashion to recruit downstream repair proteins like MutL to DNA, and only after all repair machinery assemble in the vicinity of the mismatch, MutS hydrolyzes ATP to recycle back to its mismatch scanning ADP bound form. In this scheme, the authentic signal directing excision machinery to the erroneous base is the conformational change of MutS associated with nucleotide exchange, which allows the protein to leave the mismatch in a sliding clamp formation towards the nicked site. Supporters of this model have compared this nucleotide exchange to the well characterized G-proteins, involved in biological signal transduction, which are in an active state when bound to GTP and an inactive state when bound to GDP. In this model, the MutS protein is on, or activated to recognize a mismatch, when bound to ADP and is off, or has compromised mismatch binding capability, when bound to ATP. (Fishel, 1998)

**Trans/Stationary/DNA Bending Model:**

Both of the former models for MMR have suggested that MutS moves away from the mismatch upon ATP binding in order to signal downstream events; however, while it has been shown that MutS has a weakened affinity for a mismatch in the presence of ATP, it has not been directly shown to move away from the erroneous base upon recognition. Moreover the question remains as to why a protein whose function is to recognize a target site should move away from it before the repair reaction is underway. In the third ‘stationary’ or ‘DNA bending’ model, it has been proposed that MutS must simultaneously bind the mismatch and ATP in order to faithfully initiate the repair process through activation of MutH. Specifically, this
model proposes that MutS, with its weakened affinity for both homo- and heteroduplex DNA when bound to ATP, utilizes nucleotide binding to verify the presence of a mismatch. If the protein is in contact with a homoduplex, ATP binding will induce the protein to fall off the duplex; however, if MutS is anchored to a mismatch via its mismatch binding domain, ATP binding will weaken but, not stimulate rapid dissociation of the protein from the DNA. Thus, in this model the ATPase activity of MutS is coupled to mismatch validation and targeting to an authentic mismatch site. In the MutS-DNA-ATP complex is formed, bending of DNA by MutS is thought to facilitate collision between the distance mismatch and GATC sites, bringing them into closer proximity to initiate nicking and further mismatch repair.
Figure 1.6: Summary of mismatch repair models

A) Translocation Model

B) Sliding Clamp Model

C) Stationary “DNA Bending” Model
**Figure 1.6: Summary of mismatch repair models:** A) The Translocation Model proposes that MutS rapidly hydrolyzes ATP after recognizing a mismatch. The energy generated from its ATPase activity allows the protein to move as a motor away from the mismatch and subsequent rounds of hydrolysis drive the two flanking strands of DNA through the protein’s clamp domain to simultaneously sequester the erroneous nucleotide in an α-like loop and bring the nicked site in closer proximity to the mismatch and direct downstream players for its excision. B) The Molecular Switch Model proposes that the MutS scans for mismatches in an ‘inactive’ ADP-bound state and switches ADP for ATP upon binding to a mismatch. Binding to ATP weakens the protein’s affinity for the mismatch, allowing it to travel away from the site while additionally ‘activating’ the dimer to signal downstream players to excise the error. Movement away from the mismatch allows for additional MutS proteins to bind to the mismatched site, further activating more dimers and allowing for a robust signal for downstream MMR components. C) The Stationary/Trans/’DNA Bending Model’ proposes that MutS simultaneously binds ATP and the mismatch, and through a coordinated conformational change physically brings the nicked site towards the mismatch to allows for MutS to signal for downstream players without moving away from the mismatch.
Conflict in current MMR models

Despite extensive efforts to elucidate the mechanism through which MMR is initiated, the three current models conflict as to the mode through which MutS directs downstream repair processes to the mismatch. Specifically, the coordinated recognition and ATPase activities of the protein are poorly understood, highlighting the need for the development of new in-solution assays to directly monitor MutS action on mismatched DNA. While previous studies have observed MutS binding to a mismatch in real time through monitoring a signal from the DNA substrate, a detailed kinetic analysis from the perspective of the protein has been lacking, limiting our understanding of the mechanistic steps governing this binding event. We have developed a novel fluorescent probe directly on the MutS to monitor the protein’s interactions with mismatched DNA for the first time, in the hopes of offering new insights into the poorly understood initiation step of mismatch repair.

Experimental Design and Objectives

Our project focuses primarily on MutS interactions with a mismatched DNA substrate. The mechanism through which this protein can bind to a variety of substrates and the coupling of its mismatch recognition and ATPase activities is poorly understood. An extensive amount of research has been dedicated to understanding this crucial first step in the MMR pathway, for it is the point at which the repair machinery initiate the excision of a mismatch or IDL, additionally over 40% of patients with Lynch Syndrome have mutations in the MutSα dimer,
highlighting the immediate need for a detailed understanding of how this protein’s function is related to genomic stability and prevention of tumor growth.

While useful kinetic information regarding the MutS-mismatch binding event has been obtained utilizing fluorescent probes on a DNA substrate, the interaction has not been viewed in real time from the perspective of the protein. In order to observe the actions of MutS upon binding to a mismatch and during the ATPase reaction, a novel fluorescently labeled MutS protein has been developed. It is hypothesized that this MutS will serve as a useful tool to gain new kinetic and thermodynamic information about the reaction mechanism, and thus help resolve the wide differences between the current proposed models for mismatch repair.

Firstly, *Thermus aquatics*, an archaic thermophile, was chosen as the model system based on the fact that its MMR components are conserved in structure and function with higher eukaryotic systems, including humans. *T. aquaticus* MutS protein can be efficiently overexpressed and purified from *E. coli* cells and yield over 95% pure protein in milligram quantities. *T. aquaticus* MutS has been studied previously in the Hingorani laboratory under equilibrium and pre-steady-state conditions; thus a significant database is available on its working which makes it a good candidate for further in-solution analysis.

For this study, 2-(4’-(iodoacetoamido)anilino) naphthalene-6-sulfonic acid (IAANS), a thiol-reactive fluorophore, which was recently shown to be a useful probe for monitoring MutS interactions with a mismatch when conjugated to the T469 residue in the clamp domain, was chosen as the reporter (Ban et al. 2006). Wildtype *T. aquaticus* MutS protein has only one native cysteine residue per monomer at the 42
position in the DNA mismatch binding domain. This residue, however, is adjacent to the conserved Phe-X-Glu mismatch binding residues and would most likely compromise the binding activity of the protein if it was conjugated to a large fluorophore. Thus, the C42 residue was mutated to an alanine and, through careful structural and sequential alignments of MutS homologues by a F. Noah Biro, a graduate student in the Hingorani laboratory, the M88 position in the mismatch binding domain was ultimately chosen as the site for IAANS conjugation.

The M88 position was determined to be an ideal position for IAANS labeling given that it is positioned on the mismatch binding domain and can be expected to report any changes in MutS conformation during interactions with DNA coupled with the ATPase reaction (Figure 1.3). M88 is also a non-conserved residue across MutS homologues, suggesting that its substitution would not greatly compromise the protein’s function. MutS C42A/M88C double mutant was prepared and initially characterized under equilibrium conditions by F. Noah Biro (unpublished, Biro and Hingorani). In-solution assays can now be performed to monitor MutS actions under pre-steady-state conditions and derive kinetic information on the initiation of mismatch repair crystal structures and immobilized substrate experiments such as gel-shift assays and surface plasmon resonanse (SPR) experiments have not provided thus far.

Finally, given the growing field of fluorescence resonance energy transfer (FRET) and single molecule methods to monitor biochemical processes, including protein-protein and protein-DNA interactions, this research will ideally facilitate investigation of DNA mismatch repair under the single molecule regime.
2: Materials and Methods.

**Nucleotides, DNA, and other Reagents**

BL21 (DE3) *E. coli* cells used to overexpress the MutS (both wildtype and C42A/M88C mutant) protein were purchased from Stratagene. The Q-sepharose resin used to purify the MutS protein was purchased from Amersham Biosciences (GE Healthcare) and the P-6 media used for gel filtration was purchased from Bio-Rad. All nucleotides (ATP, ADP, ATPγS) were purchased from Sigma Chemical Co.

Synthetic single stranded DNAs were ordered from Integrated DNA Technologies, Inc. and purified by denaturing polyacrylamide gel electrophoresis (PAGE), electroelution, and ethanol precipitation. Single stranded DNA’s were annealed by mixing the ‘base’ strand and the ‘top’ single strands in a ratio of 1:1.15 to minimize single-stranded +T strand, and all mixing was done in annealing buffer (20mM Tris-HCl, pH 7.7, 100mM NaCl). The samples were placed into an H₂O bath at 95°C and then allowed to cool to 25°C for 8 – 10 hours. Double stranded DNA products were run on a non-denaturing PAGE gel and found to be >95% pure. Typically, 100μM double stranded DNA stocks were prepared and used in the fluorescent experiments including equilibrium and pre-steady-state binding assays.

**DNA Purification**

Note: I would like to acknowledge and thank Laura Pierce, a laboratory technician in the Hingorani group, for purifying all the single stranded oligos used in these experiments.
Vigorously vortex lyophilized DNA in 150mL of 1X formamide buffer (1X TBE, 50% formamide). Load DNA onto 40cm 16% acrylamide gel and run at room temperature for 4 – 6 hours at 400V. With a long-wavelength (366 nm) UV light, visualize DNA, cut out the band with a clean blade and slice into 2 x 15 mm pieces.

Load pieces into an electroelution compartment (Schleicher and Schuell) with a DNA collection chamber prepared between BT1 and BT2 membranes. Perform electroelution at 100V in 1X TBE buffer at 4°C, overnight. Briefly reverse the direction of the current (~15 sec) and immediately collect the DNA. Precipitate the DNA in a 1 mL high-adhesion eppendorf tube with 3X volume of 90% ethanol and 0.1X sodium acetate, pH 5.2, at -80°C for >6 hours. Centrifuge at 14,000 rpm for 30 min at 4°C, discard supernatant, air dry the pellet for 30 minutes, and resuspend the pellet in 20mM Trish-HCl, pH 7.5. Measure the concentration by UV absorption spectroscopy ($\lambda_{EX}=260$ nm)

**Preperation of E. coli Biomass**

**Transformation of E. coli cells**

Gently mix 0.5μL of 150ng/mL C42A-M88C or C42A-T469C taqMutS clone 1 plasmid into 100μL of BL21 (DE3) competent cells. Incubate solution on ice for 10 minutes then heat shock at 42°C for 50 seconds and immediately incubate on ice for 2 minutes. Add 300μL of sterile LB under sterile conditions and shake at 200 rpm at 37°C for 15 minutes. Plate the cells onto LB/Ampicillin resistant plates and incubate the plates overnight at 37°C.
Seeding

Autoclave six 2L flasks containing 40g of LB nutrient in 2L of H₂O, allow the flasks to cool and add 2 mL of 1000xAmp. Pick a single colony from the transformation plate and resuspend colony in 1mL of sterile LB solution. Add 30μL of this solution to each flask and shake flasks at 175 rpm at 37°C overnight.

Induction

Induce flasks at A₆₀₀ = 0.6 with 1mL 1M IPTG (final concentration 1 mM) and shake the flasks at 200rpm at 37°C. Samples were obtained before and after induction, run on a 10% SDS PAGE gel and stained with Coomassie blue to determine the success of induction.

Centrifugation

Centrifuge biomass at 5000 rpm at 4°C. Pool and resuspend pellets in 10% w/v sucrose and 25 mM Tris-HCl pH 8, for a final volume of ~90 mL. Store biomass in -80°C until purification.

