Kinetic Analysis of *Saccharomyces cerevisiae* Msh2-Msh6 DNA Binding and ATPase Activities

by

Jie Zhai

Faculty Advisor: Dr. Manju M. Hingorani

A Dissertation Submitted to the Faculty of Wesleyan University in Partial Fulfillment of the Requirements for the Degree of Master of Philosophy

Middletown, Connecticut  May 2010
Abstract

The DNA mismatch repair system (MMR) identifies replication errors and damaged bases in DNA and takes corrective actions to preserve genomic integrity. MutS performs the task of locating mismatched base pairs, loops and lesions and initiating MMR, and the fundamental question of how this protein targets specific sites for MMR is unresolved. To address this question, we examined the interactions between \textit{Saccharomyces cerevisiae} Msh2-Msh6, a eukaryotic MutS homolog, and DNA in real time. The reaction kinetics reveal that Msh2-Msh6 binds a variety of sites at similarly fast rates ($k_{\text{ON}} \sim 10^7 \, \text{M}^{-1} \, \text{s}^{-1}$). Msh2-Msh6 selectivity manifests in differential dissociation rates; e.g., the protein releases a 2-Aminopurine:T base pair $\sim 90$-fold faster than a G:T mismatch. On releasing the weak 2-Ap:T site, Msh2-Msh6 is able to move laterally on DNA to locate a nearby G:T site. The long-lived Msh2-Msh6•G:T complex triggers the next step in MMR—formation of an ATP-bound clamp—more effectively than the short-lived Msh2-Msh6•2-Ap:T complex. Mutation of glutamate in the conserved Phe-X-Glu DNA binding motif stabilizes Msh2-Msh6\textsubscript{E339A}•2-Ap:T, and the mutant can signal 2-Ap:T repair as effectively as Msh2-Msh6 signals G:T repair. These findings suggest that Msh2-Msh6 employs a targeting mechanism whereby it interrogates base pairs by rapid, transient contacts and pauses at potential MMR sites, and the longer the pause the greater the likelihood of MMR.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Vita</td>
<td>ii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter I - Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II - Materials and methods</td>
<td>28</td>
</tr>
<tr>
<td>Chapter III - Results</td>
<td>79</td>
</tr>
<tr>
<td>Chapter IV - Discussion</td>
<td>128</td>
</tr>
<tr>
<td>References</td>
<td>148</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></td>
<td>Chapter I - Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Prokaryotic MMR model</td>
<td>9</td>
</tr>
<tr>
<td>1.2</td>
<td>Eukaryotic MMR model</td>
<td>14</td>
</tr>
<tr>
<td>1.3</td>
<td>Crystal structure of <em>Thermus aquaticus</em> MutS</td>
<td>17</td>
</tr>
<tr>
<td>1.4</td>
<td>Crystal structures of <em>T. aquaticus</em> MutS and <em>H. sapiens</em> MutSα (Msh2-Msh6)</td>
<td>19</td>
</tr>
<tr>
<td>1.5</td>
<td>MMR mechanism models</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Chapter II - Materials and methods</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Overexpression and purification of <em>S. cerevisiae</em> Msh2-Msh6 from <em>E. coli</em></td>
<td>43</td>
</tr>
<tr>
<td>2.2</td>
<td>Purification of phosphate binding protein (PBP)</td>
<td>60</td>
</tr>
<tr>
<td>2.3</td>
<td>Labeling efficiency of PBP with MDCC</td>
<td>67</td>
</tr>
<tr>
<td>2.4</td>
<td>MDCC labeling quality check with SDS-PAGE gel and UV scan</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Chapter III - Results</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Equilibrium fluorescence anisotropy measurements with TAMRA end-labeled DNAs</td>
<td>83</td>
</tr>
<tr>
<td>3.2</td>
<td>Msh2-Msh6 binds MMR targets rapidly and with varying affinities</td>
<td>85</td>
</tr>
<tr>
<td>3.3</td>
<td>Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis and Pi release</td>
<td>89</td>
</tr>
<tr>
<td>Figure number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.4</td>
<td>Msh2-Msh6 interactions with +T insertions in DNA</td>
<td>91</td>
</tr>
<tr>
<td>3.5</td>
<td>Msh2-Msh6 Pi release with IDLs</td>
<td>95</td>
</tr>
<tr>
<td>3.6</td>
<td>Equilibrium 2-Ap fluorescence measurements</td>
<td>100</td>
</tr>
<tr>
<td>3.7</td>
<td>An on-site 2-Ap reporter confirms that Msh2-Msh6 dissociation rates</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>underlie target site selectivity</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>ATP hydrolysis is partially suppressed by G:C&lt;sub&gt;2-Ap&lt;/sub&gt; DNA</td>
<td>104</td>
</tr>
<tr>
<td>3.9</td>
<td>Fluorescence anisotropy shows Msh2-Msh6 binding mismatched 2-Ap:T/7bp/G:T&lt;sub&gt;TAMRA&lt;/sub&gt; and matched 2-Ap:T/7bp/G:C&lt;sub&gt;TAMRA&lt;/sub&gt; DNA rapidly</td>
<td>107</td>
</tr>
<tr>
<td>3.10</td>
<td>2-Ap fluorescence yields the same rate for 2-Ap:T/7bp/G:C&lt;sub&gt;TAMRA&lt;/sub&gt;, but 2-Ap:T/7bp/G:T&lt;sub&gt;TAMRA&lt;/sub&gt; shows biphasic kinetics</td>
<td>109</td>
</tr>
<tr>
<td>3.11</td>
<td>2-Ap fluorescence reports rapid release of Msh2-Msh6 from 2-Ap:T in 2-Ap:T/7bp/G:T&lt;sub&gt;TAMRA&lt;/sub&gt; with an external trap or buffer added after 0.25 s</td>
<td>113</td>
</tr>
<tr>
<td>3.12</td>
<td>Msh2-Msh6 moves from 2-Ap:T to G:T without dissociating from DNA</td>
<td>115</td>
</tr>
<tr>
<td>3.13</td>
<td>Msh2-Msh6 interactions with G:T&lt;sub&gt;2-Ap&lt;/sub&gt; and G:C&lt;sub&gt;2-Ap&lt;/sub&gt; DNAs in the absence and presence of nucleotide cofactors</td>
<td>117</td>
</tr>
<tr>
<td>3.14</td>
<td>Msh2-Msh6 interaction with +T DNA in the absence and presence of nucleotide cofactors</td>
<td>121</td>
</tr>
<tr>
<td>3.16</td>
<td>Msh2-Msh6&lt;sub&gt;E339A&lt;/sub&gt; Pi release</td>
<td>127</td>
</tr>
<tr>
<td>Figure number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>Chapter IV - Discussion</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>A model for Msh2-Msh6 actions during the search phase of MMR</td>
<td>140</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chapter I - Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Prokaryotic mismatch repair proteins and their eukaryotic homologs</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Chapter II - Materials and methods</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>23mer 2-Ap and TAMRA labeled DNA substrates</td>
<td>29</td>
</tr>
<tr>
<td>2.2</td>
<td>37mer 2-Ap labeled DNA substrates</td>
<td>30</td>
</tr>
<tr>
<td>2.3</td>
<td>37mer 2-Ap and TAMRA labeled DNA substrates</td>
<td>31</td>
</tr>
<tr>
<td>2.4</td>
<td>37mer TAMRA and Cy3 labeled DNA substrates</td>
<td>32</td>
</tr>
<tr>
<td>2.5</td>
<td>PBP labeling buffer exchange chart</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Chapter III - Results</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Kinetic parameters for Msh2-Msh6-DNA interactions</td>
<td>122</td>
</tr>
</tbody>
</table>
Mismatch repair

During DNA replication, DNA polymerase generates an error at a rate of one in $10^4 – 10^5$ nucleotides when it adds nucleotides to newly synthesized daughter strands (Kunkel, 2004). After this nucleotide insertion step, DNA polymerase can edit the newly synthesized daughter strands to check the correction with the help of proofreading exonuclease. This proofreading process decreases the error frequency to one per $10^6 - 10^7$ nucleotides per cell division (Kunkel, 2004). However, the efficiency of this process is still not high enough to reach the spontaneous mutation rate of 1 in $10^9 – 10^{10}$ nucleotides (Iyer et al., 2006). It turns out that in addition to the proofreading process, mismatch repair increases the fidelity of DNA replication a thousand fold higher to the final one error per $10^9 – 10^{10}$ nucleotides level (Iyer et al., 2006).

To maintain the integrity of the whole genome, mismatch repair (MMR), a highly conserved DNA repair process, is mainly responsible for resolving various base pair mismatches and insertion-deletion loops (IDLs) that are generated by DNA polymerase and have escaped from DNA polymerase proofreading (Iyer et al., 2006). Mismatch repair is also capable of correcting the mismatches caused by DNA damage, such as O6MeG and 8-oxoguanine (Iyer et al., 2006; Karran and Bignami,
1994; Mazurek et al., 2009). It also corrects mismatches generated during genetic recombination (Reenan and Kolodner, 1992).

In the MMR pathway, the MutS homolog searches, recognizes and binds mismatches on DNA of newly synthesized daughter strands, and recruits the MutL homolog to initiate the replacement of mismatches. The mechanism of MMR in *Escherichia coli (E. coli)* has been well characterized, while substantial information has been found on eukaryotes (Iyer et al., 2006).

The genetic inactivation of MMR can increase the mutation frequency by 50 – 1000 folds (Kolodner and Marsischky, 1999). The defects of the MMR proteins cause increasing rates of not only base substitution but also frameshift mutations (Kolodner and Marsischky, 1999), generate incorrect recombination in quasi-homologous sequences, and render mammalian cells resistance to DNA damaging agents (Harfe et al., 2000; Stojic et al., 2004).

The core MMR proteins have been conserved through evolution. In eukaryotes, there are several MutS (Msh) and MutL (Mlh/Pms) homologs. Defects in their functions lead to severe mutator phenotypes and genomic instability; in humans, hundreds of *hMSH2, hMSH6, hMLH1*, and *hPMS2* mutations have been linked to hereditary nonpolyposis colon cancer (HNPCC) and sporadic cancers (Peltomaki, 2001). Every year about 500,000 people from all over the world are diagnosed with colon cancer and about 3 - 4 % of them are from HNPCC (Hsieh and Yamane, 2008). Currently HNPCC can be diagnosed by genetic screening of microsatellite instability (MSI) which can reflect the MMR gene mutations from germline cells (Hsieh and
Yamane, 2008). Microsatellites are short repeated-sequence motifs that showed instability in MMR-deficient cells and were not present in normal cells (Jiricny, 2006; Peltomaki, 2001). MSI is a recently well-established phenotype found in HNPCC tumor cells (Jiricny, 2006). Although MSI is a hallmark of HNPCC (Hsieh and Yamane, 2008), the cause of MSI in HNPCC cells remains unclear. The basic research of eukaryotic MMR mechanism will support the development of new therapies for HNPCC and other MMR-deficient tumors (Li, 2008).

**MMR proteins**

MMR proteins are highly conserved from prokaryotes to eukaryotes with significant homology in structure as well as function (Fishel and Wilson, 1997).

As showed in Table 1.1, *E. coli* MMR proteins are MutS, MutL, MutH, γ–δ complex, β-clamp, UvrD (DNA helicase II), ExoI, ExoX, ExoVII, RecJ, single-stranded DNA binding protein (SSB), DNA polymerase III and DNA ligase (Kunkel and Erie, 2005). MutS forms homodimers that search and recognize mismatches and initiate MMR (Kunkel and Erie, 2005). MutS recruits MutL to the mismatch location and MutL works as a matchmaker to coordinate between mismatch recognition and excision initiation (Kunkel and Erie, 2005). MutH makes nicks on daughter strands to initiate excision. Other MMR proteins including helicases, SSB and ligases are responsible for strand excision and DNA resynthesis (Kunkel and Erie, 2005).

In eukaryotes, MMR protein homologs are highly conserved between yeast and human (Hsieh and Yamane, 2008). In *Saccharomyces cerevisiae* (*S. cerevisiae*),
six homologs (MSH1-6) of the *E. coli* MutS have been identified (Fishel and Wilson, 1997). As shown in Table 1.1, MMR proteins are more complex in eukaryotes than in prokaryotes. In human and yeast, MutS homologs are mainly MutSα (Msh2-Msh6) and MutSβ (Msh2-Msh3). The functions of both complexes are to search for mismatches during DNA replication and to initiate the MMR process. However, Msh2-Msh6 mainly recognizes single base mismatches and 1-to-2-base IDLs, while Msh2-Msh3 recognizes IDLs longer than three bases (Kunkel and Erie, 2005). MutL homologs are MutLα (Mlh1-Pms2), MutLβ (Mlh1-Mlh2) and MutLγ (Mlh1-Mlh3) and there is no MutH homolog in eukaryotes (Kunkel and Erie, 2005). Instead of MutH, eukaryotic MutL can work as an endonuclease in addition to its molecular matchmaker function (Kunkel and Erie, 2005; Li, 2008). The other components of eukaryotic MMR are the RFC complex, PCNA, ExoI, DNA Pol δ and DNA ligase, whose functions are DNA excision and resynthesis (Li, 2008).
Table 1.1 Prokaryotic MMR proteins and their eukaryotic homologs

<table>
<thead>
<tr>
<th>E. Coli Protein</th>
<th>MMR Function</th>
<th>Homologs in yeast/human</th>
<th>MMR Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutS</td>
<td>Recognizes mismatches in dsDNA</td>
<td>Msh2-Msh6 (MutSα)</td>
<td>Recognizes single base-base mismatches and 1-2 base IDLs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Msh2-Msh3 (MutSβ)</td>
<td>Recognizes some single base IDLs and IDLs ≥2 bases</td>
</tr>
<tr>
<td>MutL</td>
<td>“Matchmaker” that coordinates mismatch recognition with base excision initiation</td>
<td>Mlh1-Pms2 (MutLα)</td>
<td>Coordinates mismatch recognition with base excision initiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mlh1-Mlh2 (MutLβ) and Mlh1-Mlh3 (MutLγ)</td>
<td>Suppress some IDL mutagenesis, functions unclear</td>
</tr>
<tr>
<td>MutH</td>
<td>Nick daughter DNA strand to initiate base excision</td>
<td>None</td>
<td>—</td>
</tr>
<tr>
<td>γ-δ complex</td>
<td>Loads β-clamp onto DNA</td>
<td>RFC complex</td>
<td>Loads PCNA, modulates base excision</td>
</tr>
<tr>
<td>β-clamp</td>
<td>Interacts with MutS; may recruit MutS to mismatches and/or the replication fork</td>
<td>PCNA</td>
<td>Interacts with MutS and MutL; Recruits MMR proteins to mismatches; Participates in signaling, excision, and repair synthesis</td>
</tr>
<tr>
<td>UvrD helicase</td>
<td>Unwinds DNA to allow excision of daughter strand</td>
<td>None</td>
<td>—</td>
</tr>
<tr>
<td>ExoI, ExoX</td>
<td>Performs 5’-3’ excision of daughter strand</td>
<td>EXOI (Rth1)</td>
<td>Performs excision of dsDNA</td>
</tr>
<tr>
<td>RecJ, ExoVII</td>
<td>Performs 3’-5’ excision of daughter DNA strand</td>
<td>3’ exonuclease of Pol δ; 3’ exonuclease of Pol ε</td>
<td>Performs excision of daughter DNA strand</td>
</tr>
<tr>
<td>SSB</td>
<td>Stabilizes single stranded repair intermediate</td>
<td>RPA</td>
<td>Stabilizes single stranded repair intermediate</td>
</tr>
<tr>
<td>DNA Pol III</td>
<td>Resynthesizes daughter DNA strand</td>
<td>DNA Pol δ</td>
<td>Resynthesizes daughter DNA strand</td>
</tr>
<tr>
<td>DNA ligase</td>
<td>Seals nick after DNA resynthesis</td>
<td>DNA ligase</td>
<td>Seals nick after DNA resynthesis</td>
</tr>
</tbody>
</table>

Adapted from Kunkel and Erie (2005).
Prokaryotic MMR mechanism

In *E. coli*, MMR process is a methyl-directed pathway that has been extensively studied (Kunkel and Erie, 2005; Schofield et al., 2001b). All proteins involved in this pathway have been purified and the whole MMR system has been reconstituted *in vitro* in the Modrich laboratory at the Duke University Medical Center (Kunkel and Erie, 2005; Schofield et al., 2001b). Figure 1.1 shows a current model of prokaryotic MMR.

The first step of MMR in *E. coli* is MutS searching and binding to mismatched DNA strands (Grilley et al., 1989). MutS has DNA binding activity and ATPase activity. *E. coli* MutS can recognize G:T, A:C, G:A, T:C, A:A, G:G and T:T mismatches and also G:A, C:T mismatches with weaker affinity (Dohet, 1985; Kramer et al., 1984; Su et al., 1988). In addition MutS can recognize 1-4-base IDLs (Dohet, 1985; Parker and Marinus, 1992).

The second step is the recruitment of the matchmaker MutL by MutS to the mismatch site, which forms a complex with MutS. Then the MutS•MutL complex can stimulate the endonuclease activity of the third MMR protein MutH to make a nick at a GATC recognition sequence (Au et al., 1992; Modrich and Lahue, 1996). This nick can be made either upstream or downstream from the mismatch site. The N6 position of adenine in palindromic GATC sequences is methylated only on the parent strand during DNA replication (Junop et al., 2001). The nick made by MutH is specifically on the unmethylated, mismatch-containing daughter strand of a hemimethylated GATC while not on the methylated parent strand (Junop et al., 2001; Modrich and
Lahue, 1996). These hemi-methylated GATC sites determine the repair strand specificity.

After MutH makes the nick, UvrD helicase is recruited to unwind the DNA. Then four exonucleases (ExoI, ExoV, RecJ and ExoVII) excise the mismatched daughter strand to about a hundred bases past the mismatch site (Cooper et al., 1993; Grilley et al., 1993). The single-stranded region generated can be bound quickly by SSB (Cooper et al., 1993; Grilley et al., 1993). Finally, DNA polymerase III resynthesizes the missing daughter strand and DNA ligase can seal the nick to finish MMR (Iyer et al., 2006; Kunkel and Erie, 2005; Li, 2008).

This MMR pathway in *E. coli* has three main features (Li, 2008). First, MMR is strand-specific therefore it can only cut the newly synthesized daughter strand. Second, MMR is a bi-directional process. The nick can be either upstream or downstream from the mismatch site. Excision can be either from 5’ to 3’ or from 3’ to 5’ to the mismatch site. Third, MMR can distinguish a broad spectrum of substrate mismatches, not only base-base mismatches but also short IDLs. Functional MutS, MutL and MutH are required to achieve these three properties of MMR. Since the MMR process is highly conserved throughout evolution, the *E. coli* MMR mechanism is an excellent model for both prokaryotes and eukaryotes. However, there are striking differences between *E. coli* and many other bacteria and archaea. They do not methylate their daughter strands during DNA replication and do not have *E. coli* MutH homologs (Kunkel and Erie, 2005; Modrich, 1989). Thus their MMR mechanism must be different from the *E. coli* mechanism.
**Figure 1.1. Prokaryotic MMR model.** A) The MutS protein recognizes and binds to a mismatch in DNA. B) MutS recruits MutL, which activates the MutH endonuclease activity; MutH makes a nick on the error-containing daughter DNA strand at a hemimethylated GATC sequence. C) MutS and MutL proteins also recruit the UvrD helicase, which unwinds the DNA from the nick site, allowing exonuclease to remove the error-containing DNA strand to approximately 100 base pairs past the mismatch. D) Tethered to the DNA by the β-clamp, DNA polymerase resynthesizes the daughter strand, and DNA ligase seals the nick to complete MMR. (Adapted from Jacobs-Palmer 2006)
Figure 1.1.

A) Mismatch recognition

B) Communication between mismatch site and nick site

C) DNA excision

D) DNA resynthesis
Eukaryotic MMR

The eukaryotic MMR process is much more complicated than in prokaryotes although it has several similarities such as bidirectionality and nick-directed specificity. Major difference is that there are no hemi-methylated GATC sites in eukaryotes, compared with prokaryotes (Kunkel and Erie, 2005; Li, 2008). Figure 1.2 shows a eukaryotic MMR model.

First, Msh2-Msh6 searches for the mismatches on DNA and binds to the mismatch sites to initiate the MMR process (Kunkel and Erie, 2005; Li, 2008). In yeast, Msh2-Msh6 mainly binds base-base mismatches and 1-2 IDLs, while Msh2-Msh3 recognizes larger IDLs (Marsischky et al., 1996). For both of them, ATPase activity is very important for mismatch recognition and initiation of repair (Antony and Hingorani, 2003). For human MutS homologs, another complex, hMsh4-hMsh5, was found to be involved in MMR during meiotic DNA recombination. It is known to recognize Holiday junctions (Hollingsworth et al., 1995; Novak et al., 2001; Snowden et al., 2004).

After Msh2-Msh6 binds to DNA, the MutL homolog, MutLα (Mlh1-Pms1 in yeast) is recruited to the system (Williamson et al., 1985). Among all the MutL homologs, MutLα is essential for MMR, MutLγ is involved in meiosis, and the function of MutLβ is still unknown (Kunkel and Erie, 2005). MutS and MutL homologs can form a MutS•MutL•DNA ternary complex based on in vitro findings and MutL is a matchmaker in MMR that coordinates mismatch recognition and excision initiation (Drotschmann et al., 2002; Hall et al., 2003; Hall et al., 2001).
Recent study indicated that PCNA may be involved in the localization of MutS to mismatched strands (Shell et al., 2007). The role of MutL homologs is still not clearly defined, but they are essential components for the MMR process.

