Title
Mechanisms of initiation of DNA mismatch repair in Saccharomyces cerevisiae

Permalink
https://escholarship.org/uc/item/6qm8h5z6

Author
Shell, Scarlet Sara

Publication Date
2008

Peer reviewed|Thesis/dissertation
Mechanisms of Initiation of DNA Mismatch Repair in

Saccharomyces cerevisiae

A Dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Biomedical Sciences

by

Scarlet Sara Shell

Committee in charge:

Professor Richard D. Kolodner, Chair
Professor John M. Carethers
Professor Karen F. Oegema
Professor Susan S. Taylor
Professor Joseph M. Vinetz

2008
The Dissertation of Scarlet Sara Shell is approved, and it is acceptable in quality and form for publication on microfilm:

__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________

Chair

University of California, San Diego

2008
DEDICATION

For my family, whose love and constancy are my greatest treasures.
## TABLE OF CONTENTS

DEDICATION .................................................................................................................. iv

TABLE OF CONTENTS .................................................................................................. v

LIST OF FIGURES ......................................................................................................... ix

LIST OF TABLES ........................................................................................................... xi

ACKNOWLEDGEMENTS ................................................................................................. xii

VITA ...................................................................................................................................... xiv

ABSTRACT OF THE DISSERTATION ............................................................................ xvii

CHAPTER 1 ...................................................................................................................... 1

Overview of DNA mismatch repair ................................................................. 1

1.1 SUPPRESSION OF MUTATION AND HUMAN DISEASE .............................. 2

1.2 OVERVIEW OF THE PROKARYOTIC MISMATCH REPAIR PATHWAY . 4

1.3 RECOGNITION OF MISPairs BY THE EUKARYOTIC MUTS HOMOLOGUES . . 7

1.4 TRANSMISSION OF SIGNALS FROM THE MUTS HOMOLOGUES TO DOWNSTREAM FACTORS ................................................................. 10

1.5 THE ROLES OF THE MULTIPLE MUTL HOMOLOGUE COMPLEXES IN EUKARYOTIC MISMATCH REPAIR ......................................................... 13

1.6 ROLES OF PCNA AND THE N-TERMINAL REGIONS OF MSH6 AND MSH3 IN MISMATCH REPAIR .......................................................... 17

1.8 SUMMARY AND AIMS ......................................................................................... 22

1.9 REFERENCES ........................................................................................................ 23
CHAPTER 2...........................................................................................................34

Dominant effects of an Msh6 missense mutation on DNA repair and cancer

susceptibility.............................................................................................................34

2.1 SUMMARY.........................................................................................................35

2.2 SIGNIFICANCE..................................................................................................36

2.3 INTRODUCTION.................................................................................................37

2.4 RESULTS ............................................................................................................42

Analysis of MSH6 missense mutations in S. cerevisiae........................................42
Generation of Msh6TTD mutant mice. .................................................................45
Mismatch repair activities in Msh6TTD mutant extracts. .....................................47
Microsatellite Instability in Msh6TTD mutant mice. ............................................49
In vivo mutation frequencies and spectra in Msh6 mutant mice. .........................51
Survival and cancer phenotype in Msh6TTD mutant mice. .................................53
DNA damage response in Msh6TTD MEF cells. ................................................57

2.5 DISCUSSION......................................................................................................60

2.6 EXPERIMENTAL PROCEDURES .....................................................................66

2.7 ACKNOWLEDGEMENTS.................................................................................73

2.8 REFERENCES....................................................................................................74

CHAPTER 3.............................................................................................................81

The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to

PCNA.......................................................................................................................81

3.1 SUMMARY.........................................................................................................82

3.2 INTRODUCTION.................................................................................................83

3.3 RESULTS............................................................................................................86

The Msh6 NTR is hypersensitive to proteolysis.....................................................86
N-terminal deletions of Msh6 define two mutation classes....................................86
Deletion of the N-terminal 123 residues of Msh6 completely abrogates
interactions with PCNA.........................................................................................91
The Msh6 NTR is not required for nuclear localization, DNA binding, or mispair discrimination. ................................................................. 91
The Msh3 and Msh6 NTRs are interchangeable and the Msh6 NTR can function when placed on Msh2. ......................................................... 92
The Msh6 NTR is an unstructured peptide that can interact with PCNA trimers at a 3:1 ratio. ................................................................. 94
PCNA does not induce substantial structure in the Msh6 NTR. .................. 99
Solution structure of the Msh2-Msh6-PCNA complex ................................ 100
Most functional PIP boxes in S. cerevisiae proteins are located in disordered regions. ................................................................. 105
3.6 ACKNOWLEDGEMENTS .......................................................... 115