Purification of MutSₘ₈₈₉
**MutS Purification Buffer Preparation [Buffer C]**

Buffer 1xC

Composition: 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 5% glycerol

---

**Column Packing: Q-Sepharose**

Wash Q-Sepharose resin (a white slurry in ethanol) with ~100mL of ddH2O and decant the supernatant. Repeat this washing step three times to ensure that all of the ethanol is washed away. Pack 14 mL of the washed resin in a column at room temperature, leaving a few milliliters of head volume of ddH2O to ensure that the column resin does not dry out. Ensure that the column is not leaking when finished and place column in the cold room. Prime a peristaltic pump by purging air through the tubing (at this point, not connected to the column) and then run ddH2O through the tubing. Connect the tubing to the top of the column and run ~10 bed volumes (140mL) of ddH2O through the column at 2.00 mL/min. Equilibrate the column with ~10 bed volumes (140mL) of 1xC50 buffer (Buffer C + 50mM NaCl) at 2.00 mL/min.

Note: Add about 100μL of PMSF (prepared by adding a few flakes to 100μL of isopropanol) to the C50 buffer immediately before equilibrating.

---

**Purification: Cell Lysis**

Before beginning, turn on water bath to 37°C. Thaw 6 L of biomass at room temperature and add, 1mg/mL of lysozyme (45mg in 300μL of ddH2O), 50mM NaCl (563 μL of 4M NaCl), 1.4mM β-mercaptoethanol (52.5 μL of pure B-mercaptoethanol), a few flakes of PMSF in 300μL of isopropanol, and fill to a final volume of ~ 45mL with 1xC0 buffer. Stir solution gently in cold room in a plastic
beaker on a stir plate for 30 min. Note: From this point onwards, ensure that the protein solution is kept at 4°C. Freeze-thaw the solution 3 times in a 50 mL tube to lyse the cells: 2 min in liquid N₂, 5 - 8 min in 37°C bath (until the solution is partially thawed, but not completely liquid) and finally at room temperature until the solution is completely liquid (~10 - 15 min). After freeze-thaw is complete, turn up the water bath to 70°C. Dounce the lysate vigorously with a loose glass piston in ice bucket for 15 – 20 minutes, being sure to minimize bubble formation. Note: The solution should be a light yellow color and with a significantly lower viscosity by the end of douncing. Centrifuge solution at 15,000 rpm at 4°C for 1.25 hours and remove supernatant, saving the solution in a 50 mL falcon tube. Heat the tube at 70°C water bath for 30 min and centrifuge at 17,000 at 4°C for 45 min.

**Protein Purification: Q-Sepharose Column**

Load the clear, yellow supernatant on a pre-packed Q-sepharose column in the cold room at 0.5mL/min and collect the flow through. Note: There is no wash step after loading. Elute the protein solution with a 140mL gradient of Buffer 1x C₅₀ – C₄₀₀. Prepare the gradient by adding ~2mL of C₅₀ to the lower gradient chamber, and open the valve connecting the lower and upper gradient chambers to prime the channel connecting the two chambers. Close the valve. Transfer any C₅₀ buffer that entered the upper gradient back to the lower gradient chamber. Fill the lower gradient chamber with 70mL of C₅₀ and the upper gradient chamber with 70 mL of C₄₀₀ (Buffer C + 400mM of NaCl). Connect the lower gradient chamber to the pump. Connect the bottom of the column to a fraction collector (in the cold room). Add
100uL of PMSF and a small stir bar to the lower gradient chamber. Simultaneously open the gradient chamber and collect 80 fractions of 1.75mL (~65 drops per tube, but can vary) running at 1.5mL – 2mL / min. Be sure to check on the column often to ensure that there is no leakage. Analyze the even fractions, load, flow through, and pellet on a 10% SDS PAGE gel.

**Protein Purification: Ammonium Sulfate Precipitation**

Pool the peak fractions into a 50mL glass beaker and add 0.2423 g/mL of ammonium sulfate to the solution. Add the ammonium sulfate in faction while gently stirring for 25 minutes at room temperature and allow solution to sit for 30 minutes at room temperature. Centrifuge at 5000 rpm for 10 minutes and save the supernatant. Resuspend the white pellet in 3 mL of Buffer C and analyze the pool, supernatant, and resuspended pellet on a 10% SDS PAGE gel. Dialyze the resuspended pellet in 3L of Buffer C + 50mM NaCl for 3 hours in the cold room, determine concentration by the Bradford method, aliquot and store final protein solution at -80 °C.

**Of special note:** Freeze thawed six times for MutS<sub>M88C</sub> purification instead of three times. Collected 80 fractions running at 63 drops / tube and took 10μL samples from even fractions to run on a 10% SDS PAGE gel. Concentration of protein was found by Bradford Assay at λ595, and then the solution was stored in aliquots at -80C.

**Labeling of MutS<sub>C42A/M88C</sub> with IAANS**

**P-6 Column Gel Filtration**
To prepare the P-6 gel filtration resin, weigh out 5g of dry Bio-gel P-6 media and hydrate with 60mL of degassed Buffer L (50mM Tric-HCl pH 7.5, 50mM NaCl, 5mM EDTA, 5% w/v glycerol) for 4 – 6 hours. After hydration, decant half of the supernatant transfer the gel solution to a sidearm flask and degas for 10 minutes with gentle stirring every two minutes. After degassing, add 60mL of Buffer L to the solution and swirl gently. Allow the suspended resin to settle and decant the supernatant and hydrate two more times with Buffer L. After the last wash step, rinse out a 1.5 x 20cm Econo-Column (Bio-Rad) and fasten the column with a Teflon valve at the column outlet and a funnel to the column reservoir. Fill column with Buffer L until it is one-third full. Open the column outlet and slowly pour the slurry in evenly to minimize discrete layer formation. Store the column in the cold room, connect to a pump with primed tubing and finally equilibrate with ten bed volumes of Buffer L running at 0.35mL/min.

Conjugation of Thermus aquaticus MutS\textsubscript{M88C} with IAANS

Dilute the MutS\textsubscript{M88C} protein to a 10\mu M stock solution with Buffer L and reduce the solution with a 10-fold molar excess of 20 mM TCEP (prepared by weighing out 1 – 2 mg of TCEP and made to the final concentration with dH\textsubscript{2}O). Attach tubing to a roller drum at room temperature and spin for 10 minutes. Add a 10-fold molar excess of IAANS (prepared by weighing out 1 – 2 mg of IAANS and made to the final concentration of 10 mM with DMF) to the protein solution by adding one-fourth of the IAANS solution every 15 minutes at room temperature and spinning the solution on a roller drum. Once the last fraction of IAANS is added,
move the roller drum into the cold room (45 minutes after the initiation of the
couplage reaction) and allow the reaction to continue for 10 – 12 hours. Terminate
the coupling reaction by adding 5mM DTT and spin samples on the roller
drum for an additional 10 minutes in the cold room.

Removal of Excess IAANS Dye: Concentration

To remove the excess dye, concentrate the sample by filling a 30 molecular
weight cut off (MWCO) concentrator with 3mL of the labeled protein solution and
centrifuge for two hours at 6,000 rpm at 4°C using a fixed-angle rotor. Repeat the
process if necessary. It typically takes a total of 5 hours to reduce the volume to
500μL. Transfer the concentrate to a microcentrifuge tube and rinse the sample
reservoir with 100mL of Buffer L and combine with the concentrate.

Removal of Excess Dye: Gel Filtration

Connect the equilibrated P-6 gel filtration column to a fraction collector. Open
the column outlet to allow the Buffer L head volume to fall to just about the level of
the P-6 resin and then close the column outlet to prevent the resin from drying out.
Remove the column top and carefully layer the concentrated labeled protein solution
onto the upper bed surface. Open the column outlet and allow the sample to drain into
the bed. Immediately wash the upper bed surface with 2 mL of Buffer L to wash the
sample into the bed (do not disturb bed surface). Reattach the column top when the
Buffer L reaches a level a few millimeters above the upper bed surface. Wash the
sample through the column at 0.35mL/min and collect 80 fractions of 250μL (about
10 drops). Analyze the fractions using a fluorescence microplate reader (SpectraMax M5), exciting the samples at 326nm (IAANS) and determining the peak labeled fractions from the emission readings at 445nm.
Figure 2.1: Labeling of MutS with IAANS at the M88C Position
Figure 2.1: Labeling of MutS with IAANS at the M88C Position: A) The absorbance from the MutS\textsubscript{M88C-IAANS} fractions after gel filtration through a P-6 column. From this graph, fractions 32 – 52 were determined to be the peak fractions to visualize on a SDS-PAGE gel. B): Overexpression, purification and labeling of MutS\textsubscript{M88C-IAANS}. SDS-PAGE analysis of purified MutS\textsubscript{M88C-IAANS} proteins (the molecular mass of the MutS monomer is 89.3 kDa, shown in the Coomassie gel; the molecular mass of the MutS\textsubscript{M88C-IAANS} dimer is 178.6 kDa). The bottom gel shows the same SDS-PAGE gel visualized under a UV illuminator, excited at 302nm and shown to have successfully labeled the MutS monomer with few impurities. C): Absorption scans with readings taken every 0.5nm wavelengths of the Protein in Buffer (orange trace), Buffer alone (red trace) and the calculated scan of only the labeled protein found by subtracting the values of the red trace from the orange trace (blue trace). From these traces, the maximum absorbance at 326nm and 280nm was obtained to calculate the degree of labeling. D) A crystal structure of \textit{T. aquaticus} MutS with the M88C position labeled with magenta spheres to represent the location of the IAANS probe.
Analyze the peak samples on a 10% SDS-PAGE gel along with the unlabeled, labeled, concentrate, and filtrate samples and first visualized under a UV transilluminator to qualitatively assess the labeling reaction at 302nm and 306nm and then stain the gels with Coomassie blue. Determine the concentration by the Bradford method and determine the degree of labeling by the method described below. Aliquot and store the final sample at -80°C.

**Calculating the Degree of MutS Labeling with IAANS**

Determine the degree of labeling by performing an absorption scan from 250 – 550nm with the labeled protein sample, correcting for Buffer. All absorption scans were performed with 1mL samples in quartz cuvettes on a UV-Visible Spectrophotometer (Shimadzu) and UV Probe software. Turn on the instrument and set the machine to PC Control. Place a 1mL ddH₂O sample into the reference chamber and a 1mL ddH₂O cuvette into the sample chamber and run a baseline correction scan to bring all the points in the absorption scan to a value of zero. Place a 1mL sample of Buffer L into the reference chamber and run an absorption scan taking readings every 0.5nm wavelengths, then run a final scan with a 1mL sample of Buffer L and the amount of labeled MutS_{M88C-IAANS} to yield a maximum absorbance of 0.025, calculated from Beers Law:

\[
c = A/[l2(\varepsilon_{MutS})]
\]

where A is absorbance (0.025), l is the width of the cuvette (1cm), \(\varepsilon\) is the extinction coefficient of the MutS monomer (52,720 M\(^{-1}\)cm\(^{-1}\)), and c is the calculated concentration of MutS_{M88C-IAANS} needed to be added. Note: Since MutS is known to
form a dimer in solution, the extinction coefficient is multiplied by two. From these calculations, the concentration of MutS dimer needed to elicit an absorbance peak of 0.025 is 0.24μM. Dilute the stock labeled MutS<sub>M88C-IAANS</sub> protein (with a known concentration found by the Bradford method) to 0.24μM of Buffer L and bring to a final volume of 1mL. Run the buffer scan and subtract the Buffer L trace from the Buffer L + MutS<sub>M88C-IAANS</sub> trace to get an absorption scan corresponding to the MutS<sub>M88C-IAANS</sub> dimer. Note the maximum absorbance at 326nm (the excitation wavelength of IAANS) and at 280nm (the excitation wavelength for aromatic residues on a protein) and use these values and the equation below to calculate the degree of labeling:

$$\frac{Abs\ 326/\epsilon_{IAANS}}{[Abs 280 - (Abs\ 326 \times C. F.)/(2 \times \epsilon_{MutS})]} = \frac{[IAANS]}{[MutS]}$$

where $\epsilon_{IAANS}$ is the extinction coefficient of the dye (27,000 M<sup>-1</sup>cm<sup>-1</sup>), $\epsilon_{MutS}$ is the extinction coefficient of the monomeric protein (52,720 M<sup>-1</sup>cm<sup>-1</sup>) and C.F. is the correction factor for MutS<sub>M88C-IAANS</sub> (0.72), which corrects for any promiscuous excitation of the IAANS fluorophore at 280nm. The equation yields the ratio of IAANS per dimeric MutS protein, using Beer’s law.