The excision initiation of the mismatched daughter strand in eukaryotic MMR is not fully understood because there is no MutH homolog in eukaryotes. One dominant hypothesis is that MMR proteins bind to PCNA and use it to locate repair to the newly synthesized daughter strand. Recent studies revealed that hMutLα has PCNA/RFC-dependent endonuclease activity which is involved in 3’ nick-directed MMR (Kadyrov et al., 2007). In the proteins Msh3 and Msh6, a conserved PCNA interaction motif called the PIP box is found that indeed interacts with PCNA (Clark et al., 2000). PCNA is a circular clamp and helps DNA polymerase bind onto DNA strands during DNA replication. The Msh proteins may complete PCNA interaction with DNA polymerase, resulting in a release of the DNA polymerase from the replication fork. This provides a possible entry point for exonucleases. Several in vivo experiments showed that nicks and gaps were generated as a result of discontinuous DNA replication (Lopez de Saro and O'Donnell, 2001; Pavlov et al., 2003; Umar et al., 1996). ExoI is a prominent 5’ to 3’ eukaryotic exonuclease in MMR and needs MutS and RPA to start its 5’ directed mismatch excision (Genschel and Modrich, 2003). In mice and yeast, exoI null mutants gave a very weak mutator phenotype, which indicates that there are still unidentified exonucleases involved in eukaryotic MMR pathway (Genschel and Modrich, 2003). Following DNA excision, a DNA daughter strand is newly resynthesized by DNA polymerase δ with the help of RPA.
Then DNA ligase seals the nick (Hsieh, 2001; Kunkel and Erie, 2005; Modrich, 1991).
**Figure 1.2. Eukaryotic MMR model.** A) The Msh2-Msh6 heterodimer recognizes and binds to a mismatch in DNA. B) The Mlh1-Pms1 heterodimer facilitates communication between Msh2-Msh6 and PCNA and DNA polymerase at the replication fork or with other proteins at a persistent strand break in DNA. C) Exonuclease removes the error-containing daughter DNA strand back past the mismatch. D) Tethered to the DNA by PCNA, polymerase reincorporates the missing bases and DNA ligase seals the nick to complete MMR. (Adapted from Jacobs-Palmer 2006)
Figure 1.2.

A) Mismatch recognition

B) Communication between mismatch site and PCNA

C) DNA excision

D) DNA resynthesis
**Structures of MutS and Msh2-Msh6**

As mentioned above, the MMR mechanism is conserved between prokaryotes and eukaryotes. The first key event in MMR is mismatch recognition. How MutS recognizes mismatch sites in DNA and signals initiation of MMR is still unsolved. The crystal structure studies of MutS and Msh2-Msh6 could improve our understanding of this process.

Crystal structures have been generated for *Thermus aquaticus* (*T. aquaticus*) MutS with +T DNA (Obmolova et al., 2000) and *E.coli* MutS with G:T, G:G, C:A, A:A and +T DNAs (Lamers et al., 2000; Natrajan et al., 2003). Recently, the crystal structure of the human Msh2-Msh6 heterodimer with DNA has also been established (Warren et al., 2007). Not surprisingly, these structures are very similar.

When MutS or Msh2-Msh6 binds to a mismatched DNA strand, it forms a clamp-like structure around the mismatch site (Figures 1.3 and 1.4). Each subunit of MutS or Msh2-Msh6 has five domains, which are the mismatch binding domain, connector domain, levels domain, clamps domain and ATPase domain, respectively (Warren et al., 2007).
Figure 1.3. Crystal structure of *Thermus aquaticus* MutS. (A) Panel A shows two views of the MutS homodimer, colored red and green respectively, bound to mismatched DNA (yellow). (B) Panel B shows the N-terminal DNA binding region where Phe36 (blue) from one of the two subunits stacks against the mismatched base, and the DNA is kinked by 60 degrees. (Adapted from Antony 2005)
Figure 1.3.

A) Crystal Structure of *Thermus aquaticus* MutS

B) Mismatch DNA recognition
Figure 1.4. Crystal structures of *T. aquaticus* MutS and *H. sapiens* MutSα (Msh2-Msh6).
Figure 1.4.

*T. aquaticus MutS* | *H. sapiens MutSα*

When MutS or Msh2-Msh6 binds mismatched DNA, it makes sequence-independent hydrogen bonding and van der Waals contacts in the vicinity of the mismatched DNA. The DNA is kinked toward the major groove by about 45-60 degrees (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). The striking similarities among these structures suggest that the kinked DNA complex is a significant species in the pathway leading to MMR. Both subunits of MutS/Msh2-Msh6 have mismatch-binding domain, but the binding to mismatch sites is asymmetric (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). Only Msh6 and one monomer of MutS have specific and direct contacts with the mismatched bases by a Phe-X-Glu motif that is highly conserved from prokaryotes to eukaryotes (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). The phenylalanine (Phe) residue from Phe-X-Glu motif stacks with a mismatched or inserted base and the glutamate (Glu) residue forms a hydrogen bond with the N-3 of a mismatched thymine or the N-7 of mismatched purine (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). One big difference between *T. aquaticus* MutS and human Msh2-Msh6 in this domain is that the non-mismatch binding monomer of *T. aquaticus* MutS has sequence-independent contacts with the DNA backbone and contributes to DNA bending (Obmolova et al., 2000). However, the human Msh2 mismatch binding domain gets away from the DNA backbone by rotating up, and there is only one contact between Msh2 and DNA (Warren et al., 2007). Hence, in the case of human Msh2-Msh6, Msh6 solely contributes to DNA bending (Warren et al., 2007).
The ATPase domain is the most highly conserved region. 48% of the yeast ATPase domain is identical to the *E. coli* hMSH2 (Warren et al., 2007). MutS protein belongs to the ATP binding cassette (ABC) superfamily of ATPase which contains highly conserved Walker A motif for ATP binding and Walker B motif for ATP hydrolysis (Hopfner and Tainer, 2003). MutS protein has two ATPase sites that are asymmetrical for ATP binding activity and have different affinities for nucleotides (Antony and Hingorani, 2003; Martik et al., 2004). The mismatch binding domain and ATPase domain are located far from each other according to the crystal structure but there is allostery between them.

Even though crystal structures contain a lot of useful information, it cannot tell us the protein’s dynamic action in time. Combined with kinetic studies, we can understand more detailed molecular mechanism of MMR.

**Current model of MMR mechanism**

Crystal structures of MMR proteins gave us amazing information about MMR mechanism, but this is only the first step. The ultimate goal is to fully understand each step in MMR so that we can answer the following questions such as how MutS recognizes different mismatches and IDLs, how MutS moves and searches mismatches and how MutS ATPase activity coordinates with DNA recognition activity.

After 30 years of research, several models were developed to describe the mechanism of MMR. From Figure 1.5, we can see that there are two main models
for signaling between the mismatched site and a strand-specific nick (the strand discrimination signal): one is the “cis-” or “moving” model and the other is the “trans-” or “stationary” model.

The “stationary” model (Figure 1.5, right) shows that MutS or Msh2-Msh6 protein recognizes a mismatch site and remains bound to the mismatch site in a nucleotide-free state. Upon ATP binding, MutS recruits MutL or MutL homolog to the mismatch site and forms a complex with MutL. DNA bending or looping can bring the mismatch site and the strand discrimination signal site close together. In this model, ATPase activity has a proofreading role (Junop et al., 2001). The stationary model is supported by experiments showing that MutS and MutL could stimulate MutH endonuclease activity even when the mismatch site and the GATC site were in two separate DNA strands (Junop et al., 2001), and that excision initiation could be generated even in the presence of a biotin-streptavidin blockage between the mismatch site and excision initiation site (Wang and Hays, 2004).

In the “cis-” or “moving” model (Figure 1.5, left), MutS binds to a mismatch site and moves away to search the nicked strand. There are two “moving” models: one is the “translocation” model, the other is the “molecular switch” or “sliding clamp” model. In the translocation model, MutS searches and binds a mismatch site in a nucleotide-free state. ATP binding reduces the binding affinity between the MutS and mismatch site and ATP hydrolysis is the cause of unidirectional translocation of MutS along DNA until it reaches the nick to form an α-like loop. This model is supported by electron microcopy study showing the loops formed at a
MutS binding site in the presence of ATP (Allen et al., 1997). In the molecular switch model or sliding clamp model, MutS searches for mismatch sites in the ADP-bound form and binds to a mismatch site to trigger exchange of ADP to ATP. This nucleotide exchange can cause bidirectional sliding of MutS away from the mismatch site to the nick site to start excision. In this model, ATP binding is a key to signal downstream events. In SPR experiments, MutS had a much faster dissociation rate from +T DNA in the presence of ATP than in the absence of nucleotide or in the presence of ADP (Selmane et al., 2003).

To date, currently available data cannot exclude any of these three models of the MMR mechanism. We still need to do more research to figure out how the mechanism of MMR truly works.
Figure 1.5. MMR mechanism models.
Figure 1.5.

Adapted from Li (2008)
Outline of my thesis

The main question of my thesis is how Msh2-Msh6 targets specific sites in DNA. In order to measure transient Msh2-Msh6-DNA interactions at a higher resolution, both with respect to the location and timing of the binding events, we developed stopped-flow assays with an on-site fluorescence reporter to detect Msh2-Msh6 arrival at and departure from particular sites on DNA.

In this research, comparison of the kinetic parameters revealed striking differences in Msh2-Msh6 interactions with various sites, which will be described in detail in the following chapters. The key findings in my thesis research are:

1. The reaction kinetics revealed that Msh2-Msh6 binds a variety of sites at similarly fast rates ($k_{ON} \sim 10^7$ M$^{-1}$ s$^{-1}$).

2. Msh2-Msh6 selectivity manifests in differential dissociation rates; e.g., the protein releases a 2-Aminopurine:T base pair ~90-fold faster than a G:T mismatch.

3. On releasing the weak 2-Ap:T site, Msh2-Msh6 is able to move laterally on DNA to locate a nearby G:T site.


repair as effectively as Msh2-Msh6 signals G:T repair. These findings suggest that Msh2-Msh6 employs a targeting mechanism whereby it interrogates base pairs by rapid, transient contacts and pauses at potential MMR sites, and the longer the pause the greater the likelihood of MMR.

My results illuminated the process by which Msh2-Msh6 targets mismatched base pairs and other defects in DNA, and initiates MMR.
Chapter II

Materials and methods

DNA

*The 23mer and 37mer DNA strands*

Synthetic single stranded labeled with 2-Ap (2-Aminopurine) and TAMRA (5-(6)-carboxytetramethylrhodamine) and Cy3 (cyanine 3), and corresponding unlabeled 23mer and 37mer DNAs were purchased from Integrated DNA Technologies (IDT), Inc., and purified by denaturing polyacrylamide gel electrophoresis (PAGE), electroelution, and ethanol precipitation. All 23mer 2-Ap labeled and TAMRA labeled DNA sequences used for *T. aquaticus* MutS DNA binding and ATPase activities are listed in Table 2.1. All 37mer 2-Ap labeled DNA sequences used for *S. cerevisiae* Msh2-Msh6 DNA binding and ATPase activities are listed in Tables 2.2 and 2.3. All 37mer TAMRA and Cy3 labeled DNA sequences used for *S. cerevisiae* Msh2-Msh6 DNA binding and ATPase activities are listed in Tables 2.3 and 2.4.
Table 2.1. 23mer 2-Ap and TAMRA labeled DNA substrates.

<table>
<thead>
<tr>
<th>2Ap+T mismatch</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>23merNoAp+T</td>
<td>5'-CC GGC AGC TAT _TA CCG TCG CGC GG CCG TCG ATA TAT GGC AGC GCG-5'</td>
<td>EJP22ATA EJP23ATA-NoAp+T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2ApA:T match</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>2ApG:T mismatch</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>23merNoApG:T</td>
<td>5'-CC GGC AGC TAT GTA CCG TCG CGC GG CCG TCG ATA TAT GGC AGC GCG-5'</td>
<td>EJP22ATA-G-base EJP23ATA-NoAp+T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TAMRA DNA</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>23TAMRA +T</td>
<td>5'-CC GGC AGC TAT TA CCG TCG CGC AmM-GG CCG TCG ATA TAT GGC AGC GCG-5'</td>
<td>EJP22ATA EJP23TAMRA+T</td>
</tr>
<tr>
<td>23TAMRA A:T</td>
<td>5'-CC GGC AGC TAT ATA CCG TCG CGC AmM-GG CCG TCG ATA TAT GGC AGC GCG-5'</td>
<td>EJP23ATA-A-base EJP23TAMRA+T</td>
</tr>
</tbody>
</table>
Table 2.2. 37mer 2-Ap labeled DNA substrates.

<table>
<thead>
<tr>
<th>3Ap+T mismatch</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>2Ap:T/+T</td>
<td>5’-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATAp TAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP36 EJP37Ap+T</td>
</tr>
<tr>
<td>NoAp/+T</td>
<td>5’-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATAp TAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP36 EJP37+T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3ApG:C match</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoAp/G:C</td>
<td>5’-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATAp CAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP37-Gbase EJP37+C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3ApG:T match</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoAp/G:T</td>
<td>5’-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATAp TAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP37-Gbase EJP37+C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>37(2Ap:T) G26:T</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoAp/7bp/G:T</td>
<td>5’-ATT TCC TTC AGC AGA TAT GTA CCA TGC TGA TTC ACA T TAA AGG AAG TCG TCT ATAp CAT GGT ATG ACT AAG TGT A-5’</td>
<td>JZ37GbaseG26 EJP37+C</td>
</tr>
</tbody>
</table>
Table 2.3. 37mer 2-Ap and TAMRA labeled DNA substrates.

<table>
<thead>
<tr>
<th>37(2Ap:C) G:C</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-ATT TCC TTC AGC AGA TAC GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA</td>
<td>JZ37GbaseC18</td>
</tr>
<tr>
<td></td>
<td>5’-ATT TCC TTC AGC AGA TAC GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP37+Ap+C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No(2Ap:C) G:C</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-ATT TCC TTC AGC AGA TAC GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA</td>
<td>JZ37GbaseC18</td>
</tr>
<tr>
<td></td>
<td>5’-ATT TCC TTC AGC AGA TAC GTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP37+Ap+C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>37O’MeG 2-Ap strands</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’MeG ApG:C</td>
<td>5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA</td>
<td>JZ37O’MeGbase</td>
</tr>
<tr>
<td></td>
<td>CAT GGT ATG ACT AAG TGT A-5’</td>
<td>EJP37Ap+C</td>
</tr>
</tbody>
</table>

| O’MeG NoApG:C        | 5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA | JZ37O’MeGbase |
|                      | CAT GGT ATG ACT AAG TGT A-5’ | EJP37+Ap+C |

| O’MeG ApG:T          | 5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA | JZ37O’MeGbase |
|                      | CAT GGT ATG ACT AAG TGT A-5’ | EJP37+Ap+T |

| O’MeG NoApG:T        | 5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA | JZ37O’MeGbase |
|                      | CAT GGT ATG ACT AAG TGT A-5’ | EJP37+T |

<table>
<thead>
<tr>
<th>37O’MeG TAMRA strands</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’MeG:C</td>
<td>5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A</td>
<td>JZ37O’MeGbase</td>
</tr>
<tr>
<td>AmM-TAA AGG AAG TCG TCT ATA</td>
<td>37O’MeGbase</td>
<td>JZ37TAMRA+C</td>
</tr>
</tbody>
</table>

| O’MeG:T               | 5’-ATT TCC TTC AGC AGA TAT O’MeGTA CCA TAC TGA TTC ACA T TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A | JZ37O’MeGbase |
| AmM-TAA AGG AAG TCG TCT ATA | 37O’MeGbase | JZ37TAMRA+T |

<table>
<thead>
<tr>
<th>37mer TAMRA strands</th>
<th>DNA Sequence</th>
<th>Individual Strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>G:C</td>
<td>5’-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T AmM-TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5’</td>
<td>JZ37TAMRA+C</td>
</tr>
<tr>
<td>G:T</td>
<td>5’-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T AmM-TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5’</td>
<td>JZ37TAMRA+T</td>
</tr>
</tbody>
</table>
Table 2.4. 37mer TAMRA and Cy3 labeled DNA substrates.

<table>
<thead>
<tr>
<th>37mer TAMRA strands</th>
<th>DNA Sequence</th>
<th>Individual strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>+T&lt;sub&gt;TAMRA&lt;/sub&gt;</td>
<td>5'-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T</td>
<td>EJP36 JZ37TAMRA+T</td>
</tr>
<tr>
<td></td>
<td>AmM-TAA AGG AAG TCG TCT ATA TAT GGT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
<tr>
<td>+2T&lt;sub&gt;TAMRA&lt;/sub&gt;</td>
<td>5'-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T AmM-TAA AGG AAG TCG TCT ATA</td>
<td>EJP36 JZ37TAMRA+2T</td>
</tr>
<tr>
<td></td>
<td>TTAT GGT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
<tr>
<td>+3T&lt;sub&gt;TAMRA&lt;/sub&gt;</td>
<td>5'-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T AmM-TAA AGG AAG TCG TCT</td>
<td>EJP36 JZ37TAMRA+3T</td>
</tr>
<tr>
<td></td>
<td>ATA TTTAT GGT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
<tr>
<td>+4T&lt;sub&gt;TAMRA&lt;/sub&gt;</td>
<td>5'-ATT TCC TTC AGC AGA TAT TA CCA TAC TGA TTC ACA T AmM-TAA AGG AAG TCG TCT</td>
<td>EJP36 JZ37TAMRA+4T</td>
</tr>
<tr>
<td></td>
<td>ATA TTTTTAT GGT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
<tr>
<td>2Ap:T/7bp</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TGC TGA TTC ACA T-AmM TAA AGG AAG</td>
<td>JZTAMRA37GbaseG26 JZ37Ap+C26</td>
</tr>
<tr>
<td>/G26:C</td>
<td>GCT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
<tr>
<td>2Ap:T/7bp</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TGC TGA TTC ACA T-AmM TAA AGG AAG TCG</td>
<td>JZTAMRA37GbaseG26 EJP37Ap+C</td>
</tr>
<tr>
<td>/G26:T</td>
<td>TCT ATAp CAT GGT ATG ACT AAG TGT A-5'</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>37mer Cy3 strands</th>
<th>DNA Sequence</th>
<th>Individual strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoApG:C /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37+C</td>
</tr>
<tr>
<td>Ap:T/G:C /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATAp CAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37Ap+C</td>
</tr>
<tr>
<td>NoApG:T /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATA TAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37+T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>37mer TAMRA strands</th>
<th>DNA Sequence</th>
<th>Individual strands</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoApG:C /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATA CAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37+C</td>
</tr>
<tr>
<td>Ap:T/G:C /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATAp CAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37Ap+C</td>
</tr>
<tr>
<td>NoApG:T /Cy3</td>
<td>5'-ATT TCC TTC AGC AGA TAT GTA CCA TAC TGA TTC ACA T-Cy3 TAA AGG AAG TCG TCT ATA TAT GGT ATG ACT AAG TGT A-5'</td>
<td>Jie37Gbase/Cy3 EJP37+T</td>
</tr>
</tbody>
</table>
**PAGE purification of DNA**

**Preparation of 40% (38:2) acrylamide:bis:** First, wear a lab coat, mask, goggles and gloves, because acrylamide powder is neurotoxic. Then, put a stirring bar in a beaker and carefully weigh out 190 g acrylamide and 10 g bisacrylamide in the beaker. Pour ddH$_2$O in the beaker until 90% of final volume (500 ml) slowly and carefully. Plastic wrap the top of the beaker and stir until dissolved. Transfer solution to a graduated cylinder and fill up to 500 ml with ddH$_2$O. Mix well and transfer to a foil-covered bottle and store at 4 ºC.

**Preparation of urea gel:** For DNAs smaller than 25 bp, use a 15% gel. For DNAs around 35-40 bp, use a 19% gel. Prepare 50 ml gel solution per gel (1.5 mm thickness, 16.5 cm (w) x 22 cm (l) in a side-armed flask. For 15% gel solution, add 21 g urea, 18.75 ml 40% (38:2) acrylamide:bis, 5 ml 10 x TBE, 8.6 ml ddH$_2$O. For 19% gel solution, add 21 g urea, 23.75 ml 40% (38:2) acrylamide:bis, 5 ml 10 x TBE, 3.6 ml ddH$_2$O.

**Casting the gel:** After making the gel solution, boil at 60 ºC for about 2 -3 min in the kettle, with occasional swirling, until the urea dissolves. Cool for about 20 min to the room temperature and degas the solution (close the mouth of the flask with a rubber stopper and hook it to a vacuum source), for 10 min until you see no more air bubbles rising through the solution, swirl gently every 2 min. Cast the gels according to the manufacturers instructions (apply Vaseline to the spacers for a tighter seal). To the degassed gel solution, add 125 µl 10% APS and 30 µl TEMED, and quickly pour it
into the gel apparatus (~ within 30-40 seconds). Insert the combs and allow the gel to polymerize. Once polymerization is complete, wait at least 1 hour to use. When you are ready, flush the well with 1X TBE and pre-run the gel in 1X TBE at 400 V for 1 hour at RT.

**Sample preparation:** Resuspend the oligonucleotide in 150 µl 1 x DNA loading buffer (1ml 1 x DNA loading buffer: 450 µl 4.5% (v/v) formamide, 100 µl 10 x TBE, 450 µl ddH₂O), vortex well. Reflush the well of the gel with 1 x TBE after the pre-run and gently load the oligonucleotide sample on the gel. Add 100 µl 1 x DNA loading buffer to wash the oligonucleotide sample tube and vortex and load. Similarly, make a mock reaction with ddH₂O and formamide with a speck of bromphenol blue dye (BB migrates at ~10 base) and Coomassie brilliant blue R-250 (CBB migrates at ~40 base). Load them as markers in separate wells. Run the gel at 450 V at RT for 4-5 hours until the dye reaches about 2/3\(^{th}\) of the gel. Remove the gel, place it on a silica plate, using a hand-held UV-light trace the oligonucleotide band on the gel and excise it. Cut the gel into smaller pieces and proceed with setup of the elution chamber to extract the oligonucleotide from the gel.

**Oligonucleotide Elution:** Setup the oligonucleotide elution chamber according to the manufacturer’s instructions (Schleicher & Schuell, ELUTRAP electro-separation system). The two non-porous BT1 membranes should be inserted on the outside, and the porous BT2 membrane should be used to divide the chambers. Keep the members moist while preparing the chamber. Fill the chambers with 1 x TBE and load the gel pieces. Separate the oligonucleotides from the acrylamide gel pieces at 100 V
overnight at 4 °C. Pick the eluted oligonucleotide using a long gel loading tip and repeat the elution process for another 4 hours.