3.7 SUPPLEMENTARY MATERIALS ............................................... 116
3.8 REFERENCES ......................................................................... 134

CHAPTER 4 .................................................................................. 140

A chimeric Saccharomyces cerevisiae Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins ......................... 140

4.1 ABSTRACT ........................................................................... 141

4.2 INTRODUCTION ................................................................... 142

4.3 RESULTS ............................................................................... 146

Generation of chimeric MSH6-MSH3 alleles that function in vivo. .......... 146
The mutations suppressed by the msh6(3-MBD) allele are similar to those suppressed by MSH3. ............................................................... 150
The Msh2-Msh6(3-MBD) chimera shows specificity for insertion/deletion loops in vitro. ......................................................................... 154
The msh6(3-MBD) allele functions with an msh2 MBD deletion mutant ....... 156
Unlike Msh3, Msh6(3-MBD) does not mediate significant repair through the Mlh1-Mlh3 complex ...................................................................... 156
4.4 DISCUSSION ........................................................................ 160

4.5 MATERIALS AND METHODS .................................................. 163

4.6 ACKNOWLEDGEMENTS ........................................................ 166

4.7 SUPPLEMENTARY MATERIALS ............................................. 167

4.8 REFERENCES ........................................................................ 174
CHAPTER 5.............................................................................................................178

Summary and future directions..............................................................................178

5.1 EFFECT OF A DOMINANT MUTATION IN THE ATPASE DOMAIN OF
MSH6 ..............................................................................................................................180

5.2 STRUCTURE AND FUNCTION OF THE N-TERMINAL REGION OF MSH6
.........................................................................................................................................183

5.3 MODULARITY OF THE MSH3 MISPAIR-BINDING DOMAIN ......................187

5.4 FUTURE DIRECTIONS ............................................................................................189

5.5 REFERENCES .......................................................................................................192

APPENDIX A.............................................................................................................195
LIST OF FIGURES

Figure 1-1. Structure of *Thermus aquaticus* MutS bound to a single nucleotide insertion mispair (Protein Data Bank ID 1ewq)......................................................... 6

Figure 2-1. Genetic analysis of *msh6* mutations in yeast ..............................................44

Figure 2-2. Generation of *Msh6*<sup>T1217D</sup> mice ..........................................................46

Figure 2-3. Mismatch repair activities in Msh6 mutant cell extracts. ..............................48

Figure 2-4. MSI in genomic DNA of *Msh6* mutant mice. .............................................50

Figure 2-5. Increased mutation frequencies in *Msh6* mutant mice ...............................52

Figure 2-6. Reduced survival of *Msh6*<sup>TD</sup> mutant mice ........................................54

Figure 2-7. Tumors in *Msh6*<sup>TD</sup> mutant mice ...........................................................56

Figure 2-8. Drug sensitivity and apoptosis in *Msh6* mutant cells .................................58

Figure 3-1. Partial proteolysis of Msh2-Msh6 and the Msh6 NTR. .................................87

Figure 3-2. Size exclusion chromatography of Msh6 NTR, PCNA, and their complex. .................................................................95

Figure 3-3. SAXS of Msh6 NTR, PCNA, and their complex ...........................................97

Figure 3-4. Partial proteolysis of the NTR and full length Msh2-Msh6 complexed to PCNA and solution structure of the Msh2-Msh6 complex with PCNA. ..........101

Figure 3-5. Disorder predictions for PIP-box containing proteins. ..............................104

Supplemental Figure 3-1. Alignment of yeast and human MutS homologs. .............119

Supplemental Figure 3-1. Alignment of yeast and human MutS homologs, continued. .................................................................123
Supplemental Figure 3-2. Complementation of the mutator phenotype of an msh3Δ msh6Δ strain by low-copy number plasmids bearing MSH6 alleles indicated...125

Supplemental Figure 3-3. PCNA interaction with protein constructs corresponding to complementing MSH6 alleles. ..........................................................126

Supplemental Figure 3-4. Mispair binding by Msh2-Msh6 and Msh2-Msh6Δ2-251. ..........................................................................................................................127

Supplemental Figure 3-5. Disorder predictions indicate that human and yeast NTRs are similar. .................................................................128