**Fluorescence-based Assays of MutS-DNA Interaction**

**Fluorescence Anisotropy of DNA**

Fluorescence anisotropy experiments using a TAMRA-labeled 37-mer DNA substrate (either with a +T insertion and GC homoduplex) and either unlabeled wildtype or labeled MutS<sub>M88C-IAANS</sub> protein were performed on a FluoroMax-3 fluorometer (Jobin-Y von Horiba Group; Edison, NJ) to measure MutS-DNA interactions.
interactions. A solution of 0.01 μM 37-mer +T TAMRA labeled DNA in Buffer L was prepared in a small quartz cuvette and selectively excited with polarized light at 555 nm (the excitation wavelength of the TAMRA fluorophore) through an excitation slit width of 2.5 nm and the anisotropy (\(<r>\)), or rather the degree of polarized light emitted at 585 nm, was recorded through an emission slit width of 3.5 nm. Anisotropy measures the degree of polarized light by the equation:

\[ r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp} \]

Where I terms indicate intensity measurements parallel and perpendicular to the incident polarization. The larger the complex in solution, the slower it tumbles, and thus the more polarized light it emits. All solutions were initially incubated for 30 seconds at room temperature, and had an integration time of 15 seconds with a 0.01% standard error at 25°C, and three trials for each point. The experiment involved taking the initial anisotropy reading for the DNA TAMRA alone in solution, and subsequently taking readings with increasing amounts of the wildtype or MutS\textsubscript{M88C-IAANS}.

Ten data points were collected with concentrations of MutS ranging from 0 – 0.8 μM and the anisotropy versus MutS concentration were plotted and the apparent dissociation constant (\(K_{D, APP}\)) for MutS was determined by fitting the data to a quadratic equation describing a 1:1 ligand-macromolecule binding event. (Note: all MutS concentrations are expressed for the dimer, which has a molecular weight of 178.6 kDa).

\[ x = \frac{[ML]}{[M]_T} = \frac{([L]_T - \sqrt{([L]_T + [M]_T + K_D)^2 - 4[M]_T[L]_T})}{2[M]_T} \]
Where $[M]_T$ is the total concentration of the macromolecule, which is the DNA substrate in the anisotropy experiments, $[L]_T$ is the total concentration of the ligand, in this case the MutS protein, and $K_D$ is the dissociation constant. The data were fit by non-linear regression using KaleidaGraph, Version 3.52 (Synergy Software).

**Fluorescence Intensity of MutS$_{M88C}$-IAANS**

Interaction between the MutS$_{M88C}$-IAANS and an unlabeled 37-mer DNA substrate (both with a +T insertion and a GC homoduplex) was measured on a FluoroMax-3 fluorometer (Jobin-Y von Horiba Group; Edison, NJ). A solution of 0.1$\mu$M of MutS$_{M88C}$-IAANS protein in Buffer L was prepared in a large quartz cuvette to a final volume of 2.6mL and selectively excited at 326nm through an excitation slit width of 2nm and the Signal/Reference (S/R) fluorescence output was recorded at an emission wavelength of 355nm through a 3nm slit width. All solutions were initially incubated for 2 minutes at 25°C while stirring, and then 30 seconds at 25°C without stirring and had an integration time of 1 second with a 0.1% standard error at 25°C with one trial for each point. For the experimental cuvette initial (S/R) fluorescence reading for the MutS$_{M88C}$-IAANS alone in solution, and subsequent readings with increasing amounts of the unlabeled 37-mer DNA (with a +T or GC homoduplex) were taken. A second cuvette was used to correct for photobleaching by taking the same initial reading of the MutS$_{M88C}$-IAANS alone in Buffer L, and subsequently taking readings by adding Buffer L in the same increments as the amount of DNA added in the experimental cuvette.
Eight data points were collected with concentrations of DNA ranging from 0 – 0.2μM and the IAANS fluorescence reading versus DNA concentration were plotted for both the experimental and control cuvettes. The experimental data was corrected for photobleaching by the following equation:

\[ P_{i,corr} = \frac{B_0}{B_i} \cdot P_i \]

where \( B_0 \) is the initial control reading, \( B_i \) is the control reading at a given titration point, and \( P_i \) is the corresponding raw titration point. This equation corrects for the decreasing in fluorescence signal contributed to photobleaching of the IAANS fluorophore. The apparent dissociation constant (\( K_D \)) for MutS was determined by fitting the corrected data to a quadratic equation describing a 1:1 ligand-macromolecule binding event. See above equation in Fluorescence anisotropy methods. It should be noted in these experiments the \([M]_T\) is the concentration of the MutS\textsubscript{M88C-IAANS} and the \([L]_T\) is the concentration of the unlabeled DNA substrate).

The data were fit by non-linear regression using KaleidaGraph, Version 3.52 (Synergy Software). Experiments were run with both +T DNA and G:C DNA (a matched substrate in the position of the mismatch of the +T substrate). The same experimental assay was performed with an unlabeled 80mer+T DNA substrate.

### Kinetic Analysis of MutS-DNA Interactions

To monitor the MutS-DNA binding reaction in real time, stopped flow experiments were performed on the KinTek SF-2001 stopped-flow apparatus (KinTek Corp.; Austin, TX) which has a dead-time of 1 msec. All experiments were performed in the dark (preferably at night) at 25°C in Buffer L (20 mM Tris-HCl, pH 7.7, 50 mM...
NaCl, 1 mM MgCl₂, 5% glycerol). In single-mixing experiments, 60μL of 0.2μM MutS₉₈₈C-IANS was mixed with an equal volume of either Buffer L alone or 0.2 – 20μM 37-mer+T DNA substrate, resulting in a 2-fold dilution of the reactants. A xenon lamp was used to selectively excite the IAANS fluorophore at 326nm (entrance and exit slit width of 10nm). The change in MutS₉₈₈C-IANS fluorescence signal was monitored in the observation cell from 0 to 2 – 4 seconds, with 1000 data points collected per trace. Emission over 350 nm wavelength was collected by a photomultiplier tube (PMT) set to 600V which had a long-pass cut-off filter (LG-350-F, Corion; Franklin, MA).

At least six independent traces were obtained for each concentration of DNA and at least three traces per experiment were averaged. Averaged traces for the mixing between MutS₉₈₈C-IANS with buffer alone were obtained at the beginning and end of the experiments to correct for photobleaching of the IAANS. To correct for the decrease in signal due to photobleaching, each raw averaged DNA trace was corrected against the buffer control traces using the same equation as for from the equilibrium binding data. The corrected traces were then fit to either a first or second-order-exponential-function from which information regarding the MutS₉₈₈C-IANS apparent on rate (kₐobs) could be calculated. When fit to a first exponential below:

\[ A_t = A_0(e^{-kt}) + F_0 \]

Aₜ, which corresponds to the fluorescent intensity at time point t, A₀ is the total change in the decrease in fluorescence of amplitude, and F₀ is the baseline are all given, and thus k, or kₐobs can be calculated. The traces were initially fit to a single exponential; however, after assessing the residuals, we found a significant amount of
data at the beginning of certain traces had poor R values, which suggested that a sum of two single exponentials may fit the data more effectively, as shown below.

\[ A_t = A_{0\text{Phase}1}(\cdot e^{-k(\text{Phase}1)t}) + A_{0\text{Phase}2}(\cdot e^{-k(\text{Phase}2)t}) + F_0 \]

Traces fit to this second exponential provided \( k_{\text{obs}} \) for two distinct phases occurring upon MutS\(_{\text{M88C-IAANS}}\) and mismatch binding.

To determine the concentration dependence of the MutS-DNA binding reaction, pseudo-first order reaction conditions were established with 0.2\( \mu \)M MutS\(_{\text{M88C-IAANS}}\) (0.1\( \mu \)M final) and 2 – 20\( \mu \)M unlabeled 37-mer+T DNA (1 – 10\( \mu \)M final concentration, and in great excess of the protein). The \( k_{\text{obs}} \) rate versus DNA concentration yielded a linear relationship with a slope that is the bimolecular rate constant, based on the equation below for a typical pseudo-first order rate equation:

\[ k_{\text{obs}} = k_{\text{on}}[\text{DNA}] + k_{\text{off}} \]

In experiments with the 37-mer DNA, the linear phase reached saturation, suggesting a rate-limiting second step in the reaction that is independent of DNA concentration. Analysis of these kinetic data are described in the Results section.

**Kinetic Analysis of MutS-DNA Interactions in the Presence of Nucleotides**

To monitor the MutS interactions with nucleotides in the presence of absence of mismatched DNA in real time, stopped flow experiments were performed on the KinTek SF-2001 stopped-flow apparatus (KinTek Corp.; Austin, TX). All experiments were performed with minimal light, at 40\(^\circ\)C in Buffer L. The single-mixing experiment involved mixing 60\( \mu \)L of 0.2\( \mu \)M MutS\(_{\text{M88C-IAANS}}\) (0.1\( \mu \)M final) with an equal volume of 200\( \mu \)M ATP, ADP, or ATP\( \gamma \)S (100\( \mu \)M final), and
monitoring IAANS fluorescence as described above. The data were corrected for photobleaching as described and were fit to single or double exponential functions for initial estimates of rate constants.

\[
A_t = A_0(\text{Phase1})(1 - e^{-k(\text{Phase1})t}) + A_0(\text{Phase2})(-e^{-k(\text{Phase2})t}) + F_0
\]

The same experiment was repeated with the presence of a 37-mer+T DNA substrate by initially incubating MutS\textsubscript{M88C-IAANS} with 4μM of the mismatch (2μM final) for two minutes before mixing with the nucleotides.
3: Results.

*MutS_{M88C-IANS} fluorescence reports MutS-mismatch interaction*

Previous studies in the Hingorani laboratory have shown that wildtype *Thermus aquaticus* MutS preferentially binds to heteroduplex DNA containing a +T insertion loop with high affinity ($K_D = 8 – 12$ nM) and homoduplex DNA with substantially lower affinity. When designing a novel MutS reporter protein, it was crucial to maintain this preferential binding capability. The M88 residue was chosen as an attachment site for a fluorophore because 1) it is located in the DNA binding domain, which can be expected to go through conformational change upon mismatch binding and 2) it is a non-conserved residue throughout the MutS homologues and changing it probably would not drastically compromise its wildtype function. IAANS, a thiol reactive fluorophore, was chosen as a reporter as it is known to exhibit fluorescence self-quenching (Ban et al. 2006). In a recent study with MutS labeled at the T469 position (4A apart in the MutS dimer), the authors observed a decrease in IAANS fluorescence in the presence of mismatched DNA. They speculated that high fluorescence in the absence of DNA indicates that the MutS DNA binding domains in the dimer are dynamic, whereas low fluorescence in the presence of mismatched DNA indicates stabilization of the mismatch binding domains in close proximity to each other (Ban et al. 2006). As seen in the crystal structure of *T. aquaticus* bound to a +T insertion, the epsilon carbon of the native methionine residues are approximately 12.6Å apart (Figure 3.1). It is not clear whether IAANS self-quenching is significant at this distance; however, the probe can be expected to report changes in the environment and the longer distance between the probes makes it less likely that they
may disrupt the MutS dimer interface and function. In the experiments described in this study, MutS\textsubscript{M88C-IAANS} is labeled at an IAANS:MutS dimer ratio of 1.7:1.