**Ethanol Precipitation of Oligonucleotides:** To the eluted oligonucleotide, add 100% ethanol (3 x volume of eluate) and 3 M sodium acetate (3 M sodium acetate (NaAc): 12.3 g NaAc add ddH₂O to final 50 ml, adjust pH with acetic acid (HAc)), pH 5.2 (0.1 x volume of eluate). Remember to do the ethanol precipitation in eppendorf tubes marked ‘not-low adhesion’, for the pellet to adhere to the tube. Mix and leave the tubes overnight at -80 °C, and centrifuge the samples at 14 K rpm on a table-top centrifuge at 4 °C for 30 min. Discard the supernatant, air dry the tubes until all the ethanol has evaporated (15~20 min). Resuspend the pellet in the desired volume of 20 mM Tris-HCl, pH 8.0 (usually 0.5~1 µl for 1 nmole DNA synthesis). Measure oligonucleotides final concentration. Blank 1 ml ddH₂O at 260 nm in a quartz cuvette. Mix in 1 µl DNA sample and measure absorbance. Repeat at least twice and the difference should be less than 5%. Take the average of the reading and calculate molar concentration = (A₂₆₀average x dilution factor)/ ε (DNA extinction coefficient). Store the oligonucleotides as smaller aliquots at -80 °C. The yield is above 60%.

**TAMRA labeling of DNA**

**PAGE purification of DNA:** Use PAGE to purify DNA modified with 3’NH₂ group and get the final concentration > 20 µg/µl. DNA concentration (µg/µl) = (A₂₆₀ x 1000 µl/1 µl) / extinction coefficient (µl / µmole•cm) x MW.
From next step, always keep TAMRA-labeled DNA in the dark (i.e., turn off lights, close doors and use aluminum foil).

Label DNA with TAMRA: First, prepare 0.1 M sodium tetraborate buffer (pH 8.5) just before use (add 0.1520 g Borax powder to 4 ml ddH₂O in a 15 ml falcon tube and lower pH to 8.5 by adding ~220 µl 1 M HCl). Weigh ~800 µg TAMRA dye in amber tube and add 80 µl DMSO (dimethyl sulfoxide) and vortex well. Then add 400 µg DNA (20 µl 20 µg/µl DNA) and 300 µl 0.1 M sodium tetraborate to total 400µl per reaction. This reaction ratio is for 18~24mer DNA. Parafilm the cap, cover the tube with foil and place on the roller drum at room temperature overnight. In the dark, dissolve 800 µg TAMRA in dimethyl sulfoxide and react with 400 µg DNA in 0.1 M sodium tetraborate buffer pH 8.5 for 16 hours at 25 ºC (Jacobs-Palmer and Hingorani, 2007).

Butanol extraction: Prepare water-saturated butanol: mix 5 ml ddH₂O with 5 ml 2-butanol in a 15 ml falcon tube. Vortex at max for 30 second and centrifuge at speed setting #4 for 2 minutes in the cold room. Mix again and spin down again. Water saturated butanol is the top layer. Split 400 µl DNA reaction mixture into two 1.5 ml clear tubes (each 200 µl). Add 600 µl water saturated butanol to each tube, vortex and spin down at 13,000 rpm for 2 minutes at room temperature. Remove top layer that contains unreacted TAMRA dye in butanol to a 15 ml falcon tube with a 1 ml pipette. Repeat twice. It is not necessary to remove all of top layer during first two extractions. Be careful not to remove any of the bottom layer, which contain DNA. Save the top layer in a 15 ml falcon tube and discard later. For the last extraction,
remove as much of the butanol as possible with a 200 µl pipette when it is close to the bottom layer. Transfer bottom layer to two 1.5 ml clear tubes, poke one small hole in each tube cap. Dry samples in a speed vacuum the samples for 1 hour without heat until the butanol is gone (no smell). Don’t forget to cover the speed vacuum with foil. Be careful that the speed vacuum heat is off and the condensator and vacuum pump are on. When you start the vacuum, turn on the rotor first and then the vacuum. Always release vacuum first, then turn off the rotor. Don’t shake the sample tubes because of the hole on the cap.

**P6 gel-filtration chromatography:** Mix ~1 tsp of P6 gel filtration resin in 50 ml ddH₂O and keep in cold room overnight. Prepare a 5 ml P6 column in a plastic and disposable 10 ml BioRad column and equilibrate with 50 ml buffer P6A (buffer P6A: 10 mM Tris-HCl pH8.0 and 0.1 mM EDTA). When buffer P6A level drops to the gel surface, gently and evenly layer on ~400 µl DNA\text{\textsubscript{TAMRA}} solution by using a 200 µl pipette. Allow DNA\text{\textsubscript{TAMRA}} solution to completely enter resin and evenly layer 1 ml buffer P6A multiple times on top of column with a 1 ml pipette. At the same time, collect fractions in 1.5 ml capless tubes, 8 drops (~300 µl) per tube, until the first pink peak that contains TAMRA labeled DNA is done and a more intense second peak that contains free TAMRA is coming through. Total is about 12~15 tubes. Measure absorbance of all first peak fractions at $A_{260}$ and $A_{555}$ (dilute 1 µl in 1 ml ddH₂O) in quartz cuvettes. Pool 5~7 tubes of peak fractions that $A_{260} > 3 \times A_{555}$ (total volume is 1~1.5 ml). Split into “not low adhesion” tubes at ~300 µl (400 µl is the maximum).
Add 3 x volume 100 % ethanol and 0.1 x volume 3M NaAc pH5.2 for ethanol precipitation in -80 °C overnight. Keep column in the hood and dispose later.

**PAGE purification of DNA**<sub>TAMRA</sub>: Spin down DNA<sub>TAMRA</sub> at 13,000 rpm for 30 minutes at cold room, remove supernatant and dry in the hood. Use 0.75 mm thickness gel set to run PAGE gel to further purify DNA<sub>TAMRA</sub>. For DNAs smaller than 25 bp, use a 15% gel. For DNAs around 35-40 bp, use a 19% gel. Visualize band using a handheld UV lamp. Excise the fluorescent pink DNA<sub>TAMRA</sub> band and be careful not to cut the unlabeled DNA bands. Cut the pink band into small pieces and distribute them into four 1.5 ml tubes. Do not expose DNA<sub>TAMRA</sub> to UV light for a prolonged period time. Add 400 µl elution buffer (50 mM Tris-HCl pH7.5, 50 mM NaCl, 1mM EDTA) to each tube. Parafilm tube caps, cover tubes with foil and place on roller drum overnight at room temperature. Centrifuge tubes at 13,000 rpm for 5 minutes. Collect supernatant using “not low adhesion” 1.5 ml tubes with gel loading tip as first pick. Add another 400 µl elution buffer and repeat shaking on the roller drum for 4~5 hours and collect the supernatants as second pick.

**Concentrate DNA**<sub>TAMRA</sub> and measure DNA<sub>TAMRA</sub> concentrations: Use the speed vacuum (no heat and use foil) on first and second pick DNA<sub>TAMRA</sub> for 3~4 hours to reduce the volume to about half. Combine the solutions and use “not low adhesion” 1.5 ml tubes to ethanol precipitate overnight. Spin down DNA<sub>TAMRA</sub> at 13,000 rpm for 30 minutes at 4 °C and remove supernatant and dry in the hood. Resuspend the pellet in 1x TE buffer (10 mM Tris-HCl pH7.5 and 1mM EDTA) for a final concentration of ~300 µM. Assume 80% DNA retention from P6 column. Check final
DNA<sub>TAMRA</sub> concentration. Dilute 1 µl DNA<sub>TAMRA</sub> in 1 ml ddH<sub>2</sub>O in quartz cuvette and take A<sub>260</sub> measurement. Repeat at least 3 times and take the average reading. For DNA<sub>TAMRA</sub>, E' = E<sub>DNA</sub> + 0.34 x E<sub>TAMRA-555</sub>, E<sub>DNA</sub> is the DNA extinction coefficient and E<sub>TAMRA-555</sub> is the extinction coefficient of TAMRA at 555 nm, which is 65,000 L/mol⋅cm. For final DNA<sub>TAMRA</sub> concentrations, DNA<sub>TAMRA</sub> molar concentration = (A<sub>260average</sub> x dilution factor) / E'. Store the DNA<sub>TAMRA</sub> as smaller aliquots (usually 30 µl) at -80 °C. The yield is above 50%.

**Annealing of DNA**

The DNA strands were mixed in 1.15: 1 (unlabeled: labeled, 50~100 µM concentration in 20 mM Tris-HCl, pH 8, 100 mM NaCl), and heated to 95 °C for 1 minute in boiling water in the kettle followed by slow cooling over 6 – 8 hours in the kettle to 25 °C (non-denaturing PAGE analysis of the products revealed > 95 % duplex DNA).

**Nucleotides and other reagents**

ATP, ATPγS, and ADP were purchased from Sigma Chemical Co., as were Purine Nucleotide Phosphorylase (PNPase) and Phosphate standard solution. 7-Methylguanosine (7-MEG) was purchased from R. I. Chemical, Inc. BL21(DE3) E. coli cells for MutS overexpression were purchased from Stratagene, and BLR(DE3) E. coli cells for Msh2-Msh6 overexpression, from Novagen (EMD Biosciences). MutS:pET3b was a kind gift from Dr. Peggy Hsieh (N.I.H.). The pET11a and
pET16b vector DNAs were purchased from Novagen, and pLANT2/RIL vector was a gift from Dr. Michael O’Donnell, Rockefeller University. Q-sepharose and SP-sepharose resins for protein purification were purchased from Amersham Biosciences (GE Healthcare). Heparin resin was purchased from BioRad Laboratories.

**Protein overexpression and purification**

* *S. cerevisiae Msh2-Msh6 overexpression and purification*

**Plasmid:** Msh2:pET11a and Msh6:pLANT2b plasmids were cloned by Dr. Edwin Antony (Antony and Hingorani, 2003; Finkelstein et al., 2003) (Figure 2.1).

**Competent cell preparation:** To make Competent cell BL21(DE3), every step and everything should be done under sterile conditions, such as wear gloves, flame always on, use new autoclaved sterile tips and eppendorf tubes and centrifuge tubes. Grow 200 µl BL21(DE3) cells in 200 µl LB in 240 rpm at 37 °C for ~3 hours until reach an OD$_{600}$ of 0.4. Be careful not to over grow. Spin down at 2,500 rpm for 20 minutes at 4 °C. Discard supernatant and resuspend pellet in 24 ml 0.1 M CaCl$_2$ (filter sterilized). Keep on ice for 30 minutes and centrifuge at 2,500 rpm for 15 minutes in two 30 ml centrifuge tubes. Discard supernatant and resuspend pellet in 4 ml 0.1 M CaCl$_2$ with 15% glycerol solution by pipetting (filter sterilized). Keep on ice in the cold room overnight. Make an aliquot 100 µl into sterile eppendorf tubes and store at 80 °C.

**Transformation:** Add 1-3 µg of Msh2:pET11a and 1-3 µg of Msh6:pLANT2b plasmid DNA into 200 µl fresh BLR(DE3) competent cells, thawed on ice in a round bottom tube followed by standard heat shock transformation (30 minutes on ice, 90
seconds at 42 °C, 2 minutes on ice, add 200 µl LB media, incubate at 37 °C on shaker for an hour at 230 rpm). Fresh BLR(DE3) competent cell is very important for good transformation and efficient induction. Plate the transformed mixture onto a LB with ampicillin and kanamycin plate. When I plated, I did it gently and only 1-2 round circle was enough.
Figure 2.1. Overexpression and purification of *S. cerevisiae* Msh2-Msh6 from *E. coli*. (A) Plasmids used for overexpression. *Msh2* was inserted into pLANT2/RIL and *Msh6* into pET11a and coexpressed in *E. coli*. Alternately, *Msh2* and *Msh6* were cloned into a single pET11a vector and overexpressed in BLR(DE3) cells. (B) Purification of Msh2-Msh6 analyzed on a 10% SDS-PAGE. Lane 1, uninduced cells; lane 2, IPTG-induced cells; lane 3, cleared cell lysate; lane 4, SP-Sepharose eluate; lane 5, Heparin eluate; and lane 6, Q-Sepharose eluate. The molecular weights of Msh2 and Msh6 are 108.8 and 140.1 kDa respectively. (Adapted from Antony 2005)
Figure 2.1.

A) Plasmids used for Msh2-Msh6 overexpression in *E. coli*

B) Purification of *Saccharomyces cerevisiae* Msh2-Msh6
**Growing Biomass:** Autoclave 24 L LB media (12 x 2L in 4L flasks), cool the flasks to RT and add 2 ml of 100 mg/ml ampicillin and 2 ml of 50 mg/ml kanamycin (filter sterilized). Resuspend a colony in 850 µl of sterile LB and vortex the solution vigorously. Add 100 µl of the resuspended solution into each flask of total six flasks and repeat for another six flasks with another colony. Grow the culture at 37 °C on a shaker at ~210 rpm (note: the cultures takes approximately takes 10 - 11 hours to reach OD$_{600}$ 0.6 at 37 ºC). Induce the culture with 0.5 mM freshly dissolved IPTG (add 1 ml of 1 M IPTG stock solution to each flask) when the cultures reach an OD$_{600}$ 0.6 (remember to take a 1 ml uninduced sample for the induction gel). The cell density is very important for good induction and high yield, so OD$_{600}$ should be between 0.6 - 0.7. Grow the culture for an additional 3 hours after IPTG addition (take a 1 ml induced sample for the induction gel). Centrifuge the biomass at 2,500 rpm (Sorvall RC5 centrifuge) for 20 min at 4 ºC (use 500 ml centrifuge bottles and do not overfill bottles). Decant the supernatant and resuspend the pellet on ice with resuspension buffer (20 mM Tris-Cl, pH 8.0, 10% sucrose and ~ 120 ml for a 24 L culture). Transfer the resuspended biomass to 50 ml falcon tubes and store at -80 °C until further purification.

**Induction Gel:** Spin down the 1 ml uninduced and induced samples immediately after collection in a table-top centrifuge for 1 min. Decant the supernatant and store the pellet in -20 ºC. Resuspend the pellet with 100 µl ddH$_2$O. Add 40 µl of 5 x SDS gel loading dye and load 10 µl of the sample onto a 10 % SDS-PAGE gel.
**Cell Lysis for 24 L biomass:** Thaw the frozen biomass on ice (place tubes in a beaker filled with water placed on ice). Add 200 ml 2 M NaCl (1 M final concentration), 400 µl 1000 x protease inhibitor cocktail (Sigma Chemical Co.), 400 µl PMSF (freshly prepared by adding a small speck in 3 ml isopropanol), 160 mg lysozyme resuspended in 1 ml ddH$_2$O (0.4 mg/ml final concentration) to the thawed biomass at 4 ºC. Make the final volume to 400 ml with Buffer A (25 mM Tris-HCl pH 8.0, 1 mM EDTA, 5% glycerol, 1 x protease inhibitor, PMSF 5 ml / 2 L Buffer A) in the cold room. (PMSF: 10 ml isopropanol in 15 ml falcon tube and small spatula full of PMSF. PMSF is an irritant, so wear gloves and make up the solution in the hood.) Stir the cell mixture gently for 1 hour at 4 ºC (in the cold room). Transfer the biomass to 50 ml falcon tubes (~ 40 ml per falcon tube) and rapid freeze the tubes in liquid nitrogen for 2 minutes, followed by incubation (thawing) at 37 ºC for ~ 5-7 minutes, then place in room temperature ~5-7 minutes until the biomass is completely thawed. Repeat the freeze-thaw process 2 more times and transfer the lysed biomass to a large douncer (tissue grinder). Use the “loose” piston and dounce the biomass ~ 100 times (avoid air bubble formation). Transfer the lysate to 50 ml centrifuge tubes (remember to balance the tubes carefully), and centrifuge at 15,000 rpm for 1 hour (Sorvall RC5, SS34 rotor, 4 ºC). After centrifugation, gently transfer the supernatant to a clean beaker (the supernatant now contains the protein, so take special care and maintain ice-cold conditions). Save the pellet at 4 ºC until you see the protein eluted from the SP-Sepharose column.
**Dialysis:** To remove excess NaCl and other impurities, dialyze the supernatant against 4 L Buffer A for 3 hours. Continue further dialysis overnight against new Buffer A (4 L). Measure the conductivity of the dialyzed supernatant before proceeding with further purification (conductivity ~ 20-50 mM). When you measure conductivity, take 200 µl sample and add 1,800 µl ddH₂O and mix well and get the reading. The final conductivity is 10 times of the reading.

**Protein Purification: Column 1 – SP Sepharose:** Resuspend and wash the SP Sepharose resin in water (in a large beaker) according to manufacturer’s instructions, and pack a 25 ml column. Equilibrate the column by passing through 250 ml Buffer A₂⁵ (Buffer-A + 25 mM NaCl) at 4 °C. Check that the pump tubes are tight and well fixed. Load the supernatant slowly (~ 1.0-1.2 ml/min). Watch the column every 10 minutes to ensure that there is no leakage. After the loading is complete, wash away unbound impurities with 250 ml Buffer-A₂⁵ (~ 2.0-2.5 ml/min). Elute bound protein with 250 ml Buffer-A: 100 – 300 mM NaCl gradient and collect 80 fractions of about 3 ml (~ 128 drops) each (~ 1.5-2.0 ml/min). Analyze 10 µl of alternate fractions (mix with 5 µl SDS loading dye) on a 10 % SDS-PAGE gel. After confirming the presence of protein in the eluate, wash the column with 200 ml 2 M NaCl. Pool the necessary fractions (~ 120-160 ml) and measure its conductivity (should be ~ 220-270 mM). Before loading the pool onto the Hepain column, dilute it with Buffer A (~ 80-100 ml) so as to bring the conductivity to ~ 150 mM. Proceed with further purification on a Heparin column.
**Column 2 – Heparin:** Resuspend and wash the Heparin resin according to manufacturer’s instructions and pack a 20 ml column (note: the Heparin resin expands when packed at RT, so make sure you process this column and pack it at 4 ℃). Equilibrate the column by passing through 200 ml Buffer B$^{50}$ (Buffer-B + 50 mM NaCl) at 4 ℃. Buffer B: 20 mM potassium phosphate pH 7.2 (50 ml 1 M potassium phosphate pH 7.2: 36 ml 1 M K$_2$HPO$_4$ + 14 ml 1 M KH$_2$PO$_4$). Load the supernatant slowly (~ 0.5-1.0 ml/min). Watch the column every 5 minutes to ensure that there is no leakage. After the loading is complete, wash away unbound impurities with 200 ml Buffer-B$^{50}$ (~ 1.5-2.0 ml/min). Elute bound protein with 200 ml Buffer-B: 150 – 550 mM NaCl gradient (~ 1.0-1.5 ml/min) and collect 80 fractions of 2.5 ml each (~ 96 drops). Analyze 10 µl of alternate fractions on a 10 % SDS-PAGE gel. After confirming the presence of protein in the eluate, wash the column with 200 ml 2 M NaCl. Pool the necessary fractions (~ 110-120 ml) and measure its conductivity (should be ~ 400-500 mM). Before loading the pool onto the FFQ-Sepharose column, dilute it with Buffer B (~ 200-280 ml) so as to bring the conductivity to ~ 175 mM. Proceed with further purification on a FFQ-Sepharose.

**Column 3 –FFQ Sepharose (No glass container):** Resuspend and wash the Q-Sepharose resin according to manufacturer’s instructions and pack a ~ 6-7 ml column. Equilibrate the column by passing through 100 ml Buffer A$^{50}$ (Buffer A + 50 mM NaCl) at 4 ℃. Load the supernatant slowly (~ 0.8-1.2 ml/min). Watch the column every 5 minutes to ensure that there is no leakage. After the loading is complete, wash away unbound impurities with 100 ml Buffer-A$^{100}$ (Buffer A + 100 mM NaCl) (~ 1.2
-2.0 ml/min). Elute bound protein with 80 ml Buffer-A: 250 – 500 mM NaCl gradient (~ 1.0 ml/min) and collect 80 fractions of 1 ml each (~ 40 drops). Analyze 10 µl of alternate fractions on a 10 % SDS-PAGE gel. After confirming the presence of protein in the eluate, wash the column with 100 ml 2 M NaCl. Pool the necessary fractions and measure its concentration using the Bradford method. If the protein concentration is above 5 µM, dialyze directly to avoid loss of protein during concentration. Concentrate the pooled Msh2-Msh6 protein fractions using a Centriprep YM-30 (Amicon). Fill the centriprep with 15 ml Buffer A, and centrifuge at 3500 rpm using a SS-34 rotor to wet the membrane and to check for leaks (use the small bolt to screw the rotor onto the centrifuge). Add the pooled protein to the centriprep and spin at 3500 rpm for ~ 10 minutes at 4 °C. Continue centrifugation until the final volume is reduced to ~ 6-8 ml. Measure concentration of the concentrated protein and dialyze against 2 L Buffer A^{50} (Buffer A + 50 mM NaCl) overnight. Measure the protein concentration after the overnight dialysis, aliquot and freeze at -80 °C. The yield of Msh2-Msh6 purification was ~ 1-1.5 mg per liter of *E. coli* culture.

**T. aquaticus MutS overexpression and purification**

**Plasmid**: Full length *T. aquaticus* MutS was expressed in *E. coli* bacteria from MutS:pET3b plasmid DNA under the control of a T7 promoter overexpression system. The E41A, F39A, F39W MutS DNA binding mutants were similarly overexpressed from E41A/F39A/F39W:pET17b. These clones were gifts from Dr.
Peggy Hsieh (N.I.H.). MutS wildtype and E41A, F39A, F39W proteins were purified using a previously reported protocol (Biswas and Hsieh, 1996) with modification.

**Transformation:** Add approximately 50 ng of *T. aquaticus* MutS:pET3b plasmid DNA to 100 µl BL21(DE3) competent cells (for mutants, instead use approximately 125 ng *T. aquaticus* Mutant MutS:pET17b plasmid DNA). Incubate cells for 30 min on ice, heat shock at 42 °C for 90 sec, and return to ice for an additional 2 min. Add 80 µl sterile LB broth at room temperature and shake at 230 rpm for 30 min at 37 °C. Remove the cells with a sterile pipette tip and plate on a fresh LB-ampicillin plate. Incubate at 37 °C for 12-14 hours, and when you can see clear colonies, remove to room temperature to minimize overgrowth of colonies.

**Growing biomass:** Autoclave 12 L sterile LB media (6 x 2 L) in 4 L flasks, and cool the flasks to room temperature and add 2 ml 100 mg/ml ampicillin to each flask. Choose a single, isolated colony from the transformation and resuspend in 500 µl sterile LB and vortex well. Add 50~100 µl to each of the 6 flasks and grow the culture overnight on a shaker at 230 rpm at 37 °C. The cultures take approximately 9-10 hours to reach OD$_{600}$ = 0.6 under these conditions. When the cultures reach OD$_{600}$ = 0.6, take 1 ml uninduced sample for testing on a gel and then induce the cultures with 0.5 mM IPTG (add 1 ml of 1 M IPTG stock solution to each flask). Grow the culture for an additional 3 hours and take 1 ml induced sample for the induction gel. Spin down biomass cultures in 500 ml centrifuge tubes at 2500 rpm for 20 min at 4 °C (Sorvall RC5 centrifuge, GS3 rotor), decanting the clear supernatant and refilling with culture until none remains. Decant the last of the supernatant and
transfer centrifuge tubes to an ice bucket. Resuspend cell pellets on ice in resuspension buffer: 50 mM Tris-Cl pH 7.5, 10% sucrose. Optimal resuspended volume for 12 L biomass is 50-70 ml. Split and store biomass in 50 ml falcon tubes at 80 °C until further purification.

**Induction gel:** Spin down 1 ml uninduced and induced samples that were collected earlier in a table-top centrifuge for 1 minute. Decant the supernatant and resuspend pellets in 40 µl ddH2O and add 20 µl 5X SDS gel loading dye. Boil samples for 3 min and vortex and pipette the pellets to resuspend cells. Analyze 10 µl of each sample alongside a broad-range SDS marker on a 10% SDS gel by PAGE. Run gel at room temperature, 120 V, for 1.5 hours, stain with Coomassie Brilliant Blue dye, and look for an increase in intensity of the 89.3 kDa band (monomeric MutS) in induced samples relative to uninduced.

**Cell lysis:** Thaw 50 ml frozen biomass by placing the 50 ml falcon tubes in a beaker of water on ice. When thawed, transfer to a 200 ml beaker on ice and add 50 mM NaCl (2.5 ml 2 M NaCl), 1.4 mM β-mercaptoethanol (140 ul pure β-mercaptopoethanol in hood), PMSF (a few crystals dissolved in 100 µl isopropanol) and 1 mg/ml lysozyme (100 mg lysozyme resuspended in 0.5 ml ddH2O) to final 100 ml with 1 x Buffer C (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 5% glycerol, by volume). Stir the mixture gently for 1 hour at 4 °C, and then transfer back into 50 ml falcon tubes (~ 40 ml per falcon tube). The mixture should be extremely viscous but light-colored. Freeze tubes in liquid nitrogen for 2 minutes, thaw in a 37 °C water bath for 4-6 minutes, and then transfer to room temperature until completely thawed. Repeat 6
times, and then transfer thawed biomass to a glass tissue ginder. Dounce with a loose piston ~ 80-100 times, avoiding the formation of air bubbles. Place the lysate, which should now pour easily, into clean 50 ml centrifuge tubes, balance to within 0.1 g, and centrifuge at 15,000 rpm for 1 hour 15 min at 4 °C (Sorvall RC5 centrifuge, SS34 rotor). Immediately after centrifugation, gently transfer supernatant to a clean 50 ml centrifuge tubes without disturbing the pellets, balance, cap tightly and then incubate at 70 °C for 30 minutes. Be careful not to go above 70 °C. The contents of the tubes should become cloudy with *E. coli* proteins aggregated by the heat. Separate thermostable MutS from these host proteins by centrifuging at 17,000 rpm for 45 min (Sorvall RC5 centrifuge, SS34 rotor) at 4 °C. Place the supernatant containing MutS (it should be clear but light yellow in color) in a small beaker for further column purification.

**Purification on a Q-sepharose column (room temperature):** Pack a 14 ml Q-sepharose column at room temperature. Equilibrate the column with 140 ml 1 x Buffer C + 50 mM NaCl. Load the supernatant containing MutS onto the column at a rate of approximately 1.0-2.0 ml/min, and begin collecting “flow-through” as soon as the yellow color nears the column bottom. After the protein is completely loaded, elute at 2.0-2.5 ml/min with a 140 ml gradient from 50-400 mM NaCl: place 70 ml Buffer C + 50 mM NaCl in one chamber of a gradient maker and 70 ml Buffer C + 400 mM NaCl in the other; make sure that the chamber containing the lower salt concentration is connected directly to the pump, that connecting loop is open, and that the solution is stirring as the pump draws the gradient onto the column. Collect 80 x
1.75 ml fractions in class tubes using an automated fraction collector (approximately 50 drops per tube, but check the fraction volume every few tubes because the drop volume will change over the course of elution). Analyze every other fraction, as well as pellet, load, and flow through by SDS-PAGE on 10% gels.

**Ammonium sulfate precipitation**: Pool the fractions with MutS, taking care to avoid impurities (a lack of selectivity may lead to nuclease contamination), and place in a glass beaker. Determine the volume of the pool (Y) and measure out $Y \times 0.2423$ g of ammonium sulfate, add the ammonium sulfate gradually to the protein while stirring slowly at room temperature, and continue to stir slowly for an additional 5 min. After all ammonium sulfate has dissolved, incubate at room temperature for 30 minutes with occasional mixing to fully precipitate MutS. Centrifuge the solution at 5,000 rpm for 10 min, and decant the supernatant. Finally, resuspend the pellet in 4 ml 1 x Buffer C. Analyze 10 µl samples of the pool, supernatant, and resuspended pellet by SDS-PAGE on a 10% gel. Dialyze the resuspended solution of MutS overnight in 4L Buffer C + 50 mM NaCl at 4 °C. Measure the protein concentration, aliquot, and store at - 80 °C. The yield of MutS purification was ~ 2-2.5 mg per liter of *E. coli* culture.

**Purification of E. coli PBP-A197C**

**Solution preparation**: Autoclave one empty 50 ml flask and 1 x 100 ml LB in a 500 ml flask. Pour 10 ml sterile LB into the autoclaved 50 ml flask. Autoclave 50 ml 60 % (v/v) glycerol. Prepare three of 14% SDS-PAGE gels and keep at 4 °C until use.
Make 6 ml 12.5 mg/l (1000x) tetracycline: weigh 75 mg tetracycline and add 50 % ethanol to final volume 6 ml in a 15 ml falcon tube. Aliquot 0.2 ml to an eppendorf tube and aluminum foil wrap the rest solution and store at - 20 °C. Make 1 L of 4M NaCl: weigh 233.76 g NaCl and add ddH2O to final volume 1 L. Make 50 ml of 1 M MgCl2: weigh 10.165 g MgCl2·6H2O and add ddH2O to final volume 50 ml. Make 1 L 0.5 M EDTA pH8.0: 186.1 g EDTA·2H2O and add ~ 20 g NaOH and add 800 ml ddH2O, mix vigorously and adjust pH to 8.0 and add ddH2O to final volume 1 L, then autoclave to sterilize. Make 1 L of 1 M Tris-HCl pH 7.8, pH 7.6 and pH 7.5. The pH of Tris buffer must be well-adjusted especially for a Q-Sepharose column, because it is performed at the edge of binding conditions (Tris buffers, pH 7.8 & 7.6, and the theoretical pl of this PBP is 6.91). To get well-adjusted pH of Tris, reajust pH after ~ 1-2 hours in room temperature and add ddH2O to final volume. Make sure that you sign up the skaker on the hallway for 8 x 0.5 L cultures during second and third days, becaue you can use hot room for seeding cultures but not for expression. Pack 30 ml FFQ-Sepharose column in a XK16 column and wash with plenty ddH2O at ~ 1-5 ml/min at room temperature and equilibrate with ~ 300 ml of 10 mM Tris-HCl, pH 7.8 at 3 ml/min. After equilibration, store column in the cold room.

**PBP-Glycerol Stock Preparation:** Streak out a swab of PBP culture on a LB-tetracycline plate (tetracycline concentration is 12.5 mg/L; note: make tetracycline stock in 50 % ethanol and store at – 20 °C covered with aluminum foil) and incubate overnight (~ 16 hours) at 37 °C. Pick one well-isolated and nice-round shaped colony from the plate into 25 ml LB with tetracycline and grow culture to an OD600 of 0.4 at
250 rpm in a shaker at 37 °C for ~ 3-5 hours. Place 1.5 ml of the culture in a sterile 2 ml vial (O-ring seal screw cap, sterile, liquid nitrogen fine) with 0.5 ml 60 % glycerol (autoclaved and chilled). Vortex the samples well to ensure even distribution of glycerol and rapidly freeze the vial by immersing it in liquid nitrogen and store at -80°C.

Day 1—Growing Biomass: Thaw out one 2 ml glycerol stock of PBP in a 37 °C water bath, and add the contents to 10 ml of LB-tetracycline (10 µl of 12.5 mg/ml filter sterilized tetracycline, 12.5 mg/L final) in a 100 ml autoclaved flask. Grow the culture at 37 °C for ~ 6 hours with constant shaking at 250 rpm. After 6 hours of growth, transfer the 10 ml culture to 100 ml LB-tetracycline in a 500 ml flask (100 µl 12.5 mg/ml tetracycline stock). Grow the culture at 37 °C, with constant shaking at 250 rpm for ~ 16 hours. Set 8 x 2 L holders in the shaker for tomorrow morning.

Preparation of minimal media: Prepare 4 L minimal media (MM) according to the composition below. Minimal media of 4 L: add 100 mM Hepes (95.32 g), 20 mM KCl (40 ml 2 M stock or 5.96 g), 15 mM (NH₄)₂SO₄ (7.92 g), 1 mM MgCl₂ (4 ml 1 M stock), 10 µM FeSO₄ (11.12 mg), 1 µg/ml thiamine (4 mg), 0.25 % (v/v) glycerol (10 ml 100 % glycerol), 2 mM KH₂PO₄ (8 ml 1 M stock or 1.09 g), stir in ~ 3.9 L H₂O, adjust pH to 7.5 with 10 N NaOH (~ 20 ml), then fill up to 4 L with ddH₂O. Split the media into 8 x 2 L flasks (500 ml each) and autoclave. Cool the media to room temperature and add ~ 500-600 µl 12.5 mg/ml tetracycline stock (15 mg/L final).
**Day 2:** Prepare 0.2 M rhamnose (1.82 g of 2 M rhamnose in the desiccator to final 50 ml with ddH$_2$O). Prepare resuspension buffers (RBs) for tomorrow. Resuspension buffer I (1 L): 10 mM Tris-HCl, pH 7.6 (10 ml 1 M stock), 30 mM NaCl (7.5 ml 4 M stock), 982.5 ml ddH$_2$O. Resuspension buffer II (100 ml): 33 mM Tris-HCl, pH 7.6 (3.3 ml 1 M stock) and make volume to 100 ml with ddH$_2$O. Resuspension buffer III (100 ml): 33 mM Tris-HCl, pH 7.6 (3.3 ml 1 M stock), 40% (w/v) sucrose (40 g), 0.1 mM EDTA (20 µl 0.5 M stock) and make volume to 100 ml with ddH$_2$O. Resuspension buffer IV (ice cold): 0.5 mM MgCl$_2$ (100 µl 1 M stock) and make volume to 200 ml with ddH$_2$O and place in –20 °C freezer for 5 min and store in 4 °C. RB I, II and III can keep in room temperature. Shake 8 x 0.5 L minimal media with tetracycline at 230 rpm on the shaker at 37 °C and prewarm them for ~ 30-60 minutes. After 16 hours of growth, add 10 ml of the culture to the pre-warmed minimal media in each flask and incubate at 37 °C with rapid shaking at 230 rpm. Take 1 ml samples of the minimal media before adding the culture for baseline OD measurement, and another 1 ml sample after adding the 10 ml culture for OD measurement at 0 hour time point. Measure OD$_{600}$ of the culture after each hour until the OD$_{600}$ shows a net increase of 0.2 (note: there should have a jump in OD$_{600}$ between the first and third hour of growth). The increase in OD value over time should be similar to the values shown below: 0 hour: OD$_{600}$ is 0.039 and net change in OD is 0; 1 hour: OD$_{600}$ is 0.065 and net change in OD is 0.026; 2 hours: OD$_{600}$ is 0.151 and net change in OD is 0.112; 3 hours: OD$_{600}$ is 0.256 and net change in OD is 0.217. Sometimes net change will take ~ 6-7 hours to reach. When the net change in
OD reaches 0.15 – 0.2, add 5 ml of filter sterilized 200 mM rhamnose (2 mM final) to each flask. Incubate the culture for an additional ~ 16-24 hours at 37 ºC with rapid shaking at 230 rpm. The OD<sub>600</sub> after overnight induction should be ~ 1.5 (the cultures must have a pink color after the overnight incubation). Book and warm up the centrifuge and F10 rotor to room temperature for tomorrow morning.

**Day 3: Preparation of buffers:** To minimize Pi contamination, avoid glassware throughout purification as much as possible. Perform FFQ-Sepharose column at room temperature. Prepare elution buffer (EB) at room temperature and dialysis buffer (DB) at 4 ºC. Elution buffer: 150 ml of 10 mM Tris-HCl pH 7.6 with 0 mM NaCl and 150 ml of 10 mM Tris-HCl pH 7.6 with 200 mM NaCl. Dialysis buffer: 2 L of 10 mM Tris-HCl pH 7.5.

**Osmotic shock:** PBP is exported into the periplasm after expression. To collect the periplasmic fraction, osmotic shock is used. There are several keys to ensure a good yield of PBP protein: cluture must be pink colored; for each step of osmotic shock, remove previous buffer nicely and resuspend pellet in next buffer nicely; RBIV must be ice cold. If cell cultures are turned into pink in color, check OD<sub>600</sub>. If OD<sub>600</sub> is around 1.5, start harvesting cells. Centrifuge the culture at 5,000 rpm for 15 minutes, and resuspend the pellet in 650 ml RBI with a spatula at RT and split to two centrifuge bottles. Centrifuge the resuspended biomass at 5,000 rpm for 15 minutes at RT, and resuspend the pellet in 350 ml RBI with a spatula and combine into one bottle. Centrifuge again at 5,000 rpm for 20 minutes at RT and discard carefully the supernatant (the cell pellet tends to be very loose at this point). Resuspend the pellet
thoroughly in 100 ml RBlI with a spatula, add a stir bar and stir vigorously for 2 minutes at RT. Cool down the centrifuge and rotor to 4 °C. To the stirring solution, add 100 ml RBIII and continue stirring vigorously for an additional 10 minutes. Centrifuge at 9,000 rpm for 20 min at 4 °C. During this spin, prepare an ice-bath and move RBIV into a -20 °C freezer 5 minutes before addition. Discard the supernatant, and rapidly but thoroughly resuspend the pellet in 200 ml ice-cold RBIV with a spatula and gently stir for 15 minutes in an ice bath. During this step, periplasmic proteins are released. Remove a stirring bar before spin. Save 100 µl of the resuspended solution for the purification gel, and centrifuge the solution at 9,000 rpm for 20 minutes at 4 °C. Transfer the supernatant into a clean cylinder and it should be pinkish purple. The supernatant now contains PBP; determine the concentration of the supernatant to obtain PBP yield (~ 300-350 mg). Measure absorbance for 10 µl of the supernatant at A\textsubscript{280} and A\textsubscript{320}, and calculate the PBP concentration (molecular weight of PBP = 34.45 kDa) using the formula: Total PBP (mg) = (A\textsubscript{280} - A\textsubscript{320}) / 1.78 mg\textsuperscript{-1}\cdot ml\textsuperscript{-1} x dilution factor (100) x total volume (200 ml). If the yield of PBP protein is much less than 350 mg, repeat the step of adding RBIII and RBIV one more time and measure the concentration again.

**Protein Purification with FFQ-Sepharose column:** Perform FFQ-Sepharose column at room temperature and no wash step required. Add ~ 2 ml of 1 M Tris-HCl pH 7.8 (make up final concentration to 10 mM) into ~ 200 ml supernatant and gently but nicely mix. Take 1 ml for a gel sample of loading. Load the supernatant slowly ~ 0.5-1.0 ml/min onto 30 ml FFQ-Sepharose column. Watch the column every 5 minutes to
ensure that there is no leakage. Do not wash the column after loading the protein. Elute bound protein with 300 ml of 10 mM Tris-HCl, pH 7.6: 0 – 200 mM NaCl gradient at 1.5 ml/min and collect ~ 80 fractions of 1.5 ml each. Keep eluting and checking alternative numbered fractions by mixing 1 µl sample with 100 µl Bradford in a microplate until a first peak shows up and passes. This corresponds to PBP and determines which elution fractions have to be examined on gels. Analyze 20 µl of every alternate fraction on 14 % SDS-PAGE (Figure 2.2). After confirming the presence of protein in the eluate, wash the column with 200 ml 2 M NaCl. Pool (~ 50-65 ml) and dialyze the fractions containing PBP in a dialysis bag (MWCO 6-8,000, 3.3 ml/cm) overnight against 2 L dialysis buffer (DB) in the cold room.

**Day 4:** Transfer the PBP from a dialysis bag to a falcon tube and mix well by gentle inversion. Check PBP concentration: PBP (µM) = A$_{280}$ / 61880 M$^{-1}$·cm$^{-1}$ x dilution factor x 10$^6$; PBP (mg/ml) = A$_{280}$ / 1.78 mg$^{-1}$·ml$^{-1}$ x dilution factor. Freeze at -80 °C until labeling after measuring the concentration. The concentration of the pooled fraction should be ~ 6 mg/ml (150 – 200 mg total). Cleaning the used FFQ column: scrape off the top of the resin with a purple tint and wash the column with ~ 100 ml of 2 M NaCl at 3 ml/min. Transfer the washed resin into a 50 ml falcon tube and store in the cold room.
**Figure 2.2. Purification of phosphate binding protein (PBP).** 10 ml of the sample from alternate fractions was analyzed on 14 % SDS-PAGE (M – marker, L – load). Fractions containing PBP were pooled and labeled with MDCC.
Figure 2.2.
Labeling *E. coli* PBP-A197C with MDCC

**Preparation:** Measure the concentration of the protein after dialysis as described earlier. Labeling efficiency is optimum when the concentration of the PBP in the reaction is > 60 mM. Concentrate the protein using a YM-10 centriprep if the concentration of the protein is lower than 60 mM. The 17 ml FFQ column is used for labeling of ~ 130-200 mg PBP. If PBP is less or more than this range the column should be scaled down or up. Pack a 17 ml FFQ column in a XK16 column and wash with plenty ddH₂O at ~ 1-5 ml/min and equilibrate with ~ 200 ml washing buffer (10 mM Tris-HCl pH 8.0) at 3 ml/min and keep in the room temperature. Make 100 ml exchange buffer (EB): 20 mM Tric-HCl pH 8.1; 500 ml washing buffer (WB): 10 mM Tris-HCl pH 8.0; 100 ml WB with 50 mM NaCl and keep them in the room temperature. Presoak 2 x Centriplus10 with EB. Prepare two 14 % SDS-PAGE gels and keep at 4 °C. Equilibrate a roller drum to RT before use. Make three main stock reagents are needed. Stock 1:50 mM 7-Methyguanosine (7-MEG): MW: 283.2, dissolve 14.2 mg in 1 ml ddH₂O and aliquot 50 µl and store at - 80 °C. Stock 2: 1 U/µl Purine Nucleoside Phosphorylase (PNPase): Read the label of a PNPase bottle and add an appropriate volume (~ 200 ml) ice-cold ddH₂O to entire bottle (Sigma). Rinse out all over the inner surface of the bottle by tilting and briefly vortex. Aliquot 50 µl fractions and store at - 80°C as lab stock. Upon thawing each tube at first time, further aliquot 10 µl and keep as personal stock. Stock 3: 25 mM MDCC: MW: 383.5, add 520 µl N, N-dimethylformamide (DMF), mix by vortexing. Rinse out all over the inner surface of the bottle by titling and
vortex well. This substance is light sensitive, so turn off all lights before opening the colored bottle. Wrap the MDCC solution bottle with aluminum foil after dissolving, and store at -20 °C. Estimate how much stock reagents are needed. PNPase (µl) = \{\text{final concentration of PNPase (0.2 U/ml) x total volume of PBP (µl)}\} / 1 \text{ U/µl PNPase stock}; 7-MEG (µl) = \{\text{final concentration of 7-MEG (200 µM) / 7-MEG stock concentration (50 x 10}^{3} \text{ µM)} \} x \text{ total volume of PBP (µl) x 10}^{3}; \text{MDCC (µl) = } \{\text{PBP concentration (µM) x 1.5} \} / \{\text{MDCC stock concentration (25 x 10}^{3} \text{ µM) x total volume of PBP (µl) x 10}^{3}\}.

**MDCC Labeling (2 hours):** MDCC is relatively stable under white illumination. Labeling and purification can be carried out in the light. But protect the MDCC stock bottle from light. Thaw PBP and equilibrate to RT. Add PNPase and 7-MEG to the tube and incubate on the roller drum at moderate a moderate speed for above 30 minutes at RT. This mopping is important because apo PBP is more efficiently labeled than the Pi-bound form. In the dark, bring the MDCC stock solution bottle to room temperature, vortex well and add MDCC to the mopped PBP tube. Aluminium foil wrap the tube of the labeling reaction and incubate on the roller drum at a moderate speed for above 1 hour at RT.

**Buffer exchange (4 hours):** Split the labeling reaction into 2 x presoaked Centriprep10. Prepare a gel sample of “after labeling”; take 10 ml protein and add 10 ml 2 x no-dye SDS loading buffer. Spin (ss34 rotor) at 4k rpm 10 – 20 °C, 30 min. Transfer the filtrates to a falcon tube, and fill reservoir up to a 15 ml line with the labeling reaction if any left or EB. Every time after spin, record the volumes of the
filtrate and buffer. Repeat previous steps as needed. Usually five times is enough to get the 7/8 x volume of the initial buffer replaced. For a last spin, do not add buffer and just concentrate down to 2.5 ml x 2 (Table 2.5). Sometimes free MDCC precipitates along with spin steps (white precipitates), but it will be removed by a following filtering. Ultrafiltration removes a certain amount of free MDCC in parallel with exchanging buffer. Transfer the retentates into a falcon tube and filtrate with φ = 0.2 mm filter (low protein binding such as Tuffryn). Rinse the Centriplus units with 10 ml EB, filtrate the EB buffer with the filter used for the retentate, and combine with the previously filtered retentate. Finally, total volume is ~ 15 ml and the 23/24 x volume of the initial buffer is exchanged.