**Figure 3.1: The M88 residues are in proximity when MutS is bound to a mismatch**

Both fluorescence anisotropy and fluorescence intensity assays using DNA labeled with a 2-aminopurine (2-AP) fluorophore have previously been used in the Hingorani laboratory to monitor MutS binding to a mismatched DNA. These fluorescent assays have been shown to report MutS-DNA interactions effectively,
both under equilibrium and pre-steady-state conditions (Figure 3.2 A). Typically, these experiments have yielded an equilibrium dissociation constant for the MutS-mismatched DNA of 8 – 12 nM. In order to ensure that the MutS C42A/M88C double mutant labeled with IAANS still maintained wildtype function *in vitro*, an equilibrium fluorescence anisotropy DNA binding assay was performed initially with a 37-mer+T DNA_TAMRA substrate (Figure 3.2 B). When 50 nM of 37-mer+T DNA_TAMRA was titrated in with increasing amounts of MutS_{M88C-IAANS}, an increase in TAMRA fluorescence anisotropy was observed. This increase in anisotropy indicates formation of a large MutS-DNA complex, which tumbles relatively slower than free DNA, resulting in more polarized light emitted by the sample. Ten data points were obtained between 0 – 800 nM MutS_{M88C-IAANS} concentration and were fit to a quadratic equation for a 1:1 ligand-macromolecule reaction. A dissociation constant (K_D) of 29 nM was calculated and while this value is slightly higher than the dissociation constant measured from the 2-AP assay (Figure 3.2 A), MutS_{M88C-IAANS} still appears to bind +T DNA with high affinity.

Once we established that MutS_{M88C-IAANS} was able to bind tightly to mismatched DNA, complementary experiments were performed to monitor the same interaction as reported by the IAANS fluorophore on MutS. 0.1µM of MutS_{M88C-IAANS} was selectively excited at 326nm and titrated with increasing concentrations of unlabeled 37-mer +T/G:C DNA substrate (Figure 3.2 C). Upon addition of mismatched DNA, the fluorescence signal from the IAANS decreased significantly, suggesting that the dye molecules on the two mismatch binding domains were closer to each other in the MutS-DNA complex. It should be noted that IAANS is prone to
photobleaching, or photochemical destruction of the fluorophore after repeated exposure to light. In order to ensure that the decrease in signal was a result of the change in local environment of the probes and not a result of photobleaching, a control experiment with MutS\textsubscript{M88C-IAANS} and buffer (without DNA) was performed in parallel to correct for the decrease in signal due to photobleaching. Eight data points were obtained up to 200 nM of unlabeled 37-mer+T DNA and were fit to a quadratic equation, and a $K_D$ of approximately 8.8 nM was determined. This value is slightly lower than the dissociation constant obtained from the anisotropy experiments; however, is in closer agreement with wildtype MutS binding to a mismatch, based on the 2-AP experiments (Figure 3.2 A). MutS\textsubscript{M88C-IAANS} DNA binding experiments have been performed at least twice, but additional trails are needed to obtain reliable error estimates. Furthermore, the experiments need to be repeated at concentrations closer to the estimated $K_D$ value for MutS-DNA binding. However, these initial complementary measurements do provide some confidence that the mismatch binding ability of MutS\textsubscript{M88C-IAANS} is not compromised.

When the same titration experiment was performed with a 37-mer G:C DNA substrate (with a matched G:C base pair instead of the +T insertion) the fluorescence signal from the IAANS decreased slightly, but to a much lesser extent (about 40\% less than the +T curve, Figure 3.2 C blue trace). This slight decrease, indicating movement of the DNA binding domains, is expected since the protein is believed to make non-specific contacts with the DNA backbone as it searches for mismatches (Greene et al. 2010). The scanning of MutS along homoduplex DNA would require the protein to have some affinity for matched base pairs, which may be detectable in
this assay. Taken together, these results indicate that MutS_{M88C-IANS} can bind tightly to a mismatch, discriminate between a heteroduplex from a homoduplex substrate, and can be utilized as a reporter for obtaining quantitative in-solution information regarding MutS interactions with DNA.
Figure 3.2: MutS binds with high affinity to DNA containing a +T insertion as reported by different fluorescence-based assays

A) $\lambda_{EX,2A\text{p}} = 315$ nm
$\lambda_{EM,2A\text{p}} = 375$ nm
$[37\text{-mer+T DNA}_{2A\text{p}}] = 25$ nM

B) $\lambda_{EX,TAMRA} = 555$ nm
$\lambda_{EM,TAMRA} = 582$ nm
$[37\text{-mer+T DNA}_{TAMRA}] = 50$ nM

C) $\lambda_{EX,IA\text{ANS}} = 326$ nm
$\lambda_{EM,IA\text{ANS}} = 445$ nm
$[\text{MutS}_{M88C-IA\text{ANS}}] = 100$ nM
Figure 3.2: MutS binds with high affinity to DNA containing a +T insertion as reported by different fluorescence-based assays: A) Fluorescence titration curve monitored by fluorescence of 2-AP positioned immediately adjacent to a +T mismatch. Increasing amounts of MutS results in an increase in fluorescence signal when mixed with 25 nM +T DNA (likely from 2-AP unstacking and exposure to solvent). The data fit to a quadratic equation describing a 1:1 ligand-macromolecule binding event, yield a $K_D$ of 9.6 nM. B) Fluorescence anisotropy titration curve monitored by fluorescence of TAMRA end-labeled DNA (50 nM 37-mer+T DNA-TAMRA). Increasing amounts of MutS$_{M88C}$-IAANS result in an increase in anisotropy ($<r>$) (from formation of a larger, slower tumbling protein-DNA complex). The data fit to a quadratic equation to yield a $K_D$ of 29.4 nM C) Titration of 100 nM MutS$_{M88C}$-IAANS with increasing concentrations of unlabeled 37-mer +T DNA protein causes a decrease in fluorescence signal (red trace), yielding an apparent $K_D = 8.8$nM. Titration with increasing concentrations of unlabeled 37-mer G:C DNA, 40% less decrease in IAANS fluorescence (blue trace). The binding affinity appears similar to +T DNA, however the measurement errors are high, and additional experiments are required for an accurate estimate. (For both +T and G:C DNAs, data from three independent experiments were averaged and the standard deviation is shown by error bars).
**Pre-steady-state kinetics of MutS<sub>M88C-IAANS</sub> interaction with 37-mer+T DNA**

The MutS<sub>M88C-IAANS</sub> reporter protein enables us to monitor MutS-DNA complex in solution from the perspective of changes in protein conformation; therefore, it can be utilized to provide new information on how the interaction occurs. The data shown above provide information after all species have reached equilibrium, whereas pre-steady-state data can provide important information regarding the approach to equilibrium. Specifically, if the interaction occurs in the millisecond timescale (as is the case for many macromolecular processes), the stopped-flow method experiments can provide detailed information on the initial encounter rate and perhaps identify intermediate species that might form before equilibrium is reached. Additionally, measuring the rate constants facilitates development of a complete kinetic model and allows us to identify critical rate-limiting steps in the reaction. Thus far, this type of real-time analysis has not been performed from the perspective of the MutS protein.

Upon mixing equal volumes of 0.2μM MutS<sub>M88C-IAANS</sub> (0.1μM final) and 2μM 37-mer+T DNA (1μM final), the fluorescence signal decreased rapidly as expected from the equilibrium experiments. (Figure 3.3 B). The experiment was performed at higher concentrations of DNA, and fit the data well to a second-order exponential equation. (see Materials and Methods). In order to analyze the DNA concentration dependence of the binding kinetics, the apparent on rate of MutS<sub>M88C-IAANS</sub> (k<sub>obs</sub>) were plotted versus 37-mer+T DNA concentration (Figure 3.3 C). Two novel and potentially important findings were obtained from these plots: 1) Two distinct phases, fast and slow, are observed for this binding reaction and 2) The fastest phase reaches
saturation around 3μM 37-mer +T DNA concentration. The slow phase appears to have a linear relationship to DNA concentration which is consistent with a pseudo-first order reaction, (See Methods and Materials) and from this relationship, the $k_{on}$ can be calculated from the slope, based on the equation $k_{on} = k_{obs}[S] + k_{OFF}$. From these data, the slow step has a biomolecular rate of $0.63 \times 10^6 \text{M}^{-1}\text{s}^{-1}$.

The exact identity or mechanistic step related to the slower linear phase is not clear. Possible phenomena include conformational changes from non-specific binding of the protein to matched DNA or the ends of the linear DNA substrate, conformational changes from specific mismatch recognition, or potentially even photobleaching of the IAANS fluorophore. Control experiments mixing MutS<sub>M88C</sub>-IAANS and buffer were performed at the beginning and the end of all experimental sets in order to correct for the decrease in fluorescence associated with photobleaching. (See Materials and Methods); however, because the amplitude of the slow phase is quite low, photobleaching cannot be completely ruled out. Further experiments are planned in order to determine if the slow phase is in fact a novel mechanistic step or an experimental artifact. (Further addressed in the Discussion section)

The DNA concentration dependence of the fast phase rate fit to a hyperbolic function to yield a maximum rate of of approximately $18 \text{s}^{-1}$ and a $K_D$ of $0.09 \mu\text{M}$ (Figure 3.3 C). The initial linear portion of the curve (from 0.5 – 3μM DNA), yields an apparent biomolecular rate constant of $5 \times 10^6 \text{M}^{-1}\text{sec}^{-1}$. This binding rate is similar to that obtained from previous stopped-flow measurements of MutS binding to a 2-AP labeled +T DNA substrate, $3.6 \times 10^6 \text{M}^{-1}\text{s}^{-1}$ (Jacobs-Palmer and Hingorani 2007). The similarity between the binding rate constants indicates that the fast phase reports
MutS binding specifically to the +T site on DNA. The 2-AP experiments were performed with constant +T DNA concentration and increasing MutS up to only 0.5 μM. This restriction was due to the fact that higher MutS concentrations resulted in a poor signal to noise ratio due to excess excitation of aromatic protein residues (2-AP is excited at 315nm, which is close to the absorption maxima of Phe, Tyr, and Trp residues). Our MutS<sub>M88C-IANS</sub> assay allowed for greater concentrations of the titrant (in this case, the unlabeled DNA) to be added to the reaction and thus could explain why saturation of the binding rate was detectable in these experiments.

In summary, from the fast phase kinetics, we are now seeing for the first time an intramolecular event (whose rate is independent of reactant concentration), which potentially could be a change in protein conformation after the initial encounter. Together, these results suggest a two-step event characterized by rapid initial encounter followed by a slower step related to a specific conformational change of MutS to complete the mismatch recognition process.
Figure 3.3: MutS exhibits two-step DNA binding kinetics
**Figure 3.3: MutS exhibits two-step DNA binding kinetics.** A) A schematic of the stopped-flow experiments. Samples were excited at 326nm and all emission wavelengths greater than 350nm were collected. The experiments were performed with 0.1μM MutS<sub>M88C-IAANS</sub> and at 25°C. B) Increasing concentrations of 37-mer+T DNA (0 – 10μM final) result in a decrease in IAANS fluorescence that fits to a second order exponential after photobleaching correction. C) A plot analyzing the concentration dependence on the kinetic parameters of the binding reaction. The two apparent \( k_{on} \) rates were plotted *versus* DNA concentration. There is a fast step with an initial slope that yields \( k_{on} = 5 \times 10^6 \text{M}^{-1} \text{sec}^{-1} \) and then saturates around 3μM DNA at \( \sim 18 \text{sec}^{-1} \). There is a significantly slower step with a \( k_{obs} = 0.63 \times 10^6 \text{M}^{-1} \text{sec}^{-1} \). Two independent experiments were averaged for traces for 1, 2, and 4μM DNA and the error bars correspond to the standard deviation for those points. D) A plot of the total change in amplitude *versus* 37-mer+T DNA concentration. The data fit to a quadratic equation for a 1:1 ligand-macromolecule binding reaction, yielding a high apparent \( K_D \) of 0.57μM. Two independent experiments were averaged for traces for 1, 2, and 4μM DNA and the error bars correspond to the standard deviation for those points.
In order to obtain a deeper understanding of the MutS-mismatch interactions from these kinetic traces, a brief review of the rational underlying the rate constants for a simplified two-step enzyme-substrate binding reaction is presented below.