**FFQ column purification (4 hours at RT):** Remove the top part from the packed FFQ column. Lower the buffer level down to the surface of the resin by gravity flow. Load the labeling reaction directly onto the Q column (~ 15-17 ml). Run the column by gravity flow until the reaction completely enters the resin. Overlay washing buffer (WB) and reassemble the top part to the column. Wash at ~ 1.5-3.0 ml/min with 50 ml of WB or until yellow passes. Elute with 150 ml of 0 – 50 mM NaCl gradient in 10 mM Tris-Cl, pH 8.0 at 1.5 ml/min. Collect elution fractions (1.5 ml / 61 drops / tube). Be careful that drop size becomes double soon after the labeled PBP starts to elute. Keep your eyes on a fraction collector during elution. Once elution becomes yellow in color, start to check A$_{280}$ (protein) and A$_{430}$ (MDCC) of alternative numbered fractions by mixing 10 ml protein with 990 ml H$_2$O. Plot these absorbance values along with fraction # in an excel template “PBP_Labeling_Template” (Figure
2.3). When the protein peak passes, stop elution. Although yellow-colored fractions could still continue, these contain just free dye (only conjugated MDCC efficiently absorbs 430 nm light). PBP (Pi-free, native condition), \( \varepsilon_{280} = 61880 \) (M\(^{-1}\)·cm\(^{-1}\)), MW 34453.8 and MDCC (conjugated), \( \varepsilon_{430} = 46800 \) (M\(^{-1}\)·cm\(^{-1}\)). Pool fractions containing PBP labeled with MDCC at above 70%. Sometimes, fractions with extraordinary high labeling follow the labeled PBP peak. Do not take them because these include labeled non-PBP proteins. Check concentration and labeling efficiency of the pooled protein by measuring \( A_{260} \) and \( A_{430} \). If the concentration is above 100 \( \mu \)M, proceed to aliquot 0.5 \( \mu \)l and store at -80 °C. Otherwise concentrate protein before storage. Prepare a gel sample of “after purification”: 10 \( \mu \)l protein and add 10 \( \mu \)l 2 x no-dye SDS loading buffer. Pool # 41-57 fractions on 04/29/2008 and its volume was 24.5 ml and the final PBP-MDCC concentration was 153.5 \( \mu \)M with 81% labeling efficiency. Yield is \{153.5 \( \mu \)M (final concentration of PBP-MDCC) \times 24.5 ml (final volume PBP-MDCC)\} / \{66.76 \( \mu \)M \times 73.5 ml (initial PBP)\} = 77%. To check MDCC labeling quality, run a 14% SDS-PAGE gel (Figure 2.3). The loading volumes of “after labeling” (AL) and “after purification” (AP) samples are adjusted to be same amount of PBP. The FT+W sample in Figure 2.4 was prepared from a mixture of flow-through and washing fractions during the FFQ column. Take a picture on a UV illuminator with 365 nm (Figure 2.4). Stain gel with CBB and destain and scan. After labeling, wash the column at 3 ml/min with 3 x column volume of 2 M NaCl or until no yellow color elutes. Transfer the washed resin into a 50 ml falcon tube and store in the cold room.
### Tabel 2.5. PBP labeling buffer exchange chart.

Buffer exchange,
04/29/08

<table>
<thead>
<tr>
<th>Start volume</th>
<th>Spin</th>
<th>Filtrate</th>
<th>Retentate</th>
<th>Volume ratio of the original buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ 15 ml PBP x 2</td>
<td>→ 30 min</td>
<td>Total 12.5 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 6.5 ml PBP x 2</td>
<td>→ 30 min</td>
<td>Total 27.5 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 7.5 ml PBP x 2</td>
<td>→ 30 min</td>
<td>Total 42.5 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 7.5 ml PBP x 2</td>
<td>→ 30 min</td>
<td>Total 56 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 1 ml PBP x 2</td>
<td>→ 30 min</td>
<td>Total 70 ml</td>
<td>1/2</td>
<td></td>
</tr>
<tr>
<td>+ 6.5 ml EB x 2</td>
<td>→ 30 min</td>
<td>Total 81 ml</td>
<td>1/4</td>
<td></td>
</tr>
<tr>
<td>+ 7.0 ml EB x 2</td>
<td>→ 30 min</td>
<td>Total 95 ml</td>
<td>1/8</td>
<td></td>
</tr>
<tr>
<td>+ 7.0 ml EB x 2</td>
<td>→ 30 min</td>
<td>Total 107 ml</td>
<td>1/16</td>
<td></td>
</tr>
<tr>
<td>(No addition)</td>
<td>→ 30 min</td>
<td>Total 120 ml</td>
<td>Total 6 ml</td>
<td></td>
</tr>
</tbody>
</table>

Final total is 17 ml by adding 12 ml EB 1/32
Figure 2.3. Labeling efficiency of PBP with MDCC. Absorbance of 10 ml of the sample from the FFQ column fractions was measured at 280 and 430 nm to measure the concentration of PBP and MDCC label respective. Pink line is PBP concentration, blue line is conjugated MDCC concentration and green line is the labeling percentage.
Figure 2.3.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>A_280 (mM)</th>
<th>PBP (µM)</th>
<th>Conjugated MDCC (µM)</th>
<th>Labelling (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.0005</td>
<td>0.003</td>
<td>6.4</td>
<td>79</td>
</tr>
<tr>
<td>18</td>
<td>0.0111</td>
<td>17.8</td>
<td>12.8</td>
<td>72</td>
</tr>
<tr>
<td>22</td>
<td>0.02</td>
<td>50.1</td>
<td>25.6</td>
<td>85</td>
</tr>
<tr>
<td>26</td>
<td>0.031</td>
<td>59.8</td>
<td>42.7</td>
<td>82</td>
</tr>
<tr>
<td>30</td>
<td>0.037</td>
<td>66.3</td>
<td>49.1</td>
<td>84</td>
</tr>
<tr>
<td>34</td>
<td>0.041</td>
<td>59.8</td>
<td>49.1</td>
<td>82</td>
</tr>
<tr>
<td>38</td>
<td>0.043</td>
<td>40.4</td>
<td>32.1</td>
<td>79</td>
</tr>
<tr>
<td>42</td>
<td>0.025</td>
<td>27.5</td>
<td>21.4</td>
<td>78</td>
</tr>
<tr>
<td>46</td>
<td>0.017</td>
<td>21</td>
<td>15.1</td>
<td>71</td>
</tr>
<tr>
<td>50</td>
<td>0.013</td>
<td>16.2</td>
<td>17.1</td>
<td>81</td>
</tr>
<tr>
<td>54</td>
<td>0.013</td>
<td>12.9</td>
<td>10.7</td>
<td>66</td>
</tr>
<tr>
<td>58</td>
<td>0.013</td>
<td>11.3</td>
<td>10.7</td>
<td>79</td>
</tr>
<tr>
<td>62</td>
<td>0.013</td>
<td>9.7</td>
<td>10.7</td>
<td>83</td>
</tr>
<tr>
<td>66</td>
<td>0.018</td>
<td>565.6</td>
<td>8.5</td>
<td>95</td>
</tr>
</tbody>
</table>

PBP [Pi-free, native condition], extinction coefficient @ 280nm, 61880 (M^−1 cm^−1), MW 34453.8
MDCC [conjugated], extinction coefficient @ 430nm, 46800 (M^−1 cm^−1)
Figure 2.4. MDCC labeling quality check with SDS-PAGE gel and UV scan. AL is after labeling sample, AP is after purification sample and FT+W sample was prepared from a mixture of flow-through and washing fractions during the FFQ column. Left is CBB stained gel and right is UV scan in 365 nm.
Figure 2.4.
MutS-DNA binding at equilibrium

*T. aquaticus* MutS-DNA 2-Ap binding assay at equilibrium

*T. aquaticus* MutS interactions with 2-Ap-labeled DNAs were measured on a FluoroMax-3 fluorimeter (Jobin-Yvon Horiba Group; Edison, NJ). Emission scans of 0.1 µM 23mer2Ap+T and 23mer2ApA:T duplexes were collected in 3 ml quartz cuvettes in DNA binding buffer (20 mM Hepes-NaOH, pH 7.7, 50 mM NaCl, 5 mM MgCl₂) at 40 °C, with and without MutS dimer (0.3 µM), after mixing for 30 s and λ<sub>EX</sub> = 315 nm. Although the excitation maximum for 2-Ap is at approximately 305 nm, the DNA substrates were excited at 315 nm in order to minimize excitation of tryptophan (maximum λ<sub>EX</sub> = 280 nm) and tyrosine (maximum λ<sub>EX</sub> = 274 nm) in MutS. The excitation and emission slit widths were 2 nm and 5 nm, respectively. Signal/Reference (S/R) fluorescence output was recorded every 2 nm from 330-430 nm and the emission maximum was determined (375 nm). Background fluorescence from the buffer and MutS was subtracted from the raw data. Titrations of 0.01 µM DNAs with 0–0.6 µM MutS and three MutS mutants were performed under similar conditions, in the absence or in the presence of 150 µM ATPγS, ATP, or ADP, with emission recorded at 375 nm. The data were corrected for intrinsic MutS fluorescence by subtracting data from parallel experiments with unlabeled DNA. Fluorescence intensity was plotted *versus* the concentration of MutS and the apparent dissociation constant (K<sub>D</sub>) for the interaction obtained by fitting the data to a quadratic equation:

\[
[D\cdot M]=F_0+(F_{\text{max}}-F_0)\left\{\frac{[K_D+[D_t]+[M_t]^-]}{[K_D+[D_t]+[M_t]]^2-4[D_t][M_t]}\right\}/2[D_t]
\]
where $D \cdot M$ is the fraction of MutS•DNA, $F_0$ is 23mer2Ap+T fluorescence in the absence of protein and $F_{\text{max}}$ is maximal fluorescence, and $D_t$ and $M_t$ are total molar concentrations of DNA and MutS, respectively. The data were fit by non-linear regression using KaleidaGraph (Synergy Software) (Jacobs-Palmer and Hingorani, 2007).

**T. aquaticus MutS-DNA anisotropy binding assay at equilibrium**

Anisotropy measurements of MutS–DNA interaction were performed under the same conditions as above by titrating 0.01 µM 23merTAMRA+T and 23merTAMRA A:T duplexes with 0–0.6 µM MutS at 23 °C, using vertically polarized light at $\lambda_{\text{EX}} = 555$ nm, and calculating anisotropy from the emitted vertical ($I_{VV}$) and horizontal ($I_{VH}$) polarized fluorescence intensities at $\lambda_{\text{EM}} = 582$ nm ($I_{VV} - GI_{VH}/I_{VV} + 2GI_{VH}$; where $G$ is the calculated grating correction factor). The binding isotherms obtained from plots of observed anisotropy versus the concentration of MutS were fit to a quadratic equation as described above (Jacobs-Palmer and Hingorani, 2007). The excitation and emission slit widths were 2.5 nm and 3.5 nm, respectively. Use individual small cuvettes for different MutS concentration.

**S. cerevisiae Msh2-Msh6-DNA 2-Ap binding assay at equilibrium**

Msh2-Msh6 binding to 2-Ap-labeled DNA was measured on FluoroMax-3 (Jobin-Yvon Horiba Group; Edison, NJ). DNA (0.025 µM) was titrated with Msh2-
Msh6 (0-0.6 µM) in binding buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl2), at 25 °C (λ_\text{EX} = 315 nm, λ_\text{EM} = 375 nm). The data were corrected for protein fluorescence, plotted versus [Msh2-Msh6] and fit to a quadratic equation for $K_D$ (Zhai and Hingorani, 2010). The excitation and emission slit widths were 2 nm and 5 nm, respectively.

**S. cerevisiae Msh2-Msh6-DNA anisotropy binding assay at equilibrium**

Fluorescence anisotropy was measured with TAMRA-labeled DNA (0.01 µM) titrating with Msh2-Msh6 (0-0.6 µM) in binding buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl2) at 25 °C (λ_\text{EX} = 555 nm, λ_\text{EM} = 582 nm). The binding isotherms obtained from plots of observed anisotropy versus the concentration of Msh2-Msh6 were fit to a quadratic equation as described above. The excitation and emission slit widths were 2.5 nm and 3.5 nm, respectively. Use individual small cuvettes for different Msh2-Msh6 concentration (Zhai and Hingorani, 2010).

**Msh2-Msh6-DNA binding kinetics**

**S. cerevisiae Msh2-Msh6-DNA 2-Ap binding kinetic assay**

DNA binding kinetics was measured on a KinTek SF-2001 stopped-flow (KinTek Corp.; Austin, TX), which has a mixing dead time of 1 millisecond. Single mixing experiments were performed by mixing Msh2-Msh6 (0.8 µM) with labeled DNA (0.12 µM) to measure association rates (λ_\text{EX} = 315 nm, λ_\text{EM} > 350 nm for 2-Ap
fluorescence) at 25 °C. Final reaction concentrations are 0.4 µM Msh2-Msh6 and 0.06 µM labeled DNA because the mixing ratio is 1:1. 2-Ap was selectively excited at 315 nm with Xenon lamp, with both entrance and exit slits set at 20 nm. The change in fluorescence intensity was monitored over time (up to 6 seconds) at all wavelengths above 350 nm using a photo-multiplier tube (PMT) at 700 V and a 350 nm long-pass cut-off filter (LG-350-F, Corion; Franklin MA). For each DNA substrate, at least five independent mixing events were monitored, and the resulting data averaged. The baseline fluorescence of Msh2-Msh6 mixed with buffer was subtracted from the fluorescence of Msh2-Msh6 mixed with 2-Ap-labeled DNAs to correct for intrinsic fluorescence of Msh2-Msh6. The increase in fluorescence over time with 2Ap DNA was fit to a single exponential equation in order to derive an apparent on-rate ($k_{observed}$) for its interaction with Msh2-Msh6:

$$A_t = A_0 (1 - e^{-kt}) + F_0$$

where $A_t$ corresponds to the fluorescence intensity at time $t$, $A_0$ is the amplitude of the fluorescence increase, $k$ is the apparent on-rate ($k_{observed}$), and $F_0$ is the baseline fluorescence of 2Ap DNA in the absence of protein. The data were fit by non-linear regression using KaleidaGraph.

Dissociation rates were measured by incubating Msh2-Msh6 (0.8 µM) and DNA (0.12 µM) for 10 minutes, then mixing with 8 µM unlabeled G:T DNA. Final concentrations are 0.4 µM Msh2-msh6, 0.06 µM DNA and 4 µM unlabeled DNA
because the mixing ratio is 1:1. The decrease in 2-AP fluorescence was fit to a single-
exponential equation:

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

where \( A_t \) corresponds to the fluorescence intensity at time \( t \), \( A_0 \) is the amplitude of the fluorescence decrease, \( k \) is the dissociation rate (\( k_{OFF} \)), and \( F_0 \) is the baseline fluorescence of 2-Ap DNA bound to Msh2-Msh6 protein. The data were fit by non-linear regression using KaleidaGraph.

Sequential mixing experiments were performed by mixing Msh2-Msh6 (0.8 \( \mu \)M) with labeled DNA (0.12 \( \mu \)M) for 0.25 s followed by addition of 10 \( \mu \)M unlabeled G:T DNA or buffer (mixing ratio: 1:1:1). Three or more kinetic traces were averaged and corrected for background signal, and fit to a single exponential function. Experiments with nucleotides were performed similarly except with final 0.5 mM ATP, 0.15 mM ATP\( \gamma \)S or 0.25 mM ADP added to Msh2-Msh6 and/or DNA.

**S. cerevisiae Msh2-Msh6-DNA anisotropy binding kinetic assay**

Anisotropy kinetics was measured on a KinTek SF-2001 stopped-flow (KinTek Corp.; Austin, TX), which has a mixing dead time of 1 millisecond. Single mixing anisotropy experiments were performed by mixing Msh2-Msh6 (0.8 \( \mu \)M) with labeled DNA (0.12 \( \mu \)M) to measure association rates (polarized \( \lambda_{EX} = 555 \) nm, \( \lambda_{EM} > 570 \) nm for TAMRA anisotropy) at 25 °C. Final reaction concentrations are 0.4 \( \mu \)M.
Msh2-Msh6 and 0.06 μM labeled DNA because the mixing ratio is 1:1. TAMRA was selectively excited at 555 nm with Xenon lamp, with both entrance and exit slits set at 10 nm. The change in both vertical ($I_{VV}$) and horizontal ($I_{VH}$) polarized fluorescence intensities were monitored over time (up to 6 seconds) at all wavelengths above 570 nm using two photo-multiplier tubes (PMT1 and PMT2) at 700 V and two 570 nm long-pass cut-off filters. Anisotropy was calculated from the emitted vertical ($I_{VV}$) and horizontal ($I_{VH}$) polarized fluorescence intensities at $\lambda_{EM} = 582$ nm ($I_{VV} - G I_{VH}/I_{VV} + 2G I_{VH}$; where $G$ is the calculated grating correction factor). For each DNA substrate, at least five independent mixing events were monitored, and the resulting data averaged. The baseline anisotropy fluorescence of Msh2-Msh6 mixed with buffer was subtracted from the anisotropy fluorescence of Msh2-Msh6 mixed with TAMRA-labeled DNAs to correct for intrinsic anisotropy fluorescence of Msh2-Msh6. The increase in anisotropy fluorescence over time with TAMRA DNA was fit to a single exponential equation in order to derive an apparent on-rate ($k_{observed}$) for its interaction with Msh2-Msh6:

$$A_t = A_0 (1 - e^{-kt}) + F_0$$

where $A_t$ corresponds to the fluorescence intensity at time $t$, $A_0$ is the amplitude of the fluorescence increase, $k$ is the apparent on-rate ($k_{observed}$), and $F_0$ is the baseline anisotropy fluorescence of TAMRA DNA in the absence of protein. The data were fit by non-linear regression using KaleidaGraph.
Dissociation rates were measured by incubating Msh2-Msh6 (0.8 μM) and DNA (0.12 μM) for 10 minutes, then mixing with 8 μM unlabeled G:T DNA. Final concentrations are 0.4 μM Msh2-msh6, 0.06 μM DNA and 4 μM unlabeled DNA because the mixing ratio is 1:1. The decrease in anisotropy fluorescence was fit to a single-exponential equation:

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

where \( A_t \) corresponds to the anisotropy fluorescence intensity at time \( t \), \( A_0 \) is the amplitude of the anisotropy fluorescence decrease, \( k \) is the dissociation rate (\( k_{OFF} \)), and \( F_0 \) is the baseline anisotropy fluorescence of 2-Ap DNA bound to Msh2-Msh6 protein. The data were fit by non-linear regression using KaleidaGraph.

Sequential mixing experiments were performed by mixing Msh2-Msh6 (0.8 μM) with TAMRA-labeled DNA (0.12 μM) for 0.25 s followed by addition of 10 μM unlabeled G:T DNA or buffer (mixing ratio: 1:1:1). Three or more kinetic traces were averaged and corrected for background signal, and fit to a single exponential function. Experiments with nucleotides were performed similarly except with final 0.5 mM ATP, 0.15 mM ATP\(\gamma\)S or 0.25 mM ADP added to Msh2-Msh6 and/or DNA.

*S. cerevisiae Msh2-Msh6 ATPase kinetics*  
Pre-steady state ATPase kinetics were measured by stopped-flow experiments using MDCC-labeled PBP (20 μM) to detect Pi release by Msh2-Msh6 (4 μM) on
mixing with ATP (1 mM) (mixing ratio 1:1) in the absence or presence of DNA (12 µM). So the final concentrations are 10 µM MDCC-labeled PBP, 2 µM Msh2-Msh6, 0.5 mM ATP and 6 µM DNA. The kinetic traces were fit to a burst equation:

\[
[\Pi] = \frac{(A_0 e^{-kt} + Vt + F_0)}{m_{\Pi}}
\]

\([\Pi]\) is Pi concentration, \(A_0\) is the amplitude and \(k\) is the observed burst rate constant, \(V\) is the velocity of the linear phase, \(F_0\) is initial fluorescence intensity, and \(m_{\Pi}\) is the slope of the Pi standard curve measured under the same conditions (Zhai and Hingorani, 2010). MDCC-PBP was excited at 425 nm, and the change in fluorescence upon binding to inorganic phosphate (Pi) was monitored over time (up to 2 seconds) at all wavelengths above 450 nm with a PMT at 400 V and a 450 nm long-pass cut-off filter (LL-450F, Corion; Franklin, MA). Because MDCC-PBP is sensitive to nM concentrations of Pi, all reactions and syringes were “mopped” for at least 30 minutes with a coupled enzyme reaction containing 200 µM 7-Methyl guanosine (7-MEG) and 0.001 U/µl Purine Nucleoside Phosphorylase (PNPase). PNPase catalyzes the addition of an inorganic phosphate to 7-MEG, sequestering contaminating Pi as ribose-1-phosphate before the initiation of reactions. To relate the fluorescence signal from MDCC-PBP to Pi concentration, a calibration curve was generated immediately prior to each experiment by mixing 350 µl of 20 µM MDCC-PBP in ATPase buffer (20 mM Tris-HCl pH 8.0, 110 mM NaCl, and 5 mM MgCl₂) with increasing
concentrations of Pi (0-8.0 µM), and monitoring the fluorescence increase over 0.2 seconds at 30 °C (final concentrations: 10 µM MDCC-PBP, 0-4 µM Pi).
Chapter III

Results

Msh2-Msh6•DNA complex stability underlies MMR signaling efficacy.

In equilibrium binding experiments, Msh2-Msh6 binds G:T, O\textsuperscript{6}MeG:T DNA with a higher affinity than +T DNA and almost no binding to G:C DNA.

The fundamental question of how Msh2-Msh6 targets specific sites in DNA is unresolved. To answer this question, we developed a fluorescence anisotropy assay that is an excellent general assay to measure the protein and DNA interactions. S. cerevisiae Msh2-Msh6 binding to MMR target sites was measured in solution by the fluorescence anisotropy of a 37 bp duplex DNA with a mismatch (G:T), a mismatch plus lesion (O\textsuperscript{6}Methylguanine:T), or a base insertion (+T) in the center and a TAMRA label at one end. In the equilibrium anisotropy binding experiments, 0.01 µM TAMRA-labeled DNA was titrated with Msh2-Msh6 (0 - 0.3 µM) in binding buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl\textsubscript{2}) at 25 °C. According to equilibrium binding data shown in Figure 3.1, anisotropy of G:T\textsubscript{TAMRA}, O\textsuperscript{6}MeG:T\textsubscript{TAMRA} and +T\textsubscript{TAMRA} DNA increases on Msh2-Msh6 addition and yields a dissociation constant ($K_D$) = 4.4, 3.3 and 28 nM, respectively, for the interaction; in contrast, there is barely any change in signal with fully matched G:C\textsubscript{TAMRA} DNA. The data are consistent with the reported high selectivity of Msh2-Msh6 for mismatches/IDLs/lesions over matched DNA (Gradia et al., 2000; Jiang et al., 2005;
In real time, Msh2-Msh6 binds rapidly to G:T, +T and O⁶MeG:T and dissociates more slowly from G:T and O⁶MeG:T than from +T.

Msh2-Msh6 binding to DNA was also monitored in real time by mixing the two reactants (0.8 µM Msh2-Msh6 and 0.12 µM TAMRA-labeled DNA) in a stopped-flow instrument. Final concentrations for these experiments are 0.4 µM Msh2-Msh6, 0.06 µM labeled DNA. The increase in fluorescence anisotropy reflected rapid association of Msh2-Msh6 with G:T\textsubscript{TAMRA}, O⁶MeG:T\textsubscript{TAMRA} and +T\textsubscript{TAMRA} DNA, and the data fit to a single exponential function yielded apparent binding rates, $k_{\text{observed}} = 3.8$ s\textsuperscript{-1}, 12 s\textsuperscript{-1}, and 7.1 s\textsuperscript{-1}, respectively; there was no detectable interaction with G:C\textsubscript{TAMRA} DNA (Figure 3.2A).

In complementary experiments, DNA dissociation was measured by mixing pre-formed Msh2-Msh6•DNA complexes (0.8 µM Msh2-Msh6 and 0.12 µM TAMRA-labeled DNA) with excess unlabeled G:T DNA trap (8 µM unlabeled G:T DNA trap) and monitoring decrease in anisotropy over time. Final concentrations for these experiments are 0.4 µM Msh2-Msh6, 0.06 µM labeled DNA, 4 µM unlabeled G:T trap. The data fit to a single exponential yielded $k_{\text{OFF}} = 0.013$ s\textsuperscript{-1}, 0.005 s\textsuperscript{-1} and 0.15 s\textsuperscript{-1} for G:T\textsubscript{TAMRA}, O⁶MeG:T\textsubscript{TAMRA} and +T\textsubscript{TAMRA} DNA, respectively (Figure 3.2B), indicating significant differences in the stabilities of these complexes (t\textsubscript{1/2} = 53, 138 and 4.6 s). Assuming that the kinetics reflect a one-step reversible interaction...
between Msh2-Msh6 and DNA, we calculated bimolecular binding constants, $k_{ON} = 1, 3$ and $1.7 \times 10^7 \text{ M}^{-1} \text{s}^{-1}$ for G:T_{TAMRA}, O6MeG:T_{TAMRA} and +T_{TAMRA} DNA, respectively ($k_{observed} = k_{ON}[\text{Msh2-Msh6}] + k_{OFF}$).

Thus, the transient kinetic data reveal that Msh2-Msh6 binds at similarly rapid rates to DNAs containing a mismatch (-/+ lesion) or base insertion but dissociates at different rates, resulting in complexes with long but varied lifetimes.

In contrast, Msh2-Msh6 binding to fully matched DNA is not detectable, likely because of fast dissociation (Msh2-Msh6 would diffuse over 37 bp within 0.6 ms based on a 0.25 $\mu$m$^2$ s$^{-1}$ diffusion coefficient) (Gerland et al., 2002; Gorman et al., 2007).

It should be noted that our in-solution method yields rate constants that are an order of magnitude faster than those obtained from surface-based methods such as SPR, despite the use of similar sized DNAs (Mazurek et al., 2009; Selmane et al., 2003); variables such as DNA surface density and mass transport effects may underlie slower kinetics in the latter case (Schuck and Minton, 1996).
Figure 3.1. Equilibrium fluorescence anisotropy measurements with TAMRA end-labeled DNAs show higher Msh2-Msh6 affinity for G:T\textsubscript{TAMRA} (dark blue), O\textsuperscript{6}MeG:T\textsubscript{TAMRA} (red) and lower affinity for +T\textsubscript{TAMRA} (green) ($K_D = 4.4 \pm 0.5, 3.3 \pm 0.5$ and $28 \pm 1.6$ nM, respectively), and almost no binding to G:C\textsubscript{TAMRA} (light blue).
Figure 3.1.
Figure 3.2. Msh2-Msh6 binds MMR targets rapidly and with varying affinities. (A) Kinetic measurements after mixing Msh2-Msh6 with DNA yield binding rate constants for the three target DNAs ($k_{ON} = 0.95, 3$ and $1.7 \times 10^7 \text{ M}^{-1} \text{s}^{-1}$, respectively), but no detectable binding to G:C TAMRA. (B) Mixing Msh2-Msh6•DNA$_{TAMRA}$ complexes with an unlabeled G:T trap yields $k_{OFF} = 0.013, 0.005$ and $0.15 \text{ s}^{-1}$, respectively.
Figure 3.2.

A Msh2-Msh6 binds rapidly to G:T, +T, O\textsuperscript{6}MeG:T in DNA

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3a}
\caption{Change in anisotropy over time for Msh2-Msh6 binding.}
\end{figure}

B Msh2-Msh6 dissociates more slowly from G:T and O\textsuperscript{6}MeG:T than +T

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3b}
\caption{Change in anisotropy over time for Msh2-Msh6 dissociation.}
\end{figure}
**Msh2-Msh6•DNA stability correlates with the inhibition of ATP hydrolysis.**

We have shown previously that in the absence of DNA one subunit of the MutS/Msh2-Msh6 dimer catalyzes rapid ATP hydrolysis and phosphate (Pi) release. Then a slow step in the reaction limits steady state turnover (i.e., Msh2-Msh6 exists predominantly in an ADP-bound state). On binding a mismatch/IDL, ATP hydrolysis is suppressed and Msh2-Msh6 persists in an ATP-bound state (Antony and Hingorani, 2003, 2004; Antony et al., 2006). Past reports from several research groups indicate that ATP-bound Msh2-Msh6 is responsible for signaling MMR (Gradia et al., 2000; Mazur et al., 2006; Selmane et al., 2003; Studamire et al., 1998). Given the above data on Msh2-Msh6 DNA-binding kinetics, we asked whether variation in Msh2-Msh6•DNA complex lifetimes has any impact on ATPase kinetics.

Pre-steady-state ATPase kinetics were measured by stopped-flow experiments using MDCC-labeled PBP (20 µM) to detect Pi release by Msh2-Msh6 (4 µM) on mixing with ATP (1 mM) with mixing ratio 1:1 in the absence or presence of DNA (12 µM) (Antony and Hingorani, 2004; Jacobs-Palmer and Hingorani, 2007). Figure 3.3 shows Msh2-Msh6 catalyzed Pi release in the absence and presence of target DNAs (Antony and Hingorani, 2004; Jacobs-Palmer and Hingorani, 2007). Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis at 1.2 s⁻¹ and the subsequent steady state turnover rate is 0.18 s⁻¹ (burst amplitude = 1.9 µM or 1 site per dimer). G:T, O⁶MeG:T and +T all suppress ATP hydrolysis but with differing efficiencies: O⁶MeG:T > G:T > +T (Figure 3.3). These data indicate a positive correlation between
Msh2-Msh6•DNA stability and the next step in MMR—formation of a ternary Msh2-Msh6•ATP•DNA complex.
Figure 3.3. Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis and Pi release \((k_{\text{burst}} = 1.2 \text{ s}^{-1})\) followed by steady state turnover \((k_{\text{cat}} = 0.18 \text{ s}^{-1})\). The target DNAs inhibit ATP hydrolysis to varying degrees.
Figure 3.3.
**Figure 3.4.** Msh2-Msh6 interactions with +T insertions in DNA. (A) Fluorescence anisotropy measurements of Msh2-Msh6 mixed with TAMRA end-labeled +T, +2T and +3T DNA yield binding rate constants $k_{ON} = 1.7$, 4.4 and $4 \times 10^7$ M$^{-1}$ s$^{-1}$, respectively. (B) Mixing Msh2-Msh6•DNA complexes with excess unlabeled G:T DNA trap yields $k_{OFF} = 0.15 \pm 0.002$ s$^{-1}$, $0.43 \pm 0.02$ s$^{-1}$ and $2 \pm 0.12$ s$^{-1}$ for the three DNAs, respectively. Final concentrations: 0.4 µM Msh2-Msh6, 0.06 µM labeled DNA, 4 µM unlabeled G:T trap.
Figure 3.4.

A Msh2-Msh6 binds rapidly to +T, +2T and +3T loops in DNA

\[ k_{\text{ON}} = 2 \times 4 \times 10^7 \text{ M}^{-1} \text{s}^{-1} \]

Change in anisotropy

Time, s

B Msh2-Msh6 dissociates at different rates from various +T loops

\[ \begin{align*}
  k_{\text{OFF}} & \text{s}^{-1} \\
  +T & 0.15 \\
  +2T & 0.4 \\
  +3T & 2 \\
\end{align*} \]

Change in anisotropy

Time, s
**Msh2-Msh6 interactions with +T, +2T and +3T insertions in DNA**

The question of how Msh2-Msh6 distinguishes IDLs in DNA is unclear. To answer this question, we designed 37 bp duplex TAMRA-labeled DNAs with a base insertion (+T), two base insertions (+2T) and three base insertions (+3T) to detect their binding interactions with Msh2-Msh6.

Fluorescence anisotropy measurements of Msh2-Msh6 mixed with TAMRA end-labeled +T, +2T and +3T yield binding rate constants $k_{ON} = 1.7, 4.4$ and $4 \times 10^7$ M$^{-1}$ s$^{-1}$, respectively (Figure 3.4A). In complementary experiments, mixing Msh2-Msh6•DNA complexes with excess unlabeled G:T DNA trap yields $k_{OFF} = 0.15 \pm 0.002$ s$^{-1}$, $0.43 \pm 0.02$ s$^{-1}$ and $2 \pm 0.12$ s$^{-1}$ for the +T, +2T and +3T DNAs, respectively (Figure 3.4B). The final concentrations for these experiments are 0.4 µM Msh2-Msh6, 0.06 µM labeled DNA, 4 µM unlabeled G:T trap.

Stopped-flow experiments measuring ATP hydrolysis and phosphate (Pi) release by Msh2-Msh6 were performed by mixing the protein -/+ DNA with ATP and MDCC-PBP. Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis and Pi release at $k_{burst} = 1.2$ s$^{-1}$ and amplitude 2 µM (1 site per Msh2-Msh6), followed by a linear, steady state phase at a rate of 0.36 µM s$^{-1}$ ($k_{cat} = 0.18$ s$^{-1}$). DNA binding to Msh2-Msh6 inhibits the burst of ATP hydrolysis to varying extents, with +T being the most effective ATPase inhibitor of the three (Figure 3.5). Moreover, +T in a different sequence context inhibits Msh2-Msh6 ATPase activity as effectively as G:T (Antony and Hingorani, 2003). Final concentrations are 2 µM Msh2-Msh6, 500 µM ATP, 6 µM DNA, 10 µM MDCC-PBP.
Analysis of larger IDLs, which are poor Msh2-Msh6 targets (Marsischky et al., 1996), supports that Msh2-Msh6•DNA complex stability underlies MMR signaling efficacy, as +2T and +3T-bound complexes exhibit shorter lifetimes ($k_{OFF} = 0.4$ and $2 \text{ s}^{-1}$) and correspondingly poor suppression of ATP hydrolysis (Figure 3.5).
Figure 3.5. Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis and Pi release at $k_{\text{burst}} = 1.2 \text{ s}^{-1}$ and amplitude 2 $\mu$M (1 site per Msh2-Msh6), followed by a linear, steady state phase at a rate of 0.36 $\mu$M s$^{-1}$ ($k_{\text{cat}} = 0.18 \text{ s}^{-1}$). DNA binding to Msh2-Msh6 inhibits the burst of ATP hydrolysis to varying extents, with +T being the most effective ATPase inhibitor of the three. Moreover, +T in a different sequence context inhibits Msh2-Msh6 ATPase activity as effectively as G:T (Antony and Hingorani, 2003).
Figure 3.5.

![Graph showing the effect of different DNA sequences on Pi concentration over time.](image)

- **ATP**
- **No DNA**
- $k_{\text{burst}} = 1.2 \text{ s}^{-1}$
- $k_{\text{cat}} = 0.18 \text{ s}^{-1}$

Parameters:
- $[\text{Pi}] \mu\text{M}$
- Time, s
2-Aminopurine serves as an on-site reporter of Msh2-Msh6 interactions with target sites.

Fluorescence anisotropy measurements with end-labeled DNA provide a generic report of Msh2-Msh6 binding, but do not reveal its specific location on DNA. In order to track Msh2-Msh6 more precisely, we used a 2-Aminopurine:Thymine base pair (2-Ap:T) as an on-site reporter of the interaction. 2-Ap is a fluorescent adenosine analogue that pairs with T in Watson-Crick geometry (Nordlund et al., 1989). 2-Ap:T has much lower stability (~ 0.5 kcal mol\(^{-1}\)) (Law et al., 1996) and, from base pair opening kinetics, a shorter lifetime (~ 6-fold) (Lycksell et al., 1987) than A:T. Many proteins, including DNA polymerase, substitute 2-Ap:T for A:T, although the efficiency may vary (e.g., T4 polymerase \(K_m\) is 25 \(\mu\)M for 2-Ap and 5 \(\mu\)M for A) (Clayton et al., 1979). Since 2-Ap fluorescence is exquisitely sensitive to the local environment and is quenched in duplex DNA (Jean and Hall, 2002), we positioned 2-Ap:T adjacent to G:T or G:C, anticipating that it would report Msh2-Msh6 binding to the site (Jacobs-Palmer and Hingorani, 2007; Nag et al., 2007).

2-Aminopurine reports Msh2-Msh6 binding to G:T and 2-Ap:T sites in DNA.

Msh2-Msh6 binding to 2-Ap DNA was measured on FluoroMax-3. 0.025 \(\mu\)M 2-Ap-labeled DNA was titrated with Msh2-Msh6 (0-0.5 \(\mu\)M) in binding buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl\(_2\)) at 25 °C. Equilibrium binding experiments show fluorescence intensity of G:T\(_2\)-Ap DNA increasing with Msh2-Msh6, and the isotherm yields a \(K_D\) of 13 nM (Figure 3.6), comparable with
G:T\textsubscript{TAMRA} DNA (Figure 3.1). Interestingly, we also observed Msh2-Msh6 binding to the 2-Ap:T site in G:C\textsubscript{2-Ap} DNA, albeit with a much lower affinity (Figure 3.6; $K_D = 70$ nM). Since we did not detect any Msh2-Msh6 binding to fully matched G:C\textsubscript{TAMRA} (Figure 3.1, 3.2A), it appears that the protein can distinguish 2-Ap:T from a Watson-Crick base pair, but not as effectively as G:T.


Binding kinetics, measured by monitoring increase in 2-Ap fluorescence over time after mixing 0.8 µM Msh2-Msh6 and 0.12 µM DNA, yielded similar association rates for G:T\textsubscript{2-Ap} and G:C\textsubscript{2-Ap} DNAs (Figure 3.7A; $k_{ON} \sim 2 \times 10^7$ M\textsuperscript{-1} s\textsuperscript{-1}).

In complementary experiments, DNA dissociation was measured by mixing pre-formed Msh2-Msh6•DNA complexes (0.8 µM Msh2-Msh6 and 0.12 µM 2-Ap-labeled DNA) with excess unlabeled G:T DNA trap (8 µM unlabeled G:T DNA trap) and monitoring decrease in 2-Ap fluorescence over time. Final concentrations for these experiments are 0.4 µM Msh2-Msh6, 0.06 µM labeled DNA, 4 µM unlabeled G:T trap. The dissociation rates, however, are strikingly different. Msh2-Msh6 releases the mismatch in G:T\textsubscript{2-Ap} slowly (Figure 3.7B; $k_{OFF} = 0.013$ s\textsuperscript{-1}; $t_{1/2} = 53$ s), consistent with G:T\textsubscript{TAMRA} anisotropy data (Figure 3.2B); +T\textsubscript{2-Ap} DNA, which contains a T insertion adjacent to 2-Ap:T, yields similar results ($k_{ON} = 1.8 \times 10^7$ M\textsuperscript{-1} s\textsuperscript{-1}) as +T\textsubscript{TAMRA} ($k_{ON} = 1.7 \times 10^7$ M\textsuperscript{-1} s\textsuperscript{-1}) as well (Figure 3.2A; Figure 3.14B). In contrast, Msh2-Msh6 release from 2-Ap:T in G:C\textsubscript{2-Ap} occurs ~ 90-fold faster (Figure
3.7B; $k_{\text{OFF}} = 1.2$ s$^{-1}$; $t_{1/2} = 0.6$ s). Thus, the on-site reporter confirms that differences in Msh2-Msh6 affinity for various sites in DNA result, apparently exclusively, from differences in Msh2-Msh6 dissociation from them.

**2-Ap:T can inhibit ATP hydrolysis, but less effectively than G:T.**

Figure 3.8 showed that free Msh2-Msh6 catalyzes a burst of ATP hydrolysis at 1.4 s$^{-1}$ and the subsequent steady state turnover rate is 0.22 s$^{-1}$. 2-Ap:T can inhibit ATP hydrolysis, but less effectively than G:T (Figure 3.8). Therefore, ATP hydrolysis is partially suppressed by G:C$_{2\text{-Ap}}$ DNA.

Moreover, partial inhibition of Msh2-Msh6 catalyzed ATP hydrolysis by G:C$_{2\text{-Ap}}$ DNA affirms that Msh2-Msh6•DNA complex stability correlates with formation of the ATP-bound ternary complex (Figure 3.8).
Figure 3.6. Equilibrium 2-Ap fluorescence measurements yield $K_D = 13 \pm 2$ nM for G:T$_{2-Ap}$ (dark blue) and 70 $\pm$ 7 nM for G:C$_{2-Ap}$ DNA (light blue).
Figure 3.6.
Figure 3.7. An on-site 2-Ap reporter confirms that Msh2-Msh6 dissociation rates underlie target site selectivity. (A) Kinetic measurements yield $k_{ON} = 2.4$ and $1.5 \times 10^7$ M$^{-1}$ s$^{-1}$, 2-Ap:T (purple dot) is located next to G:T or G:C, respectively and (B) $k_{OFF} = 0.013$ and $1.2$ s$^{-1}$ for G:T$_{2-Ap}$ and G:C$_{2-Ap}$, respectively.
Figure 3.7.

A  Msh2-Msh6 binds rapidly to both 2-Ap:T and G:T

B  Msh2-Msh6 releases 2-Ap:T faster than G:T
**Figure 3.8.** ATP hydrolysis is partially suppressed by G:C$_{2-Ap}$ DNA.
Figure 3.8.

[Diagram showing the reaction of ATP with DNA and the resulting burst and turnover rates.]

- ATP
- 2-6
- k_cat = 0.22 s^{-1}
- k_burst = 1.4 s^{-1}
- No DNA
- 2-Ap:T
- G:T

[Graph showing the concentration of Pi (µM) over time (s).]
Msh2-Msh6 moves on DNA in search of MMR target sites.

Since 2-Aminopurine is an effective on-site reporter for Msh2-Msh6 DNA binding/release, we could use it to find out how the protein locates particular sites on DNA. A DNA substrate was designed in which a 2-Ap:T base pair and G:T mismatch were positioned 7 bp apart, to allow only one Msh2-Msh6 binding per DNA, to either site (2-Ap:T/7bp/G:TAMRA); human Msh2-Msh6 contacts are detected 6 bp and 4 bp from G:T in the crystal structure (Warren et al., 2007). The substrate was also end-labeled with TAMRA, in order to detect Msh2-Msh6 binding both by fluorescence anisotropy (generic signal) and 2-Ap fluorescence (site-specific signal), and monitor transient interactions that facilitate the selection of one site over the other.

Anisotropy reports rapid Msh2-Msh6 binding to G:T or 2-Ap:T sites.

Kinetic anisotropy measurements show that Msh2-Msh6 binds both 2-Ap:T/7bp/G:TAMRA DNA and the control DNA with no G:T (2-Ap:T/7bp/G:C_TAMRA), at similar rapid rates (Figure 3.9; $k_{ON} \sim 1.5 \times 10^7$ M$^{-1}$ s$^{-1}$), consistent with previous data (Figure 3.2A, 3.7A). The lower amplitude of the 2-Ap:T/7bp/G:C_TAMRA trace is due to relatively rapid dissociation of Msh2-Msh6 from 2-Ap:T (Figure 3.7B), resulting in less Msh2-Msh6•DNA complex at equilibrium.
Figure 3.9. Fluorescence anisotropy shows Msh2-Msh6 binding mismatched 2-Ap:T/7bp/G:T TAMRA and matched 2-Ap:T/7bp/G:C TAMRA DNA rapidly ($k_{ON} \sim 1.5 \times 10^7 M^{-1} s^{-1}$).
Figure 3.9.

$2\rightarrow 6$

$2\text{-Ap:T/7bp/G:T}$

$2\text{-Ap:T/7bp/G:C}$

$k_{\text{ON}} \sim 1.5 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$
Figure 3.10. 2-Ap fluorescence yields the same rate ($k_{\text{ON}} \sim 1.6 \times 10^7 \, \text{M}^{-1} \, \text{s}^{-1}$) for 2-Ap:T/7bp/G:C_{TAMRA}, but 2-Ap:T/7bp/G:T_{TAMRA} shows biphasic kinetics.
Figure 3.10.

![Graph showing fluorescence change over time with annotations for 2-Ap:T/7bp/G:C and 2-Ap:T/7bp/G:T. The graph includes points at 0.25 s and a rate constant $k_{ON} = 1.6 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$. ]

2-Ap fluorescence measurements also show Msh2-Msh6 binding 2-Ap:T/7bp/G:C_TAMRA with apparent $k_{ON} = 1.6 \times 10^7$ M$^{-1}$ s$^{-1}$ (Figure 3.10). However, in case of G:T mismatch-containing 2-Ap:T/7bp/G:T_TAMRA, 2-Ap fluorescence increases rapidly and then decreases nearly to the baseline (Figure 3.10). One explanation for the biphasic kinetics is that any Msh2-Msh6 bound to the 2-Ap:T site releases it rapidly (Figure 3.7B; $t_{1/2} = 0.6$ s), and the G:T site on the DNA serves as an endogenous trap for the protein (Figure 3.7B; $t_{1/2} = 53$ s), preventing its return to 2-Ap:T.

The G:T site serves as an internal trap for Msh2-Msh6 released for 2-Ap:T.