Ideally, these equations show how a saturation event for a $k_{\text{obs}}$ versus reactant concentration plot can be interpreted as showing a two-step binding event.

**Rapid Equilibrium Binding (2 step reaction scheme):**

$$
E + S \leftrightarrow \begin{array}{c}
\text{ES} \\
\text{ES}'
\end{array} \quad \begin{array}{c}
K_1 \\
k_2
\end{array} \quad \begin{array}{c}
k_2 \\
K_1
\end{array}
$$

**Rate constants and total enzyme concentration:**

$$
K_1 = \frac{[\text{ES}]}{[E][S]} \quad K_2 = \frac{[\text{ES}']}{[\text{ES}]} = \frac{[\text{ES}']}{K_1[E][S]} \quad [E]_0 = [E] + [\text{ES}] + [\text{ES}']
$$

**Ratio of [ES] and [ES'] species to total enzyme concentration:**

$$
\frac{[\text{ES}]}{[E]_0} = \frac{K_1[S]}{1+K_1[S]+K_1K_2[S]} \quad \frac{[\text{ES}']}{[E]_0} = \frac{K_1K_2[S]}{1+K_1[S]+K_1K_2[S]}
$$

**Rate of change of [ES] and [ES'] species (proof not shown):**

$$
\frac{[\text{ES}]}{[E]_0} = \frac{K_1[S]}{1+K_1[S]+K_1K_2[S]} \left(1-e^{-k_{\text{obs}}t}\right) \quad \frac{[\text{ES}']}{[E]_0} = \frac{K_1K_2[S]}{1+K_1[S]+K_1K_2[S]} \left(1-e^{-k_{\text{obs}}t}\right)
$$

**All species decay by the same rate constant, $k_{\text{obs}}$:**

$$
k_{\text{obs}} = \left(\frac{K_1[S]}{1+K_1[S]}\right)k_2 + k_2
$$
A theoretical plot of a two-step binding plot of $k_{obs}$ versus substrate concentration:

From these equations, it appears that MutS$_{M88C-IAANS}$ interaction with a mismatched DNA involves a rapid initial step that establishes equilibrium between the protein and DNA followed by a slow step correlated to the formation of a specific enzyme-substrate complex. While the saturating phase is left open to interpretation, it is likely to be a specific conformational change of the protein before completion of the reaction (more completely addressed in the Discussion section).

The stopped-flow data can additionally offer information on the degree of binding between MutS$_{M88C-IAANS}$ and mismatch DNA. The maximum total change in amplitude of a trace can be interpreted as the total amount of MutS$_{M88C-IAANS}$ bound to mismatched DNA under our experimental conditions. Any amplitude change that is lower than the maximum change corresponds to a fraction of protein bound or active species in the reaction as shown in the equation below:

$$\text{Degree of Binding} = \bar{X} = \frac{[L]_f}{(K_D + [L]_f)}$$
where \([L]_f\) is the concentration of free ligand and \(K_D\) is the dissociation constant of the reaction. Under our experimental conditions (i.e. pseudo-first order rate reaction in which the [DNA or ligand] >> [MutS or macromolecule]) it is assumed that \([L]_f\) is equal to the total concentration of the ligand added. The total change in amplitude (i.e. the summation of the change in amplitude corresponding to the fast phase and the slow phase) \textit{versus} DNA concentration were plotted and fit to a quadratic equation (Figure 3.3 D). The amplitude increased with increasing amounts of DNA concentration until saturating at around 2µM 37-mer+T DNA concentration. The data fit to a quadratic equation yields a \(K_D\) of 0.56µM, approximately fifty-fold higher than the value obtained in the equilibrium binding assays. While this \(K_D\) may correspond to a different species formation that was not seen in the equilibrium binding experiments, it appears that more data points are needed for lower concentrations of mismatched DNA (i.e. between 0 – 1µM) in order to fit a more accurate curve and confirm these results. (See Discussion for further explanation).

\textit{Length Dependence on the Kinetic Phases}

Our model system measures MutS interactions with DNA using a relatively short substrate (37 base pairs); however, a recent single molecule fluorescence resonance energy transfer (smFRET) study has shown that MutS can bind to a long homoduplex DNA substrate and slide along the duplex (Lee, Fishel et al. 2006). To begin investigating MutS actions on a longer DNA substrate, similar stopped-flow experiments were performed with an 80-mer+T substrate containing the 37-mer+T sequence as a central core. An initial equilibrium binding experiment was performed.
in order to ensure the protein could still bind tightly to a longer DNA substrate. As with the 37-mer +T data, a fluorescent titration experiment was performed using MutS_{M88C-IAANS} as a reporter and the unlabeled 80-mer+T DNA as a substrate. The fluorescence signal decreased when 100 nM MutS_{M88C-IAANS} was mixed with increasing amounts of the mismatched DNA (Figure 3.4). Seven titration points were obtained from 0 – 200 nM 80-mer+T DNA concentration and the data fit to a quadratic equation corresponding to a 1:1 ligand-macromolecule binding event and yield a $K_D$ of 0.6 nM, suggesting a slightly tighter affinity of the protein for a longer DNA substrate.
Figure 3.4: MutS$_{M88C}$-IAANS binds a longer DNA substrate with high affinity.

Titration of MutS$_{M88C}$-IAANS (0.1μM) with increasing concentrations of unlabeled 80-mer +T DNA protein causes a decrease in IAANS fluorescence and yields an apparent $K_D = 0.6$ nM +/- 1.7 nM. Three independent traces were corrected for photobleaching and averaged. (The standard deviation is shown by error bars.)
In order to achieve a high signal and minimize background noise, the MutS concentration used in the equilibrium binding experiments (100 nM) is approximately ten times greater than the expected dissociation constant, and thus the reaction is under stoichiometric conditions. The apparent $K_D$ determined under these conditions is most likely greater than the actual dissociation constant. Thus, a more accurate $K_D$ for this experiment most likely falls between 0.6 – 6 nM, which is in the range of the expected $K_D$ of 8 – 12 nM for this binding reaction. While stoichiometric conditions are non-ideal for obtaining accurate $K_D$ measurements, the data shows that MutSM88C-IAANS can still bind tightly to the mismatched substrate. Additionally, the experiment shows the 80-mer+T substrate can be utilized to obtain kinetic data under our experimental conditions.

Once the equilibrium binding constant for this protein-DNA interaction was determined, a real time kinetic experiment was performed on the KinTek stopped-flow apparatus to initiate an investigation into the DNA length dependence of rate constants governing the MutS-DNA binding event. All experiments were performed under the same conditions as the 37-mer+T DNA experiments. Equal volumes of 0.2µM MutSM88C-IAANS (0.1µM final) and 0 – 20µM 80-mer +T DNA (0 – 10µM final) were mixed and observed for 4 seconds, taking 500 points between time 0 – 1 sec (to catch points associated with the fast phase) and 500 points between 1 – 4 sec. Upon rapid mixing, the signal from the IAANS fluorophore decreased at all concentrations higher than 1µM 80-mer+T DNA, indicating binding (Figure 3.5 B). Unlike the 37-mer data, these traces fit better to a first order exponential. It appears that the longer DNA substrate resulted in a loss or masking of the second distinct
slower phase, further highlighting the difficulty of interpreting the slow phase in the 37-mer+T data.

To assess the 80-mer+T concentration dependence on MutS\(_{M88C-LAANS}\) and DNA interactions, the apparent on rates gathered from the first exponential fit \textit{versus} DNA concentration were plotted (Figure 3.5 C). Strikingly, the results yield a slower binding constant and the fast phase does not saturate, even at 10\(\mu\)M 80-mer +T, the highest concentration of DNA used in our experiments. The observed rates showed a linear dependence on DNA concentration and from the equation \(k_{\text{obs}} = k_{\text{on}}[\text{DNA}] + k_{\text{off}}\), and a bimolecular rate constant of \(0.76\times10^6\ \text{M}^{-1}\text{sec}^{-1}\) was determined (assuming for the moment that \(k_{\text{off}}\) is a relatively low value). The rate of the binding event between MutS and mismatched DNA decreased by nearly seven-fold (compared to the on rate calculated for the 37-mer+T DNA) and did not reach the maximum rate of \(18\text{sec}^{-1}\). Taken together, these data suggest that MutS binds more slowly to a mismatch site in the context of a longer stretch of matched DNA.
Figure 3.5: MutS exhibits relatively slow one-step DNA binding kinetics on a longer DNA substrate

A) MutS<sub>M88C-IANS</sub> with 80-mer+T DNA

\[ \lambda_{Ex} = 326 \text{ nm} \]  
\[ \lambda_{Em} > 350 \text{ nm} \]

B) Pre-Steady State Binding of MutS<sub>M88C-IANS</sub> with a 80-mer+T DNA Substrate

C) 80-mer +T DNA Concentration Dependence on the Observed Binding Rates for MutS<sub>M88C-IANS</sub>

\[ k_{obs} = 0.78 \times 10^4 \text{ M}^{-1} \text{ sec}^{-1} \]

D) Total Change in Amplitude of MutS<sub>M88C-IANS</sub> and 80-mer+T DNA Pre-Steady State Binding
Figure 4.5: MutS exhibits relatively slow one-step DNA binding kinetics on a longer DNA substrate A) A schematic of the stopped-flow experiment. Samples were excited at 326nm and all emission wavelengths greater than 350nm were collected. The experiments were performed with 0.1µM MutS_{M88C-IAANS} at 25°C. B) Increasing concentrations of 80-mer+T DNA (0 – 10µM final) result in a decrease IAANS in fluorescence that fits to a first order exponential after photobleaching correction. C) A plot analyzing the concentration dependence on the kinetic parameters of the binding reaction. The observed on rate was plotted versus DNA concentration. There is kinetic step has a slope that yields a $k_{on} = 0.76 \times 10^6 \text{M}^{-1} \text{sec}^{-1}$ and does not saturate under our experimental conditions. D) The total amplitude was plotted versus DNA concentration. While the data was too noisy to fit a hyperbolic function, the maximum amplitude is similar to the maximum amplitude in the 37-mer+T experiments.
The total change in amplitude was also monitored for this 80-mer+T interaction and while the data was too noisy to fit quadratic binding curve, the maximum change in amplitude (~0.12) is similar to the 37-mer+T amplitude maximum (~0.11) which suggests that the ultimate complex is the same for both 37-mer and 80-mer DNA. (Figure 3.5 D). Further 80-mer+T experiments are needed to obtain reliable data sets and additionally experiments need to be repeated with DNA concentrations below and closer to the estimated $K_D$ value for MutS-DNA binding.

**Nucleotide Interactions with MutS$_{M88C-IAANS}$**

*T. aquaticus* MutS has a C-terminal ATPase domain opposite and about 70Å apart from the DNA binding domain (Figure 1.3), and the allostERIC relationship between its ATPase activity and mismatch binding is poorly understood. Previous reports have shown that MutS can rapidly hydrolyze ATP in the absence of a mismatch; however, MutS binding to a mismatch significantly inhibits ATP hydrolysis. (Anthony and Hingorani 2004; Jacobs-Palmer and Hingorani 2007). In order to investigate how nucleotides affect MutS conformation as it interacts with DNA, we performed stopped-flow experiments with MutS$_{M88C-IAANS}$ and nucleotides (both with and without mismatched DNA).