To test this hypothesis, we designed a sequential mixing experiment that would monitor 2-Ap:T-bound Msh2-Msh6 over time. The protein was initially mixed with 2-Ap:T/7bp/G:T_TAMRA for 250 msec (peak fraction of Msh2-Msh6 bound to 2-Ap:T; Figure 3.10), followed by addition of excess unlabeled G:T DNA as an exogenous trap for free Msh2-Msh6. A similar experiment was performed in which only buffer was added to the reaction after 250 msec, in order to test whether the G:T site on 2-Ap:T/7bp/G:T_TAMRA serves as a comparable endogenous trap.

Data collected after the second mixing step show rapid decrease in 2-Ap fluorescence as Msh2-Msh6 releases 2-Ap:T and is trapped by G:T (in the absence of any G:T trap, Msh2-Msh6 rebinds 2-Ap:T and fluorescence increases; Figure 3.7B). The kinetics are similar whether G:T is present in cis (2-Ap:T/7bp/G:T_TAMRA) or in
*trans* (excess unlabeled G:T) (Figure 3.11; $k_{OFF} = 0.8 - 1.2 \text{ s}^{-1}$). These results confirm the hypothesis that after leaving 2-Ap:T, Msh2-Msh6 can be trapped by a G:T site present on the same DNA.

**Msh2-Msh6 moves from 2-Ap:T to G:T without dissociating from DNA.**

We could now query the mode of Msh2-Msh6 transfer from 2-Ap:T to G:T. To this end, the sequential mixing experiment was performed again by incubating Msh2-Msh6 with 2-Ap:T/7bp/G:T_{TAMRA} for 250 msec followed by the addition of the G:T trap, except this time the interaction was monitored by fluorescence anisotropy.

These data show that most of the bound Msh2-Msh6 dissociates at a slow rate from 2-Ap:T/7bp/G:T_{TAMRA} DNA (Figure 3.12; $k_{OFF} = 0.04 \text{ s}^{-1}$). An analogous experiment performed with 2-Ap:T/7bp/G:C_{TAMRA} confirms rapid Msh2-Msh6 dissociation in the absence of an internal G:T site (Figure 3.12; $k_{OFF} = 1 \text{ s}^{-1}$).

Thus, we can conclude that on releasing 2Ap:T (at ~ 1 s⁻¹), Msh2-Msh6 does not dissociate from DNA, rather, it moves laterally on the duplex and is trapped by the G:T mismatch, and then dissociates slowly as expected for the Msh2-Msh6•G:T complex (Figure 3.2B). These results provide the first direct evidence for Msh2-Msh6 movement on DNA en route to formation of a stable complex with a target site.
Figure 3.11. 2-Ap fluorescence reports rapid release of Msh2-Msh6 from 2-Ap:T in 2-Ap:T/7bp/G:T_{TAMRA} with an external trap (unlabeled G:T) or buffer added after 0.25 s ($k_{OFF} = 0.8 – 1.2 \text{ s}^{-1}$).
Figure 3.11.
**Figure 3.12.** Msh2-Msh6 moves from 2-Ap:T to G:T without dissociating from DNA. Corresponding TAMRA anisotropy shows slow release of Msh2-Msh6 from 2-Ap:T/7bp/G:TAMRA ($k_{OFF} = 0.044 \text{ s}^{-1}$).
Figure 3.12.
**Figure 3.13.** Msh2-Msh6 interactions with G:T<sub>2-AP</sub> and G:C<sub>2-AP</sub> DNAs in the absence and presence of nucleotide cofactors. (A) ADP-bound Msh2-Msh6 binds G:T<sub>2-AP</sub> and G:C<sub>2-AP</sub> rapidly, as in the absence of nucleotides ($k_{ON} = 3 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$). Pre-incubation of Msh2-Msh6 with ATPγS results in no detectable interaction with either DNA substrate. (B) ADP-bound Msh2-Msh6 releases G:T ($k_{OFF} = 0.04 \pm 0.0002 \text{ s}^{-1}$) and 2-AP:T ($k_{OFF} = 1.5 \pm 0.03 \text{ s}^{-1}$) rapidly, as in the absence of nucleotides. Mixing Msh2-Msh6•DNA complex with excess unlabeled G:T DNA trap and ATP results in faster Msh2-Msh6 release from G:T ($k_{OFF} = 0.4 \pm 0.004 \text{ s}^{-1}$) while the 2-AP:T release rate barely changes ($k_{OFF} = 2 \pm 0.03 \text{ s}^{-1}$). Final concentrations: 0.2 µM Msh2-Msh6, 0.03 µM 2-AP DNA, 250 µM nucleotide, 4 µM unlabeled G:T trap.
Figure 3.13.

A  Effect of nucleotide cofactors on Msh2-Msh6 binding to G:T and 2-Ap:T

\[ k_{\text{ON}} = 3 \times 10^7 \text{ M}^{-1} \text{ s}^{-1} \]

B  Msh2-Msh6 releases 2-Ap:T rapidly with or without nucleotide cofactors

\[ k_{\text{OFF}} = 0.4 \text{ s}^{-1} \]

\[ k_{\text{OFF}} = 0.04 \text{ s}^{-1} \]

\[ k_{\text{OFF}} = 1.5 - 2 \text{ s}^{-1} \]
Nucleotide cofactors appear to have minimal impact on Msh2-Msh6-DNA interactions during the search phase of MMR.

As noted earlier, MutS/Msh2-Msh6 ATPase kinetics suggests that the proteins exist in ADP-bound state when not bound to a mismatch/IDL (Antony and Hingorani, 2003, 2004; Gradia et al., 1997); therefore, we examined Msh2-Msh6 interactions with DNA in the presence of ADP. Under these conditions, Msh2-Msh6 binds at $k_{ON} \sim 3 \times 10^7 \text{M}^{-1}\text{s}^{-1}$ and dissociates at $k_{OFF} = 0.04 \text{s}^{-1}$, 0.35 s$^{-1}$ and 1.5 s$^{-1}$ from G:T, +T and 2-Ap:T, respectively (Table 3.1; Figure 3.13 & 3.14). Thus, both nucleotide-free and ADP-bound Msh2-Msh6 exhibit similar DNA binding kinetics.

Previous studies have also shown that ATP binding weakens MutS/Msh2-Msh6 interactions with DNA (Gradia et al., 1997), and it has been proposed that differential loss of affinity for matched versus mismatched DNA yields a net increase in Msh2-Msh6 specificity (Junop et al., 2001).

We find that pre-incubation of Msh2-Msh6 with ATP$_{\gamma}$S results in no detectable binding to either DNA (Table 3.1; Figure 3.13A & 3.14A). Mixing Msh2-Msh6•DNA with ATP and G:T trap leads to faster Msh2-Msh6 dissociation from G:T ($k_{OFF} = 0.4 \text{s}^{-1}$) but not +T ($k_{OFF} = 0.5 \text{s}^{-1}$) or 2-Ap:T ($k_{OFF} = 2 \text{s}^{-1}$) (Table 3.1; Figure 3.13B & 3.14B).

These data do not support the above hypothesis, in that ATP-bound Msh2-Msh6 does not have higher affinity for G:T over 2-Ap:T; however, ATP-induced specificity may manifest in later steps in the pathway. For example, the switching of ATP-bound Msh2-Msh6 to a conformation that diffuses away from a mismatch/IDL
is considered a key post-mismatch/IDL recognition step (Gradia et al., 1999; Mendillo et al., 2005). Previous studies indicate that this switch occurs specifically with Msh2-Msh6•ATP•DNA complexes formed at a mismatch (Mendillo et al., 2005).

The finding that ATP-bound Msh2-Msh6 releases both G:T and +T at similar rates (~ 0.5 s⁻¹) compared with 2-Ap:T (2 s⁻¹) supports the possibility of a common Msh2-Msh6 conformation triggered by interaction with an MMR target site.
Figure 3.14. Msh2-Msh6 interaction with +T DNA in the absence and presence of nucleotide cofactors. (A) Msh2-Msh6 binds +T_{2,Ap} at $k_{ON} = 1.8 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ in the absence of nucleotides and $3.7 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ in the presence of ADP; no DNA binding is detectable when Msh2-Msh6 is pre-incubated with ATPγS. (B) Mixing Msh2-Msh6•+T_{2,Ap} complex with excess unlabeled G:T DNA -/+ nucleotides yields $k_{OFF} = 0.13 \pm 0.002\text{ s}^{-1}$ (no nucleotide), $0.35 \pm 0.004\text{ s}^{-1}$ (ADP), and $0.49 \pm 0.007\text{ s}^{-1}$ (ATPγS). Final concentrations: 0.2 µM Msh2-Msh6, 0.03 µM 2-Ap DNA, 250 µM nucleotide, 4 µM unlabeled G:T trap.
Figure 3.14.

A Effect of nucleotide cofactors on Msh2-Msh6 binding to +T

2-Ap:T/+T

\[ k_{ON} = 2 \times 4 \times 10^7 \, \text{M}^{-1} \, \text{s}^{-1} \]

Time, s

2-Ap fluorescence change

B Msh2-Msh6 release from +T is stimulated slightly by nucleotides

2-Ap:T/+T

ADP

No nucleotide

ATPγS

Time, s

2-Ap fluorescence change

ADP

ATP

No nucleotide

\[ k_{OFF} = 0.35 \, \text{s}^{-1} \]

\[ k_{OFF} = 0.5 \, \text{s}^{-1} \]

\[ k_{OFF} = 0.13 \, \text{s}^{-1} \]
Table 3.1. Kinetic parameters for Msh2-Msh6–DNA interactions.

<table>
<thead>
<tr>
<th>Protein (nucleotide)</th>
<th>$k_{ON}$ ($\times 10^7$ M$^{-1}$ s$^{-1}$)</th>
<th>$k_{OFF}$ (s$^{-1}$)</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Msh2-6 (none)</td>
<td>1 1.7 1.5</td>
<td>0.013 0.15 1.2</td>
<td>1.3 8.8 80</td>
</tr>
<tr>
<td>Msh2-6 (ADP)</td>
<td>3 3.7 3</td>
<td>0.04 0.35 1.5</td>
<td>1.3 9.4 50</td>
</tr>
<tr>
<td>Msh2-6 (ATP/ATPgS)</td>
<td>— — —</td>
<td>0.4 0.5 2</td>
<td>— — —</td>
</tr>
<tr>
<td>Msh2-6E339A (none)</td>
<td>5.7 2 6.8</td>
<td>0.06 0.13 0.14</td>
<td>1 6.5 2</td>
</tr>
</tbody>
</table>
**Impact of Msh6 Glu339 on recognition of MMR target sites.**

Glutamate in the conserved Phe-X-Glu DNA binding motif of MutS/Msh2-Msh6 proteins form hydrogen bonds with the mis-paired or inserted base in DNA (Lebbink et al., 2006; Obmolova et al., 2000; Warren et al., 2007). We assessed the contribution of *S. cerevisiae* Msh6 Glu339 to mismatch recognition by examining an alanine mutant, Msh2-Msh6_{E339A}.

Figure 3.15A showed that the Msh2-Msh6_{E339A} mutant binds G:T and 2-Ap:T at slightly faster rates than wild type Msh2-Msh6 ($k_{ON} = 6 \times 10^7$ M$^{-1}$ s$^{-1}$). More significant differences manifest during DNA dissociation as Msh2-Msh6_{E339A} releases 2-Ap:T at ~7-fold slower rate ($k_{OFF} = 0.14$ s$^{-1}$) and G:T at ~4-fold faster rate ($k_{OFF} = 0.06$ s$^{-1}$) than Msh2-Msh6; +T binding kinetics appear to be unaffected (Table 3.1; Figure 3.15B).

Free Msh2-Msh6_{E339A} catalyzes a burst of ATP hydrolysis at 1.4 s$^{-1}$ and $k_{cat}$ 0.25 s$^{-1}$ like the wild type Msh2-Msh6; however, in case of the mutant, 2-Ap:T binding inhibits ATP hydrolysis as strongly as G:T (Figure 3.16), consistent with the longer lifetime of the Msh2-Msh6_{E339A}•2-Ap:T complex.

These data indicate that Msh6 Glu339 affects the stability of Msh2-Msh6•DNA complexes and thereby influences the selection of MMR target sites in DNA.
**Figure 3.15.** E339A mutation in Msh6 affects Msh2-Msh6 interactions with G:T and 2-Ap:T, but not +T in DNA. (A) Fluorescence anisotropy measurements of Msh2-Msh6<sub>E339A</sub> mixed with DNA yield binding rate constants $k_{ON} = 5.7, 6.8$ and $2 \times 10^7$ M$^{-1}$ s$^{-1}$ for G:T<sub>TAMRA</sub>, 2-Ap:T/G:C<sub>TAMRA</sub>, and +T<sub>TAMRA</sub> DNAs, respectively. (B) Mixing the Msh2-Msh6<sub>E339A</sub>-DNA complex with unlabeled G:T DNA trap yields $k_{OFF} = 0.057 \pm 0.0005$ s$^{-1}$ for G:T, $0.14 \pm 0.002$ s$^{-1}$ for 2-Ap:T, and $0.13 \pm 0.002$ s$^{-1}$ for +T DNA. Final concentrations: 0.4 µM Msh2-Msh6, 0.06 µM G:T<sub>TAMRA</sub>, 4 µM unlabeled G:T trap.
Figure 3.15.

A. Msh2-Msh6 \textsubscript{E339A} binds rapidly to G:T, +T and 2-Ap:T DNA

\[ k_{ON} = 2 \times 6 \times 10^7 \text{ M}^{-1} \text{s}^{-1} \]

B. Msh2-Msh6 \textsubscript{E339A} releases G:T, +T and 2-Ap:T DNA at similar rates

\[ k_{OFF} \]
Figure 3.16. Like wild-type Msh2-Msh6, Msh2-Msh6<sub>339A</sub> catalyzes a burst of ATP hydrolysis and Pi release at \( k_{\text{burst}} = 1.4 \text{ s}^{-1} \) and amplitude 2 \( \mu \text{M} \) (one site per Msh2-Msh6), followed by a linear, steady state phase at a rate of 0.5 \( \mu \text{M s}^{-1} \) \( (k_{\text{cat}} = 0.25 \text{ s}^{-1}) \) (Antony and Hingorani, 2003). In case of wild-type Msh2-Msh6, G:T completely suppresses the burst ATP hydrolysis, stabilizing the complex in an ATP-bound state, and 2-Ap:T suppresses the burst partially. In case of Msh2-Msh6<sub>339A</sub>, both G:T and 2-Ap:T DNA are similarly strong suppressors of ATP hydrolysis, while +T has similar partial effect as on wild-type Msh2-Msh6. Final concentrations: 2 \( \mu \text{M} \) Msh2-Msh6, 500 \( \mu \text{M} \) ATP, 6 \( \mu \text{M} \) DNA, 10 \( \mu \text{M} \) MDCC-PBP.
Figure 3.16.

A Msh2-Msh6 ATP hydrolysis is suppressed by G:T and to lesser extent by +T and 2-Ap:T

\[
\text{2-6 ATP} \quad k_{\text{cat}} = 0.22 \text{ s}^{-1}
\]

\[
k_{\text{burst}} = 1.4 \text{ s}^{-1}
\]

B Msh2-Msh6_{\text{E339A}} ATP hydrolysis is suppressed similarly by G:T and 2-Ap:T

\[
\text{2-6 ATP} \quad k_{\text{cat}} = 0.25 \text{ s}^{-1}
\]

\[
k_{\text{burst}} = 1.4 \text{ s}^{-1}
\]
Chapter IV

Discussion

DNA mismatch repair (MMR) is a process to identify and correct a variety of base pair mismatching and insertion/deletion (IDL) that appear in DNA due to replication and recombination errors (Kunkel and Erie, 2005). MutS protein in prokaryotes or MutS homologues in eukaryotes (e.g., Msh2-Msh6) can recognize and bind mismatch sites to initiate mismatch repair (Kunkel and Erie, 2005). This crucial, initial step in MMR requires that MutS/Msh2-Msh6 employ efficient strategies to interrogate base pairs and recognize a broad spectrum of discrepancies in the structure and/or dynamics of the double helix. But, the mechanism of how MutS/Msh2-Msh6 proteins interact with DNA and recognize their target sites is still unclear.

In order to decipher this intriguing unsolved question, we developed two stopped-flow assays -- 2-Ap fluorescence assay and fluorescence anisotropy assay -- to detect Msh2-Msh6 arrival at and departure from particular sites on DNA. We also designed mismatch (G:T), mismatch plus lesion (O6MeG:T) and IDLs (+T, +2T and +3T) DNA to compare the selectivity of Msh2-Msh6 binding. Comparison of the kinetic parameters reveals that Msh2-Msh6 makes rapid, transient contact with potential target sites in DNA during the search phase of MMR. Msh2-Msh6 pauses for long periods at strong sites such as G:T or O6MeG:T and pauses for short times at other sites that have altered structure or dynamics such as 2-Ap:T. On releasing the 2-
Ap:T site, Msh2-Msh6 is able to move laterally on DNA to locate a nearby G:T site. ATP can form an ATP-bound clamp with long-lived Msh2-Msh6-G:T complex to trigger the next step in MMR.

**Binding and recognition mechanism of Msh2-Msh6**

*Equilibrium data reveal that Msh2-Msh6 has different selectivity on DNA*

To address the question how *S. cerevisiae* Msh2-Msh6 recognizes target sites in DNA and signals initiation of MMR, we performed Msh2-Msh6 equilibrium binding experiments. Msh2-Msh6 binds 37mer G:TAMRA, O\(^6\)MeG:TAMRA, +T TAMRA and G:C\(_{2-}\)Ap DNA and yields \(K_D\) = 4.4, 3.3, 28 and 70 nM, respectively. From \(K_D\) values, Msh2-Msh6 binds G:T and O\(^6\)MeG:T with similar highest affinity, ~ 10-fold lower affinity for +T DNA, and ~ 20-fold much lower affinity for G:C\(_{2-}\)Ap DNA. There is barely any change in signal with fully matched G:C\_TAMRA DNA. Msh2-Msh6 discloses high selectivity for unmatched DNA over fully matched DNA and shows preference for a mismatched base pair over a base insertion. Also Msh2-Msh6 can distinguish 2-Ap:T from Watson-Crick base pair.

Both G:T and O\(^6\)MeG:T are well-known targets for Msh2-Msh6 (Hsieh and Yamane, 2008; Jiricny, 2006). A single molecule method called unzipping force analysis has been used to measure the binding affinity between *S. cerevisiae* Msh2-Msh6 and 1.1 kb duplex DNA with G:T mismatch and +1 (G insertion) (Jiang et al., 2005). The \(K_D\) obtained for G:T was 5.7 nM and from +1 (G insertion) was 9.8 nM (Jiang et al., 2005) which is consistent with our data and proved the strong preference...
of Msh2-Msh6 for G:T mismatch. Previous research also shows that human Msh2-Msh6 binds O\textsuperscript{6}MeG:T DNA with high affinity (Duckett et al., 1996).

Until now, no structure of MutS/Msh2-Msh6 bound to fully matched DNA has been reported, so whether and how Msh2-Msh6 binds fully matched DNA is still a puzzle. In early studies using gel mobility shift assays, the $K_D$ for matched DNA was found to be 10-fold to 30-fold higher than mismatched DNA (Marsischky and Kolodner, 1999), whereas recently in unzipping force analysis it was several-thousand-fold higher (Jiang et al., 2005). This difference is likely because the single molecule method can directly measure Msh2-Msh6 interaction with the mismatch site, while mobility shift method cannot distinguish mismatch-specific from non-specific Msh2-Msh6-DNA interactions. Such weak interactions suggest very high selectivity of Msh2-Msh6 for target sites, but the mechanism by which the protein differentiates between matched DNA and target sites still remains to be determined.

It is also interesting to note that *T. aquaticus* MutS and *S. cerevisiae* Msh2-Msh6 have different DNA binding specificities. A single molecule method using atomic force microscopy (AFM) revealed that *T. aquaticus* MutS binds +T with higher affinity than G:T mismatch (Yang et al., 2005). *T. aquaticus* MutS can not recognize 2-Ap DNA as a mismatch, unlike Msh2-Msh6 (Jacobs-Palmer and Hingorani, 2007). 2-Ap can pair with thymine in canonical Watson-Crick base-pairing interactions (Nordlund et al., 1989) and 2-Ap:T structure is very similar to A:T. The only difference is the location of the amine group that is on carbon 6 in adenine and carbon 2 in 2-Ap and the 2-Ap:T base pair is a little less stable than A:T.
These data manifest that *T. aquaticus* MutS is less sensitive to mismatched base pairs than *S. cerevisiae* Msh2-Msh6 and has different binding specificities, especially a preference for base IDL.

The equilibrium data are also consistent with $K_D$ values that were calculated from kinetic on rates and off rates ($K_D = k_{OFF} / k_{ON}$). According to these measurements, Msh2-Msh6 binds 37mer $G:T_{TAMRA}$, $O^6$MeG:$T_{TAMRA}$, $+T_{TAMRA}$ and $G:C_{2-Ap}$ DNA and yields $K_D = 1.3, 0.17, 8.8$ and $80$ nM, respectively (Table 3.1). Thus the equilibrium and kinetic measurements are consistent with each other.

*Msh2-Msh6 forms complexes with mismatch, lesion, IDL DNA at different efficacy levels*

To further explore how *S. cerevisiae* Msh2-Msh6 recognizes target sites in DNA and signals initiation of MMR, we monitored Msh2-Msh6 binding with DNA in real time on the stopped-flow instrument. From kinetic stopped-flow experiments, Msh2-Msh6 binds to 37mer $G:T_{TAMRA}$, $O^6$MeG:$T_{TAMRA}$, $+T_{TAMRA}$, $+2T_{TAMRA}$, $+3T_{TAMRA}$ and $G:C_{2-Ap}$ DNA at similarly rapid rates: $k_{ON} = 1, 3, 1.7, 4.4, 4$ and $1.5 \times 10^7$ M$^{-1}$ s$^{-1}$, but dissociates at different rates. The dissociation rates for 37mer $G:T_{TAMRA}$, $O^6$MeG:$T_{TAMRA}$, $+T_{TAMRA}$, $+2T_{TAMRA}$, $+3T_{TAMRA}$ and $G:C_{2-Ap}$ DNA are $k_{OFF} = 0.013, 0.005, 0.15, 0.4, 2$ and $1.2$ s$^{-1}$, respectively (Table 3.1). There is no binding detected between Msh2-Msh6 and fully matched DNA. Fluorescence anisotropy measurements can provide a generic report of Msh2-Msh6 binding with DNA. 2-Ap fluorescence measurement is an excellent site-specific assay to report Msh2-Msh6 binding with DNA. These two assays are complementary for each other.
From 2-Ap kinetic experiments, Msh2-Msh6 binds to G:T<sub>2-Ap</sub> and +T<sub>2-Ap</sub> with similar rapid on rates (k<sub>ON</sub> = 2.4, and 1.8 x 10<sup>7</sup> M<sup>-1</sup> s<sup>-1</sup>) comparing with G:T<sub>TAMRA</sub> and +T<sub>TAMRA</sub> DNA. The dissociation rates and half-life time is also consistent: k<sub>OFF</sub> = 0.013, 0.13 s<sup>-1</sup> and t<sub>1/2</sub> = 53, 5.3 s (Table 3.1). The consistencies of the data from the two assays reveal their high quality and accuracy. Not only generic anisotropy measurement but also site-specific 2-Ap measurement can provide precise quantitative kinetic data.