Equal volumes of 0.2μM MutS$_{M88C-IAANS}$ (0.1μM final) and 200μM of ATP, ADP, or ATPγS (100μM final) were mixed and the IAANS fluorophore was selectively excited at 326nm (Figure 3.6 A, B). It should be noted that these experiments were performed at 40°C, which is closer to the temperature at which *T. aquaticus* lives at (~70°C), and ultimately all pre-steady-state data will be obtained at
this or higher temperatures. Upon rapid mixing with ADP (the red trace), the IAANS fluorescence decreased, suggesting that the DNA binding domain changes conformation upon binding the nucleotide. The trace fit well to a first-order exponential, with a $k_{\text{obs}}$ of 2.4 sec$^{-1}$. From this value, a bimolecular rate constant of 0.024 $\mu$M$^{-1}$ sec$^{-1}$ can be estimated by dividing the observed rate by the concentration of ADP in the reaction (100 $\mu$M). When mixed with ATPγS (purple trace) the IAANS fluorescence increased rapidly, suggesting that the DNA binding domains were changing in a distinct manner upon ATPγS binding compared to ADP binding. The data fit to a single-exponential with a rate constant of 0.15 $\mu$M$^{-1}$ sec$^{-1}$. The IAANS signal leveled off over time, suggesting that after binding the nucleotide analogue, a stable protein-ATPγS complex was formed.

When MutS was rapidly mixed with ATP (blue trace), there was a rapid first phase with an increase in IAANS fluorescence, followed by a slower phase with a decrease in fluorescence. The initial increase in signal can be interpreted as ATP binding, for it similar to that obtained for ATPγS binding. The decrease in signal suggests ATP hydrolysis, for the decreasing phase resembles the data obtained with ADP. This interpretation is consistent with the fact that MutS binds and hydrolyzes ATP rapidly in the absence of mismatched DNA. The first up-phase had a calculated rate constant of 0.41 $\mu$M$^{-1}$ sec$^{-1}$ which is on the same order of magnitude as the ATPγS binding rate, further suggesting this phase represents the ATP binding step. From the second phase, a rate constant of 7 sec$^{-1}$ was obtained, which is similar to the 10 sec$^{-1}$ ATP hydrolysis rate measured previously for MutS under the same conditions (Jacobs-Palmer and Hingorani, 2007). These experiments show for the first time how
the environment within the mismatch binding domain is changing conformation upon binding to a nucleotide in the absence of a mismatch and explicitly show the ATP binding and subsequent ATP hydrolysis events in the reaction.

In order to investigate the interactions of MutS and nucleotides in the presence of a mismatched DNA, the same stopped-flow experiments were performed with MutS<sub>M88C-IAANS</sub> initially pre-incubated for two minutes with a 37-mer+T DNA substrate (Figure 3.6 C,D). When mixed with ATP and ATP<sub>γ</sub>S, (blue and purple traces, respectively) the IAANS signal increased in a similar fashion as the no DNA traces. These data show that while MutS still binds ATP and ATP<sub>γ</sub>S, it does so with a one hundred-fold slower rate of 0.017μM<sup>-1</sup>sec<sup>-1</sup> for both ATP and ATP<sub>γ</sub>S when MutS is on DNA. Additionally for the ATP trace, the second phase associated with a decrease in IAANS fluorescence is lost, suggesting that the protein is unable to hydrolyze ATP when bound to mismatched DNA. These results are consistent with the reported suppression of MutS ATPase activity when the protein is bound to a mismatch (Anthony and Hingorani 2004; Jacobs-Palmer and Hingorani 2007).

When mixed rapidly with ADP, the IAANS fluorescence decreased even when the protein was initially pre-incubated with mismatched DNA, and initially at a low signal (for MutS binding to a mismatch results in a lowered IAANS signal). The data fit to a single-exponential function and yielded a <i>k<sub>on</sub></i> of 0.051μM<sup>-1</sup>sec<sup>-1</sup>, which is similar to the on rate of ADP for MutS in the absence of the mismatch, suggesting that MutS binding to ADP is not affected by the presence of a mismatch. In summary, these results indicate that ATP binding is slowed and hydrolysis is suppressed by the presence of a mismatch, while ADP can still bind to MutS regardless of the presence
of DNA (the ADP dissociation rate may change, but is not measurable in this experiment). Table 3.1 summarizes the rate constants for MutS interaction with nucleotides both with and without a mismatched DNA.
Figure 3.6: ATP binding is slower and hydrolysis is suppressed when MutS is bound to a mismatch.
Figure 3.6: ATP binding is slower and hydrolysis is suppressed when MutS is bound to a mismatch: A) A schematic of the stopped-flow experiments with nucleotides without a mismatched DNA substrate. Samples were excited at 326nm and all emission wavelengths greater than 350nm were collected. The experiments were performed with 0.1μM MutS<sub>M88C-IANS</sub> and at 40°C. B) Rapid mixing with 100μM ADP (red race) results in a decrease in IAANS fluorescence that fits well to a first-order exponential after photobleaching correction and yields a rate constant of 0.024μM<sup>-1</sup>sec<sup>-1</sup>. Rapid mixing with 100μM ATPγS (purple trace) results in an increase in IAANS fluorescence that fits well to a first-order exponential and yields a rate constant of 0.15μM<sup>-1</sup>sec<sup>-1</sup>. Rapid mixing with 100μM ATP (blue trace) results in an increase in IAANS fluorescence followed by a decrease in signal that fits well to a second-order burst equation and yields a rate constant of 0.41 μM<sup>-1</sup>sec<sup>-1</sup> for the first phase and 7 sec<sup>-1</sup> for the second phase. C) A schematic of the stopped-flow experiments with nucleotides with a mismatched DNA substrate performed under the same conditions. D) Rapid mixing with 100μM ADP (red race) results in a decrease in IAANS fluorescence that fits well to a first-order exponential after photobleaching correction and yields a rate constant of 0.05μM<sup>-1</sup>sec<sup>-1</sup>. Rapid mixing with 100μM ATPγS (purple trace) results in an increase in IAANS fluorescence that fits well to a first-order exponential and yields a rate constant of 0.02μM<sup>-1</sup>sec<sup>-1</sup>. Rapid mixing with 100μM ATP (blue trace) results in an increase in IAANS fluorescence that fits well to a single-order burst equation and yields a rate constant of 0.02 μM<sup>-1</sup>sec<sup>-1</sup>.
Table 3.1: Rate constants for MutS interaction with nucleotides in the presence or absence of mismatched DNA.

<table>
<thead>
<tr>
<th></th>
<th>No Mismatched DNA $k_{on}$ (µM$^{-1}$sec$^{-1}$)</th>
<th>Matched DNA $k_{on}$ (µM$^{-1}$sec$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>ATPγS</td>
<td>0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>ATP</td>
<td>Binding: 0.41</td>
<td>Binding: 0.02</td>
</tr>
<tr>
<td></td>
<td>Hydrolysis: 7 sec$^{-1}$</td>
<td></td>
</tr>
</tbody>
</table>
IV. Discussion.

DNA mismatch repair (MMR) is responsible for recognizing post-replicative errors that are missed by the DNA polymerase proofreading mechanisms and correcting them before they are fixed into the genome (Modrich and Lahue 1996; Kunkel and Bebenek 2000). While the main proteins of the MMR pathway have been identified and crystalized, the exact mechanisms through which these proteins communicate in order to perform their crucial cellular functions continue to be under active investigation in several laboratories; for example, questions remain regarding the MutS mechanism, and a detailed understanding is important if we are to discover how hundreds of cancer-linked missense mutations in MutS affect its structure and function (a complete list of pathogenic mutations for MSH2 and MSH6 can be found through the online LOVD database). MutS interactions with a mismatched DNA in the presence and absence of nucleotides have been investigated over the past few years by the Hingorani group; however, all previous experiments have monitored this binding reaction through a signal from the DNA substrate.

Different in-solution-assays have been developed to monitor in real-time the MutS-mismatch binding event (fluorescence anisotropy of a 5’-end TAMRA labeled DNA substrate and a 2-AP adenine analogue immediately adjacent to a +T insertion; Jacobs-Palmer and Hingorani 2007; Anthony and Hingorani 2004). While these data have provided important insights into the binding kinetics, the assays used thus far have not yielded vital information on how the protein behaves or changes during the mismatch recognition process. The results of this study provide for the first time a view of MutS DNA binding and ATPase activities from the perspective of the
protein, and thus provide new insights into mismatch recognition. Through the development of this direct, sensitive biosensor, we can now gain a wealth of new molecular insights into the structure/dynamics of the MutS protein.

*MutS\textsubscript{M88C-IAANS} is an Effective Reporter for MutS Action on DNA*

In 2006, Ban et al. developed a *T. aquaticus* MutS fluorescent biosensor in order to create a sensitive and relatively simple method for detecting mismatches, or single nucleotide polymorphisms (SNPs) for genomic studies. 2-(4’-(iodoacetamido)anilino) naphthalene-6-sulfonic acid (IAANS) was chosen as a fluorescent probe as it had previously been shown to have excitation and emission spectra that are highly influenced by changes in protein conformation. Their study indicated that MutS with IAANS conjugated to the T469 position on the clamp domain is an effective reporter for DNA binding. Upon mixing labeled MutS with mismatched DNA, the fluorescence intensity of IAANS decreased, leading the group to hypothesize that when MutS binds the mismatch, the IAANS probes are brought into close proximity and undergo self-quenching. (Ban et al. 2006).

Here, we have further developed the MutS biosensor by mutating the single native cysteine at the C42 position of the *T. aquaticus* MutS to an alanine and, after careful sequence and structure alignments (performed by F. Noah Biro, a graduate student in the Hingorani group) choosing the M88 position as the site for IAANS conjugation. M88 is in the DNA binding domain, which is believed to undergo a large conformational change upon mismatch binding, based on crystallographic structures of MutS free and bound to DNA (Figure 1.3; Obmolova and Yang, 2000).
Specifically, it appears that the mismatch binding domains in each subunit of the dimer are highly dynamic in the absence of DNA, and come together to form an ordered clamp around the DNA on binding to a mismatch. A recent Molecular Dynamics simulation of the human MutS homologue, MSH2-MSH6, also suggests that the mismatch binding and clamp domains are in an open position and highly dynamic in the absence of mismatched DNA (Feig, Mukherjee 2009). Through site-directed mutagenesis, F. Noah Biro mutated the methionine to cysteine and successfully labeled the mutant protein with IAANS at a ratio of 1.7:1 (IAANS:MutS dimer; Figure 2.1). A fluorescence anisotropy titration experiment with a 37-mer+T DNA_TAMRA substrate confirmed that the MutS_{M88C-IAANS} mutant still binds a mismatched substrate with wildtype affinity, confirming that the manipulations done on the MutS protein did not compromise its mismatch binding function (Figure 3.1 B).

A complementary titration experiment, in which IAANS fluorescence was measured, a significant decrease in signal intensity was detected when MutS_{M88C-IAANS} was mixed with an unlabeled +T DNA substrate. The data fit to a 1:1 ligand-macromolecule quadratic equation yielded a dissociation constant of 9 nM, consistent with previous equilibrium experiments which typically yielded a $K_D$ for a MutS binding +T DNA in the range of as 8 – 12 nM (Figure 3.1 C). When the same experiment was performed with a 37-mer G:C DNA substrate (in which the +T was replaced by a G:C base pair), IAANS fluorescence decreased, but the change in intensity was less pronounced compared to the +T data. MutS has been shown to have a low affinity for matched DNA, although reports have been conflicting as to the
exact $K_D$ for the interaction, and the slight decrease in signal suggests that the MutS$_{M88C-IAANS}$ reporter is sensitive enough to monitor weak MutS:DNA interactions as well as high affinity binding to a mismatch.

The equilibrium binding experiments with MutS$_{M88C-IAANS}$ gives us confidence that this reporter be utilized in quantitative, in-solution based DNA binding assays, and that it may additionally be sensitive to MutS actions other than mismatch recognition.

**Mismatch Recognition by MutS Involves Two-Step Binding**

The stopped-flow method enables pre-equilibrium and pre-steady state kinetic measurements of MutS activities during the mismatch repair reaction. In order to investigate the process of mismatch recognition, specifically how MutS works to recognize an error in DNA and signal its repair, MutS$_{M88C-IAANS}$ was mixed rapidly with DNA in a KinTek stopped-flow instrument, and the reaction was monitored within the millisecond time scale. One novel finding from the results is that MutS binding to a mismatch involves two distinct steps: one dependent on and the other independent of DNA concentration.

All stopped-flow experiments were performed under pseudo-first order reactions in which the substrate concentration (unlabeled DNA) far exceeded enzyme concentration (labeled MutS$_{M88C-IAANS}$) so that a first order exponential function, $k_{on}[DNA]$, could describe the reaction, where $k_{on}$ is the bimolecular binding rate constant. Assuming that the MutS-DNA interaction is reversible, the apparent on rate should be linearly dependent on DNA concentration as described by the equation $k_{obs}$.
\[ = k_{\text{on}}[\text{DNA}]+k_{\text{off}}, \] where \( k_{\text{obs}} \) is the observed on rate and \( k_{\text{off}} \) is the dissociation constant for MutS coming off the mismatch. To determine the concentration dependence of the interaction between MutS and a mismatch, \( k_{\text{obs}} \) values obtained from the first fast phase of the reaction were plotted versus DNA concentration and yielded a hyperbolic curve that saturated at concentrations higher than 3\( \mu \)M 37-mer+T DNA; beyond that the observed rates were unchanged even at 10\( \mu \)M DNA concentration. From previous experiments performed with DNA containing a 2-AP fluorescent reporter adjacent to a +T insertion, we know that MutS dissociation from the +T site occurs at a rate of 0.08 s\(^{-1}\) (Jacobs-Palmer, Hingorani 2007). Thus, the initial slope of the hyperbola could be used to calculate a \( k_{\text{on}} \) of 5\( \times 10^6 \)M\(^{-1}\)s\(^{-1}\), which is almost the same value as the previous estimate of 3.6\( \times 10^6 \)M\(^{-1}\)s\(^{-1}\) from measurements with 2-AP labeled DNA. The consistency indicates that the two assays, MutS conformation reported by IAANS fluorescence and MutS-mismatch site specific interaction reported by 2-AP fluorescence, are measuring the same event—mismatch recognition.

In addition, the binding rate approaches a maximum value, which suggests a two-step binding process consisting of a second slow step that is independent of reactant concentrations. Two-step binding, in which all species in the reaction decay at the same rate constant, \( k_{\text{obs}} \), approaches a maximum rate which correlates to the saturation of an enzyme-substrate (ES) complex (by analogy to Michaelis-Menten kinetics). Results from the 37-mer+T DNA experiments indicate that the maximum rate of formation of the ES complex under our experimental conditions is approximately 18sec\(^{-1}\). We can speculate that the first binding step involves formation
of an initial MutS-DNA complex and second step involves a conformational change in MutS that occurs on mismatch recognition to complete the binding reaction.

It should be noted that in the earlier DNA binding experiments with 2-AP labeled +T DNA, the observed rates appeared to saturate at high reactant concentrations (MutS in that case) (Fig 4.1). However, at MutS concentrations greater than 1μM, the signal to noise ratio was high due to overlapping and nonspecific excitation of tryptophan residues at 315 nm, and the data were deemed unreliable in the absence of any independent confirmation of a saturating rate. Now, with data from MutS\textsubscript{M88C-IAANS} experiments showing the same phenomenon (despite slight differences in DNA sequences, which may result in slightly different maximal binding rates), we can confirm that MutS binding to the mismatch site (as reported by the 2-AP assay) is limited by a substrate concentration independent step, which likely involves a conformational change in the protein (as reported by the IAANS assay).

The two-step binding model also explains why the apparent $K_D$ values obtained from plots of $k_{obs}$ versus substrate concentration are so high (> 1 μM) compared to $K_D$ values measured for the same DNA substrates with the same assay under equilibrium conditions. In the case of the kinetic data, we are observing the reaction approach equilibrium, and if there is initial weak complex formation (in rapid equilibrium) during the binding process, the $K_{1/2}$ obtained from the hyperbolic function is actually the $K_D$ for the initial complex (> 1 μM) not the final high affinity complex (~10 nM).

Finally, the second slow phase observed in the binding experiments with MutS\textsubscript{M88C-IAANS} and 37-mer +T DNA cannot be incorporated into the binding model.
at this time, since there is significant uncertainty as to the nature of the event being reported by the slow decrease in MutS$_{M88C}$-IAANS fluorescence intensity (Figure 3.2 C, red trace). Interpretation of this phase as a bona fide mechanistic step must be taken with caution, for a variety of phenomena could be contributing to the signal including an on-pathway event (e.g., an intermediate binding step) or an off-pathway event (e.g., MutS binding to the ends of the short linear DNA, as has been suggested by previous AFM experiments; (Erie, Yang et al. 2005), or perhaps even photobleaching that is different in the MutS-DNA complex relative to the photobleaching correction experiments performed with MutS alone.

We have attempted initial experiments to characterize the slow phase in the kinetic data; for example, monitoring MutS$_{M88C}$-IAANS binding to a short linear DNA substrate with biotin-streptavidin blocked ends (data not shown). If the phase in question was in fact MutS binding to DNA ends, it was hypothesized that the kinetic date would show loss of the second phase and fit well to a single exponential function. Unfortunately, the signal to noise ratio was poor in these experiments due to the high concentrations of streptavidin needed to create blocked DNA ends (perhaps due to excitation of the excess aromatic residues; $\lambda_{\text{EX-TRP}}=290\text{nm}$, $\lambda_{\text{EX-IAANS}}=326\text{nm}$; or due to light scattering). Future experiments are being developed with a Cy3 labeled MutS reporter that can be excited at a longer wavelength and DNA minicircles that do not any ends that may possibly be bound by MutS.
Figure 4.1: Previously observed saturation event monitored by a 2-AP DNA substrate

![Graph showing the dependence of observed rate constant on MutS concentration](image)

**Figure 4.1: Previously observed saturation event monitored by a 2-AP DNA substrate:** A) A plot of the observed rate constant versus MutS concentration monitored by a 2-AP fluorescence signal. Note the potential saturation event occurring with the 23-mer+T DNA substrate; however, the signal/noise ratio was poor in these experiments. (Jacobs-Palmer 2007, unpublished)
The MutS Mismatch Recognition Event is Influenced by the Concentration of Matched Base Pairs

Recently, a single molecule fluorescence experiments with quantum dot-labeled MutS showed that the protein can slide along homoduplex DNA, possibly via 1-D rotational diffusion. It was proposed that this movement is required for scanning newly replicated DNA for mismatches. (Lee, Fishel et al. 2011). This result indicates that MutS makes weak or non-specific interactions with matched bases. Since the single molecule study reported by Gorman et al. was at low resolution (> 300 base pairs), we decided to explore the question of MutS binding to matched DNA using short DNA substrates of varying lengths (i.e., with more matched bases).

Experiments with an 80-mer+T DNA substrate (containing a central 37-mer+T core) yielded kinetic data that were distinctly different from the 37-mer data. For the 80-mer+T DNA, the corrected kinetic traces fit well to a single exponential function, suggesting that the slow phase observed with 37-mer +T (Fig 3.3) either does not occur, or is masked by the fast phase, which in the case of 80-mer +T DNA occurs at a relatively slow rate. Additionally, when the \( k_{\text{obs}} \) was plotted versus DNA concentration, we did not observe saturation at high DNA concentrations, suggesting that formation of the ES complex is dependent on the length of the substrate (the amount of matched base pairs present in the reaction). Furthermore, the linear slope yielded a bimolecular rate constant of \( 0.76 \times 10^6 \text{M}^{-1} \text{sec}^{-1} \), indicating slower apparent binding to the mismatch. Figure 4.2 shows a compiled \( k_{\text{obs}} \) vs. mismatched DNA concentration plot for the 37- and 80-mer+T substrates that highlight the differences in binding kinetics for these two substrates. Together, these data suggest that MutS
partitions between binding to matched and mismatched DNA, and excess matched base pairs lower the apparent rate and affinity of MutS binding to a mismatch site. MutS is known to bind matched DNA with low affinity ($K_D$ in low micromolar range); therefore in our experiments, we are likely measuring a combination of MutS binding directly to matched base pairs or the mismatch site, scanning matched base pairs for the mismatch site, and MutS dissociation from matched base pairs or the mismatch site. All these processes may lower the apparent binding rate constant for mismatched DNA in the presence of increasing matched DNA in the reaction.
Figure 4.1: MutS<sub>M88C-IAANS</sub> binds at an apparently slower rate to a substrate with a higher concentration of matched base pairs. MutS<sub>M88C-IAANS</sub> binds rapidly to a shorter DNA substrate \( (k_{on} = 5 \times 10^6 \text{ M}^{-1}\text{sec}^{-1}) \) followed by a slow saturating formation of an activated ES complex occurring around 18 sec\(^{-1}\) fits well to a hyperbolic curve (blue curve). Alternatively, MutS<sub>M88C-IAANS</sub> binds to a longer DNA substrate at a much slower rate \( (k_{on} = 0.71 \times 10^6 \text{ M}^{-1}\text{sec}^{-1})\) that does not saturate at substrate concentrations of 10\( \mu \text{M} \), suggesting that the presence of more matched bases is affecting the binding reaction rates.

Figure 4.2: MutS<sub>M88C-IAANS</sub> binds at an apparently slower rate to a substrate with a higher concentration of matched base pairs.

37-mer+T and 80-mer DNA Concentration Dependance of the Observed Binding Rates for MutS<sub>M88C-IAANS</sub>

\[ K_D \approx 1.3 \mu \text{M} +/- 0.5 \mu \text{M} \]

\[ k_{obs} \approx 5 \times 10^6 \text{ M}^{-1}\text{sec}^{-1} \]

\[ k_{obs} = 0.71 \times 10^6 \text{ M}^{-1}\text{sec}^{-1} \]
Further experiments with different lengths of DNA are necessary before we can develop a quantitative model that describes MutS interactions with a mismatch site in the context of excess matched DNA. Accordingly, additional experiments are planned with 23-mer and 60-mer +T DNA substrates.

**Effects of Nucleotides on MutS Interactions with a Mismatch**

Stopped-flow experiments performed by rapid mixing between MutS<sub>M88C-IAANS</sub> and nucleotides have shown, from the perspective of the protein, how nucleotide binding and hydrolysis is coupled to mismatch recognition for the first time. From previous experiments monitoring MutS ATPase activity, the Hingorani group has shown that MutS can bind ATP rapidly in the absence of a mismatch with a bimolecular binding rate constant of 2.5x10<sup>5</sup> M<sup>-1</sup>s<sup>-1</sup> and hydrolyze it at 10 s<sup>-1</sup> in the absence of DNA (40 °C) (Anthony and Hingorani 2004). The similarity in kinetic parameters obtained from previous measurements of P<sup>32</sup>-ADP formation as well as phosphate (Pi) release and current measurements of MutS conformation by IAANS fluorescence is striking; therefore, we assign specific MutS actions (conformational changes) in the absence and presence of DNA to specific events in the ATPase reaction with confidence.