As a conclusion, all DNAs tested were bound at similarly fast rates, but the dissociation rates vary over a hundred-fold such that G:T and O<sup>6</sup>MeG:T form the most stable complexes while 2-Ap:T and +3T form the least stable complexes with Msh2-Msh6. The differences in dissociation rates suggest that recognition involves formation of an initial complex whose stability determines whether the site is a target for MMR.

The half-life time for 37mer G:T<sub>TAMRA</sub>, O<sup>6</sup>MeG:T<sub>TAMRA</sub>, +T<sub>TAMRA</sub>, +2T<sub>TAMRA</sub>, +3T<sub>TAMRA</sub> and G:C<sub>2-Ap</sub> DNA are t<sub>1/2</sub> = 53, 138, 4.6, 1.7, 0.35, 0.6 s, respectively. Msh2-Msh6 can pause to form stable long lived complexes with sites such as G:T or O<sup>6</sup>MeG:T, and short lived complexes with IDLs or 2Ap:T DNA. The formation of long-lived complexes of Msh2-Msh6 with strong MMR targets (G:T, O<sup>6</sup>MeG:T) and short-lived complexes with weak MMR targets (+3T, 2-Ap:T) suggests that the stability of interaction with Msh2-Msh6 is a defining feature of sites that trigger MMR.
Msh2-Msh6 forms complexes with +T, +2T and +3T and the stability is +T > +2T > +3T and no binding for +4T. As mentioned in the introduction, in *S. cerevisiae*, six MutS homologs have been identified. Msh2 and Msh6 can form a Msh2-Msh6 heterodimeric complex and Msh2-Msh3 is the alternate complex formed. Both Msh2-Msh6 and Msh2-Msh3 can recognize IDL mismatches. There is overlap between Msh2-Msh6, which processes 1–2 base loops, and Msh2-Msh3, which processes 1–2 base and longer loops in DNA (Marsischky et al., 1996). This separation of function between Msh2-Msh6 and Msh2-Msh3 may be the cause of the shorter life spans of +2T and +3T-bound complexes and low efficiency between IDL DNAs and Msh2-Msh6.

A single-molecule optical microscopy study showed that Msh2-Msh6 can slide along the λ DNA (48,502 bp) by one-dimension diffusion (Gorman et al., 2007). Our kinetic assays could not detect any binding between Msh2-Msh6 with fully matched DNA. This may be caused by the very fast dissociation of Msh2-Msh6 on fully matched DNA. Msh2-Msh6 would diffuse off 37 bp within 0.6 ms based on a 0.25 µm² s⁻¹ diffusion coefficient derived from a recent single-molecule study (Gorman et al., 2007). The minimum time for the fluorescence anisotropy assay to detect the fluorescent change is 1 ms. Therefore, 0.6 ms is not enough time to be detected on a short piece of DNA.

According to IAsys (resonant mirror biosensor) total internal reflectance (TIR) measurements performed by other research groups, *E. coli* MutS binds 37mer +T heteroduplex fast (k_{ON} = 2.2 \times 10^5 M^{-1} s^{-1}) and dissociates slowly (k_{OFF} = 7.8 \times 10^{-3}
s\(^{-1}\)) (Selmane et al., 2003) and human Msh2-Msh6 associates 41mer G:T mismatch rapidly \((k_{ON} = 7.5 \times 10^5 \text{ M}^{-1} \text{s}^{-1})\) and dissociates slowly \((k_{OFF} = 1.4 \times 10^{-3} \text{ s}^{-1})\) (Mazurek et al., 2009). Compared to these constants, our in-solution method rate constants are an order of magnitude faster than surface-based methods. It has been reported that the resonant mirror biosensor binding experiment that uses biosensors to obtain kinetic parameters can generate up to four orders of magnitude variability compared with data from other assays. Therefore, the kinetic parameters obtained from such methods should be tested and validated by other methods (Schuck and Minton, 1996). Mass transport effects may be the cause of the slightly slower constant obtained by TIR measurement (Schuck and Minton, 1996).

**The searching mechanism of Msh2-Msh6**

Our kinetic data show that Msh2-Msh6 forms long- and short-lived complexes at different sites on DNA. But the mechanism of how Msh2-Msh6 finds these target sites in DNA remains ambiguous.

The single molecule study of *S. cerevisiae* Msh2-Msh6 movement on \(\lambda\) DNA visualized fluorescently labeled Msh2-Msh6 by total internal reflection microscopy (Gorman et al., 2007). The data indicated that Msh2-Msh6 could slide on the DNA helical axis by one dimension diffusion with an estimated diffusion coefficient of about 0.25 \(\mu\text{m}^2 \text{s}^{-1}\). It was proposed that such lateral movement could enable Msh2-Msh6 to scan the duplex, perhaps in contact with the replisome (e.g., bound to PCNA), and that sites with higher local flexibility or altered structures could trap the
protein and form specific recognition complexes. This study developed a model of Msh2-Msh6 actions during the search phase of MMR that could account for its ability to rapidly locate non-Watson-Crick sites in DNA and signal an appropriate response. Thus far, however, there is not enough evidence at base pair-level resolution to address essential features of the model. As noted earlier, there is no structure of MutS/Msh2-Msh6 bound to fully matched DNA. The single molecule data were at low resolution (~ 300 base pairs), and the λ DNA was not engineered with MMR target sites (Gorman et al., 2007); therefore, critical information on non-specific interactions between Msh2-Msh6 and DNA and the transition to specific Msh2-Msh6•DNA recognition complexes remains unknown.

In order to test if the model is accurate or not and to further explore the problem, we designed 2-Ap:T/7bp/G:T TAMRA, a DNA strand that generates two different Msh2-Msh6 target sites (2-Ap:T and G:T) and is labeled by both TAMRA and 2-Ap, but allows binding of only one Msh2-Msh6 per DNA. Both TAMRA fluorescence anisotropy (generic signal) and 2-Ap fluorescence (site-specific signal) were used to detect how Msh2-Msh6 selects one site over the other and how Msh2-Msh6 moves on DNA during recognition.

Fluorescence anisotropy shows Msh2-Msh6 binding mismatched 2-Ap:T/7bp/G:T TAMRA and matched 2-Ap:T/7bp/G:C TAMRA DNA rapidly ($k_{ON} \sim 1.5 \times 10^7 \text{ M}^{-1} \text{s}^{-1}$). 2-Ap fluorescence yields the same rate for 2-Ap:T/7bp/G:C TAMRA, but 2-Ap:T/7bp/G:T TAMRA shows biphasic kinetics which is direct evidence of Msh2-Msh6 dissociation from the 2-Ap:T site. These results generate the hypothesis that after
leaving 2-Ap:T site, Msh2-Msh6 can be trapped by a G:T mismatch site on the same DNA strand. Thus the anisotropy signal remains high. Experiments performed to directly measure off rate with 2-Ap fluorescence shows rapid release of Msh2-Msh6 from the 2-Ap:T site in 2-Ap:T/7bp/G:T{TAMRA} with either an external trap (unlabeled G:T) or a buffer added after 0.25 s incubation ($k_{off} = 0.8 \sim 1.2$ s$^{-1}$). An control experiment performed with 2-Ap:T/7bp/G:T{TAMRA} confirms rapid Msh2-Msh6 dissociation from the 2-Ap:T site in the absence of an internal G:T site ($k_{off} = 1$ s$^{-1}$). However, corresponding TAMRA anisotropy shows slow release of Msh2-Msh6 from 2-Ap:T/7bp/G:T{TAMRA} ($k_{off} = 0.044$ s$^{-1}$). These results confirm the hypothesis proposed by Gorman et al. by showing that Msh2-Msh6 slides along the DNA to search mismatched bases after releasing the 2-Ap:T site, while not dissociating from DNA. Based on our findings, when Msh2-Msh6 scans DNA for discrepancies, it pauses at potential locations for mismatch repair. If it stops for a longer time, there is higher likelihood that the site is a target for MMR, and Msh2-Msh6 can recruit other proteins to the site.

One possible question related to our experimental design is whether we can increase the 7 bp distance between 2-Ap:T and G:T sites on DNA. The answer is that we could not. This is because if there is too much distance between the two sites, more than one Msh2-Msh6 will bind the DNA and we will not be able to distinguish the behavior of one Msh2-Msh6 binding to either the 2-Ap:T or G:T sites. A recent human Msh2-Msh6 crystal structure with DNA containing a G:T revealed that Msh2-Msh6 contacts beyond the mismatch, and that Msh2 binds DNA up to a distance of 6
bp and Msh6 binds DNA up to a distance of 4 bp (Warren et al., 2007). This means that the maximum distance between 2-Ap:T and G:T is 7 bp if we want only one Msh2-Msh6 binding per DNA.

The other question related to the experimental design is whether we can increase the length of 2-Ap:T/7bp/G:T TAMRA DNA. Anisotropy assay has its limits for the DNA length to yield quality data for binding experiments. Fluorescence anisotropy can be useful when there is a significant signal difference between an unbound and bound state, for example when a small fluorescent molecule binds a larger molecule. When DNA length is increased, the fluorescence anisotropy of the free DNA is higher and the difference between the bound and unbound state is much smaller, reducing data quality.

**Msh2-Msh6 DNA recognition coupled with ATPase activity**

From kinetic data, we generated a hypothesis that Msh2-Msh6 scans DNA for discrepancies and as it scans, it pauses at potential sites for MMR where the base pairs may be less stable or have a distorted structure, and a longer pause increases the likelihood of signaling MMR.

We tested this hypothesis by measuring the effects of different DNA substrates on Msh2-Msh6 ATPase activity and the effect of different nucleotides on Msh2-Msh6 DNA binding activity.

**Effects of nucleotides on Msh2-Msh6 binding activity**
To test the effects of nucleotides on Msh2-Msh6 binding activity, we added ADP, ATPγS and ATP into the DNA binding kinetic experiments. In the presence of ADP, Msh2-Msh6 binds at $k_{ON} \sim 3 \times 10^7$ M$^{-1}$ s$^{-1}$ and dissociates at $k_{OFF} = 0.04$ s$^{-1}$, 0.35 s$^{-1}$ and 1.5 s$^{-1}$ from G:T, +T and 2-Ap:T, respectively (Table 3.1). Therefore ADP-bound Msh2-Msh6 does not interfere with its DNA binding activity. In the presence of ATPγS, Msh2-Msh6 had no detectable binding to any DNA. In the presence of ATP, Msh2-Msh6 obtained faster dissociation from G:T ($k_{OFF} = 0.4$ s$^{-1}$) but not +T ($k_{OFF} = 0.5$ s$^{-1}$) or 2-Ap:T ($k_{OFF} = 2$ s$^{-1}$) (Table 3.1). In conclusion, both nucleotide-free and ADP-bound Msh2-Msh6 have similar DNA binding kinetics; ATPγS inhibits the binding between Msh2-Msh6 and mismatch/IDL DNA. ATP-bound Msh2-Msh6 has weaken interaction with the mismatch site.
**Figure 4.1.** A model for Msh2-Msh6 actions during the search phase of MMR. The schematic represents nucleotide free or ADP-bound Msh2-Msh6 making rapid, transient contact with Watson-Crick base pairs, and pausing at possible sites of discrepancies in the double helix. Prolonged pauses favor formation of the Msh2-Msh6•ATP•DNA complex that signals MMR (thicker lines indicate preferred events in the pathway).
Figure 4.1.
Based on these data, model for Msh2-Msh6 actions during the search phase of MMR was generated. In this model, nucleotide-free or ADP-bound Msh2-Msh6 can make rapid, transient contact with potential target sites in DNA during the search phase of MMR. It pauses for long period at strong sites such as G:T or O\(^6\)MeG:T, which facilitates the formation of ATP-bound ternary complexes that are hypothesized to signal MMR. Msh2-Msh6 also pauses briefly at other sites that have altered structure and dynamics, which may trigger MMR as well, albeit less effectively (Figure 4.1). Such promiscuity could be an unavoidable side effect of the requirement that Msh2-Msh6 recognize a broad range of defects in DNA. It has been proposed that MutS/Msh2-Msh6-DNA interaction is energetically favorable at sites that simply display increased local flexibility or bendability (Isaacs and Spielmann, 2004; Mazurek et al., 2009). Indeed, S. cerevisiae Msh2-Msh6 is trapped at multiple sites on \(\lambda\) DNA that does not contain any known MMR target sites (Gorman et al., 2007). Perhaps 10 \sim 100\textsuperscript{-fold} higher stability of Msh2-Msh6 interactions with mismatches/IDLs/lesions over other sites limits promiscuous MMR to acceptably low levels, and it remains possible that further selectivity for MMR targets is imposed by other proteins in the MMR pathway such as MutL/Mlh1-Pms1.

MutS/Msh2-Msh6 proteins have an ATPase activity that is essential for MMR. These proteins are members of the ATP binding cassette (ABC) family. Msh2-Msh6 has Walker A (ATP-binding P loop) and Walker B (Mg\(^{2+}\) -binding) motifs to catalyze ATP binding and hydrolysis (Antony and Hingorani, 2003, 2004; Antony et al., 2006; Iyer et al., 2006; Kunkel and Erie, 2005). In S. cerevisiae Msh2-Msh6
complex, ATP binding and hydrolysis can cause conformational change (Studamire et al., 1998) which may be coupled with DNA binding or other protein interaction. In gel mobility shift experiments, when no nucleotides are added, human Msh2-Msh6 forms stable complex with DNA; when ATP/ATPγS was added, human Msh2-Msh6 changes conformation and appears to slide off the edge of short linear duplex DNA (Gradia et al., 1997; Schofield et al., 2001a). Other studies of *E. coli*, *S. cerevisiae*, and human mismatch repair proteins, suggest that MutS/Msh2-Msh6 bound to DNA could form a complex with ATP, MutL/Mlh1-Pms1 in MMR (Grilley et al., 1989; Habraken et al., 1998; Schofield et al., 2001b). Our data are consistent with these observations, since Msh2-Msh6 forms a relatively long-lived complex at a mismatch, which may be sufficient for interaction with other proteins, and ATP binding weakens the interaction with the mismatch, which may allow the complex slide away from the mismatch and initiate MMR.

We have also investigated *T. aquaticus* MutS ATPase activity and its coupling with DNA binding and mismatch recognition. For *T. aquaticus* MutS, ADP-bound MutS can make highly dynamic interactions with DNA, whereas ATP-bound MutS forms a stable complex with mismatch DNA and dissociates slowly off the mismatch (Jacobs-Palmer and Hingorani, 2007). Compared with our Msh2-Msh6 data, ADP has a similar effect on both MutS and Msh2-Msh6; while the ATP dissociation rate is different. For Msh2-Msh6, the ATP-bound Msh2-Msh6 complex has a much faster off rate compared to either nucleotide-free or ADP-bound Msh2-Msh6-DNA complex. These data show that *T. aquaticus* MutS and *S. cerevisiae* Msh2-Msh6
ATPase may have a slightly different ATP coupling process in MMR. Further studies with addition of MutL/Pms1-Mlh1 in the reaction will be more informative about the complete MutS/Msh2-Msh6 reaction.

**DNA effects on ATPase activity of Msh2-Msh6**

Several past reports indicate that the MutS/Msh2-Msh6•ATP•DNA complex is important for the initiation of MMR (Gradia et al., 2000; Mazur et al., 2006; Selmane et al., 2003; Studamire et al., 1998).

In the absence of DNA, one subunit of Msh2-Msh6 dimer catalyzes rapid ATP hydrolysis and phosphate release, then a slow step in the reaction limits steady-state turnover (Antony and Hingorani, 2003, 2004; Antony et al., 2006). Free Msh2-Msh6 catalyzes a burst of ATP hydrolysis at 1.2 s\(^{-1}\) and the subsequent steady state turnover rate is 0.18 s\(^{-1}\). Previous studies have also shown that binding of G:T mismatch DNA inhibits ATP hydrolysis (not ATP binding). To determine whether differences in the life times of Msh2-Msh6•DNA complexes has any impact on ATPase kinetics, we used pre-steady-state experiments to measure the effect of DNA binding on ATP hydrolysis and phosphate release. According to these data, DNA suppresses ATP hydrolysis to varying degrees; O\(^6\)MeG:T is a robust inhibitor like G:T, whereas the +T, +2T, +3T loops and 2-Ap:T are partial inhibitors. Thus, strong sites do lead to more Msh6•ATP•DNA complex formation, but it appears that the fraction of Msh2-Msh6 bound to weak sites can also form the ternary complex and presumably trigger MMR.
Consistent with our *in vitro* data, treatment with 2-Ap activates MMR in *E. coli* (Burdett et al., 2001; Glickman and Radman, 1980), and increases the number of MutS•MutL complexes bound to DNA in *B. subtilis* (Smith et al., 2001).

The finding that Msh2-Msh6 interaction with +T is less stable than G:T is interesting though not without precedent (Hess et al., 2002), and its physiological impact may be minimal since Msh2-Msh3 has overlapping IDL recognition function (Marsischky et al., 1996).

We also noted that in a different sequence context +T can strongly inhibit ATP hydrolysis (Antony and Hingorani, 2003), while sequence context does not change Msh2-Msh6 binding affinity and ATPase kinetics for a G:T mismatch (Antony and Hingorani, 2003; Mazurek et al., 2009). Therefore Msh2-Msh6 may be more influenced by sequence context when binding a base insertion/deletion (IDL) as compared with base pair mismatches.

**E339A mutant**

From the crystal structures of *E. coli* and *T. aquaticus* MutS homodimers (Lamers et al., 2000; Obmolova et al., 2000) and the human Msh2-Msh6 heterodimer (Warren et al., 2007) bound to different errors/lesions, all structures showed MutS/Msh2-Msh6 encircled around DNA, making sequence-independent hydrogen bonding and van der Waals contacts in the vicinity of the mismatch/IDL/lesion as well as direct contacts with the base in question mainly via a conserved Phe-X-Glu motif from one subunit (Msh6 in eukaryotes). The DNA is kinked toward the major
groove (45 – 60°), with Phe stacked against the base and Glu hydrogen bonded to it. The striking similarities among these structures suggest that the kinked DNA complex is a significant species in the pathway leading to initiation of MMR. However, since no structure of MutS/Msh2-Msh6 bound to fully matched DNA has been reported, the mechanism by which these proteins interact with DNA and recognize their target sites is unclear.

To further investigate this mechanism, we made a mutant, Msh2-Msh6\textsubscript{E339A} in which the Glu in the Phe-X-Glu motif was changed to Ala. From the kinetic data, Msh2-Msh6\textsubscript{E339A} binds G:T and 2Ap:T at slightly faster rate than wild-type Msh2-Msh6. The significant difference is the dissociation rate. Msh2-Msh6\textsubscript{E339A} releases 2Ap:T at ~7-fold slower rate and G:T at ~ 4-fold faster rate than wild-type; while +T binding kinetics appear to be unaffected. The ATPase analysis shows that free Msh2-Msh6\textsubscript{E339A} catalyzes ATP like the wild-type Msh2-Msh6. However, for Msh2-Msh6\textsubscript{E339A}, 2-Ap:T binding inhibits ATP hydrolysis as strongly as G:T. These data indicate that Msh6 Glu339 alters Msh2-Msh6•DNA complex dissociation rates and thereby influences selection of MMR targets sites in DNA.

We also made a similar Glu to Ala mutant MutS-E41A on \textit{T. aquaticus} MutS (Tessmer et al., 2008). From the pre-steady-state analysis, wildtype MutS and MutS-E41A catalyze the same fast burst of ATP hydrolysis at 8 s\textsuperscript{-1} and similar \( k_{\text{cat}} = 0.4-0.5 \) s\textsuperscript{-1}, in the absence of DNA or in the presence of homoduplex DNA (Tessmer et al., 2008). +T DNA can suppress ATP hydrolysis by both wildtype MutS and MutS-E41A and AFM showed they have same conformation, while G:T DNA can suppress
the ATP hydrolysis by wildtype MutS but have no effect on MutS-E41A (Tessmer et al., 2008). Compared with the Msh2-Msh6E339A mutant, the loss of Glu in the conserved Phe-X-Glu motif of *T. aquaticus* MutS affects its interactions and ATPase activities with G:T mismatch but not IDLs (Tessmer et al., 2008).

Research from other groups show that in *E. coli* the MutS-E38A mutant almost loses repair function on both single base pair mismatch and IDL DNA, but in *S. cerevisiae* Msh2-Msh6E339A had nearly all repair function (Holmes et al., 2007). All these data suggest that *E. coli* MutS, *T. aquaticus* MutS and *S. cerevisiae* Msh2-Msh6 have slight differences in their ability to recognize mismatched base pairs and IDLs, and the Glu residue plays an important role in interaction with a base pair mismatch but not an IDL (Holmes et al., 2007).

**Future research directions**

In future research, we can add Mlh1-Pms1 into the system to further clarify the Msh2-Msh6 recognition mechanism and find out how its ATPase activity helps to initiate the next step in the MMR reaction. MutL/Mlh1-Pms1 is the next protein in the reaction that forms a complex with MutS and DNA, and its mechanism of action is unclear (Iyer et al., 2006).

We can also use Msh2-Msh6 ATPase mutants to study the effect of loss of ATP activity on DNA binding and mismatch recognition.

Finally, recent research revealed that PCNA interacts with Msh2-Msh6 and may have a role in the mismatch binding step in MMR (Shell et al., 2007). In later
research, we can add PCNA into the experimental system and gain further insight into how MMR may be coupled to DNA replication.