In the absence of mismatched DNA, MutS binds and hydrolyzes ATP rapidly, and we observe an increase (4.1x10<sup>5</sup> M<sup>-1</sup>s<sup>-1</sup>) and then decrease (7 s<sup>-1</sup>) in MutS<sub>M88C-IAANS</sub> fluorescence that correspond well with the ATP binding and hydrolysis events (Fig 4.3). In the presence of mismatched DNA, ATP binding is observed, though at a relatively slow rate, (MutS<sub>M88C-IAANS</sub> undergoes the same increase in fluorescence with ATP and ATP<sub>γ</sub>S), but ATP hydrolysis is apparently blocked (MutS<sub>M88C-IAANS</sub>...
remains at in the high fluorescence conformation with both ATP and ATPγS over a prolonged period). According to ATPase measurements, in the presence of mismatched DNA MutS hydrolyzes ATP at a slow rate of about 0.3 s⁻¹ (Antony and Hingorani, 2004) and this event may occur only after MutS leaves the mismatch, at 0.14 s⁻¹ (Jacobs-Palmer and Hingorani, 2007). Indeed when MutS_{M88C-IAANS} fluorescence is monitored over a long period (15 seconds), the signal eventually decreases, indicating slow ATP hydrolysis (Fig 4.3). Future experiments are planned to parse out the exact timing of MutS conformational changes relative to ATP binding, hydrolysis, and ADP and Pi release, as the data will provide vital information about rate-limiting steps in the MutS-catalyzed reaction when the work is done.
Figure 4.3: Mismatched DNA inhibits rapid ATP hydrolysis by MutS

A) When MutS<sub>M88C-IANS</sub> is rapidly mixed with ATP in the absence of DNA (green trace) it binds and hydrolyzes the nucleotide rapidly. When the protein is initially incubated with 37-mer+T DNA (blue trace) the binding of ATP is slower and the hydrolysis step is absent. B) A longer time trace shows that ATP is eventually hydrolyzed. C) A complementary Pi-release assay performed by Edwin Anthony
(2004) showing that MutS has robust ATPase activity in the absence of a mismatch (Green trace) and hydrolysis is suppressed in the presence of a mismatch (Blue trace).

**Evaluation of Current MMR Models**

Both the ‘Translocation’ and ‘Stationary/DNA bending’ models propose that MutS binds a mismatch in a nucleotide-free state, while the sliding clamp model proposes that MutS scans for and binds to a mismatch when bound to ADP (Figure 1.6). The kinetic rates determined from the MutS<sub>M88C-IAANS</sub> reporter indicate that MutS is capable of rapidly binding a mismatch in the absence of nucleotides, which supports the ‘Translocation’ and ‘Stationary/DNA Bending’ models. MutS – mismatch binding experiments have not been performed with MutS bound initially to ADP yet, but rapid mixing of MutS with ADP shows that the protein’s interaction with ADP is not affected by the presence of a mismatch, suggesting the protein can still interact with both ADP and DNA simultaneously. It has been previously shown, however, that while MutS<sub>ADP</sub> can still bind a mismatch (k<sub>on</sub> = 0.1 μM<sup>-1</sup>s<sup>-1</sup>), it also has a short lifetime on the mismatch, coming off rapidly at 1.7 s<sup>-1</sup>. Together these data suggest that ADP-bound MutS may well interact transiently with DNA in search of the mismatch site, but that the nucleotide free protein appears capable of doing so as well.

We have also shown that MutS cannot hydrolyze ATP rapidly when bound to a mismatched DNA. The Translocation model suggests that MutS utilizes its ATPase activity to move away from the mismatch and continually hydrolyze ATP to bring the two flanking DNA strands through its clamp domain. Our kinetic data show no
evidence of continual hydrolysis of ATP when the protein is bound to DNA, and thus do not support the proposed motor-like function proposed for MutS in the Translocation model.

Our kinetic data has shown that MutS can still bind ATP in the presence of a mismatch. Additionally, the data suggest there is the formation of a stable Muts-ATP-mismatch complex (indicated by the long-lived high IAANS fluorescence species). The Stationary/DNA Bending model proposes that MutS binds ATP and a mismatch simultaneously to form a stable ternary structure on the mismatch, while the Molecular Switch model proposes that ATP binding by MutS activates the protein and allows it to move away from the mismatch in a hydrolysis-independent step. Our data suggest that MutS stays on the mismatch site for a prolonged period before it moves and hydrolyzes ATP. Thus, according to our current data, both the Stationary/Bending model (during the time MutS remains in ATP-bound form at a mismatch) and a Molecular Switch model (during the time MutS moves away from the mismatch but before it hydrolyzes ATP) are feasible. Future experiments are planned to further resolve which of the proposed models is closer to the actual MutS reaction mechanism.

Finally, we have shown that the MutS-mismatch binding event is dependent on the length of the linear DNA substrate. We have interpreted these results as being an effect of the concentration of matched base pairs present in the reaction. Specifically, it appears that the rapid rate of MutS binding to a mismatch is slowed when more matched base pairs are present, suggesting that MutS interactions with matched base pairs can compete, albeit poorly, for MutS binding to a mismatch. It
appears that this length-dependence trend on the kinetic parameters most readily agrees with the sliding clamp model for mismatch repair. Both the rate of mismatch binding (fast phase) and the slow conformational change of MutS (slow, saturating phase) are affected by the length of the DNA substrate, which suggests MutS is making contact with the entire linear substrate.
5: Conclusions and Future Directions.

The research reported here has furthered development of a novel fluorescent reporter of MutS actions in DNA mismatch repair, particularly the initial stages of mismatch recognition and ATPase-coupled initiation of repair. The resulting assays have begun to yield new information on the protein mechanism, particularly by providing kinetic parameters governing protein-ligand interactions in solution. Through real time kinetic analysis monitoring MutS-mismatch binding from the perspective of the protein, we have gained a more precise understanding of the binding reaction; specifically showing it occurs in two steps—formation of an initial weak complex between MutS and matched DNA or a potential mismatch site, and then formation of a high affinity complex with a bona fide mismatch, which we propose involves a rate-limiting conformational change in MutS. Moreover, we have begun to characterize the effects of excess matched DNA on the ability of MutS to find a mismatch, which provide important information on how the protein is able to find its target sites \textit{in vivo}. Finally, experiments with nucleotides have revealed the conformation changes MutS undergoes as it couples mismatch recognition to initiation of repair. This project was initiated with the hope of developing a method to minimally monitor MutS binding to mismatched DNA, and initial studies show that we now have an effective reporter for several key events involving MutS in the DNA mismatch repair reaction mechanism.
Appendices.

Appendix A: DNA Substrates

Table A.1

Single stranded DNA oligos for MutS DNA-binding and ATPase analysis

<table>
<thead>
<tr>
<th>Strand Name</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-mer +T</td>
<td>5’- CC GGC AGC TAT _TA CCG TCG CGC 3’</td>
<td>EJP22ATA</td>
</tr>
<tr>
<td></td>
<td>3’- GG CCG TCG ATA TAT GGC AGC GCG 5’</td>
<td>EJP23ATA-NoAp+T</td>
</tr>
<tr>
<td>37-mer+T</td>
<td>5’- ATG TGA ATC AGT ATG GTA TAT ATC TGC TGA AGG AAA 3’</td>
<td>EJP 37+T</td>
</tr>
<tr>
<td></td>
<td>3’- TAC ACT TAG TCA TAC CAT _TA TAG ACG ACT TCC TTT 5’</td>
<td>EJP 36</td>
</tr>
<tr>
<td>37-merG:C</td>
<td>5’- ATG TGA ATC AGT ATG GTA GAT ATC TGC TGA AGG AAA 3’</td>
<td>EJP ATA-G-base</td>
</tr>
<tr>
<td></td>
<td>3’- TAC ACT TAG TCA TAC CAT CTA TAG ACG ACT TCC TTT 5’</td>
<td>EJP ATA-C-base</td>
</tr>
<tr>
<td>80-mer+T</td>
<td>5’- CTA AAG GGA ACA AAA GCT GGG TAC CAT GTG AAT CAG TAT GGT A T A TAT CTG CTG AAG GAA ATC TCG AGG TCG ACG GTA TC 3’</td>
<td>EJP 37core 80mer+T</td>
</tr>
<tr>
<td></td>
<td>3’- GAT TTC CCT TGT TTT CGA CCC ATG GTA CAC TTA GTC ATA CCA T T A TAT GAC GAC TTC CTT TAG AGC TCC AGC TGC CAT AG-5’</td>
<td>EJP 37core 79mer</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>2-AP</td>
<td>2-aminopurine</td>
<td></td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>adenine</td>
<td></td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine 5'-diphosphate</td>
<td></td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5' -triphosphate</td>
<td></td>
</tr>
<tr>
<td>ATPγS</td>
<td>adenosine 5' -o- (3-thiotriphosphate)</td>
<td></td>
</tr>
<tr>
<td>Bp</td>
<td>Basepairs</td>
<td></td>
</tr>
<tr>
<td>ddH₂O</td>
<td>deionized water</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
<td></td>
</tr>
<tr>
<td>ds</td>
<td>double stranded</td>
<td></td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine-N,N,N',N'-tetraacetic acid</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
<td></td>
</tr>
</tbody>
</table>
HCl  hydrochloric acid

HEPES  N-[2-hydroxyethyl]piperazine-N’-[2-ethanesulfonic acid]

IAANS  2-(4’-(iodoacetamido)anilino)naphthalene-6-sulfonic acid, sodium salt

IPTG  isopropyl-beta-D-thiogalactopyranoside

KCl  potassium chloride

$K_D$  dissociation constant

kDa  kilodalton(s)

LB  Luria-Bertani media

M  molar

MDCC  N-[2(1-maleimidyl)ethyl]-7-(diethylamino)coumarin-3-carboxamide

Mg  magnesium

MgCl$_2$  magnesium chloride

Min  minutes

mL  milliliter

Mlh  MutL homolog
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>Msh</td>
<td>MutS homolog</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBP</td>
<td>phosphate binding protein</td>
</tr>
<tr>
<td>Pi</td>
<td>inorganic phosphate</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethysulfonlfuoride</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>sec</td>
<td>second(s)</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td><em>Saccharomyces cerevisiae</em></td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>ss</td>
<td>single stranded</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>T. aquaticus</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>tris(hydroxymethyl)aminomethane hydrochloride</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
</tbody>
</table>
APPENDIX C: Equations

Degree of Labeling of MutS\textsubscript{M88C-IAANS}:

\[
\frac{\text{Abs } 326/\epsilon \text{IAANS}}{[\text{Abs} 280 - (\text{Abs } 326 \times \text{C.F.})]/(2 \times \epsilon \text{MutS})} = \frac{[\text{IAANS}]}{[\text{MutS}]} \]

Quadratic equation describing 1:1 ligand-macromolecule interaction (IAANS fluorescence):

\[
x = \frac{[\text{ML}]}{[\text{M}]_{\text{T}}} = \frac{([\text{L}]_{\text{T}} - \sqrt{([\text{L}]_{\text{T}} + [\text{M}]_{\text{T}} + K_D)^2 - 4[M]_T[L]_T})}{2[M]_T} \]

Hyperbolic equation (37-mer rate constants and amplitude plots):

\[
\bar{X} = \frac{[\text{L}]_f}{(K_D + [\text{L}]_f)} \]

Buffer Correction (IAANS fluorescence-based assays):

\[
\text{Corr } P_i = \frac{[B_o]}{[B_d]} *[P_i] \]

Single exponential decay (IAANS fluorescence for 80-mer+T DNA):

\[
A_t = A_0(\text{- } e^{-\text{ht}}) + F_0 \]

Double exponential decay (IAANS fluorescence for 37-mer+T DNA):

\[
A_t = A_{0\text{Phase1}}(\text{- } e^{-k(\text{phase1})t}) + A_{0\text{Phase2}}(\text{- } e^{-k(\text{phase2})t}) + F_0 \]

Burst + single exponential decay (ATP hydrolysis):

\[
A_t = A_{0\text{Phase1}}(1 - e^{-k(\text{phase1})t}) + A_{0\text{Phase2}}(\text{- } e^{-k(\text{phase2})t}) + F_0 \]
References.