Supplemental Figure 3-6. Disorder predictions for other PIP box-containing yeast proteins. .................................................................130

Figure 4-1. Mispair-binding domain structure and sequence alignments. ...............147

Figure 4-2. Properties of point mutants and chimeric alleles..............................149

Figure 4-3. Surface plasmon resonance analysis of DNA binding......................155

Supplemental Figure 4-1. Example of a patch test experiment to assess functionality of msh3/6 mutants.........................................................168

Supplemental Figure 4-2. Plasmids for making heteroduplex DNA substrates.......169

Figure A-1. Summary of Msh3/Msh6 chimeras..............................................201

Figure A-2. Summary of Msh3 and Msh6 deletion mutants..............................203

Figure A-3. Summary of Mlh1, Mlh3, and Pms1 mutants.................................205
LIST OF TABLES

Table 3-1. Mutation rates of Msh6 mutants.................................................89
Table 3-2. Mutation rates of N-terminal region chimeric alleles.....................93
Supplemental table 3-1. Plasmids used for complementation and expression
  experiments.................................................................117
Table 4-1. Mutation rate analysis of msh6(3-MBD)........................................151
Table 4-2. CAN1 mutation spectra.............................................................153
Table 4.3. Complementation of msh2Δ msh3Δ msh6Δ strain by MSH3/6 alleles and
  msh2ΔMBD........................................................................157
Table 4-4. Effect of MLH3 genotype on mutation rate......................................159
Supplemental table 4-1. Top and bottom strand combinations for mispaired
  substrates...........................................................................170
Table A-1. Summary of all plasmid-born mutants created and tested..............196
ACKNOWLEDGEMENTS

I give tremendous thanks to my mentor Richard Kolodner for providing me with an outstanding scientific education and for knowing how to give me as much support as I needed while still giving me the freedom to develop through my own explorations.

I thank Chris Putnam for being a terrific mentor and sharing so generously his time, patience, and intellectual enthusiasm. Special thanks also to Dan Mazur, Marc Mendillo, and Jorrit Enserink for lots of help in the lab over the years. Thanks to Chris, Dan, Marc, Jorrit, Ellen Kats, Martin Hess, Hans Hombauer, Stephanie Soltero, Hal Hoffman, James Mueller, Sandra Martinez, Jason Chan, Jill Harrington, Cecilia Harries, Eric Jaehnig, Tiki Hayes, Katie Pallis, and Ann Montgomery as well as past lab members for experimental help, intellectual discussion, and a wonderful working environment.

I am very grateful to Joe Vinetz and Carlton Evans for giving me the invaluable opportunity to study in Peru and begin the transition into the next stage of my career.

To my best friend Vanessa Crouch, thanks for being a wonderful partner all these years, and for being by my side to support me through the challenges and celebrate the successes along the way.

Thanks to my grandfather Alexander “Alice” Stieglitz for always expecting great things from me.
I am deeply thankful to my family for the love, support, inspiration and encouragement that they have given me in abundance since the day I was born. To my sisters, your love and friendship is one of the most valued treasures in my life. Mom and Dad, your wisdom and unfailing support have made me the person I am today. Because of the gifts you have given me I can dream the dreams I do, and I can fulfill them.

Chapter 2, in full, is a reprint of the material as it appears in Yang, G., Scherer, S. J., Shell, S. S., Yang, K., Kim, M., Lipkin, M., Kucherlapati, R., Kolodner, R. D., and Edelmann, W. (2004). Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. Cancer Cell 6, 139-150. The dissertation author was a secondary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Shell, S. S., Putnam, C. D., and Kolodner, R. D. (2007). The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA. Mol Cell 26, 565-578. The dissertation author was a primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in Shell, S. S., Putnam, C. D., and Kolodner, R. D. (2007). Chimeric *Saccharomyces cerevisiae* Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. Proc Natl Acad Sci U S A 104, 10956-10961. The dissertation author was a primary investigator and author of this paper.
VITA

EDUCATION

2008    Ph.D., Biomedical Sciences, University of California, San Diego.
2001    B.A, Biological Sciences, Smith College.

RESEARCH EXPERIENCE

2001-2008  Biomedical Sciences graduate program, University of California, San Diego. Conducted studies of the mechanisms of initiation of DNA mismatch repair in the laboratory of Dr. Richard Kolodner.

2007    Howard Hughes Med into Grad program, University of California San Diego and Universidad Peruana Cayetano Heredia. Performed lab experiments and data analysis for tuberculosis diagnostic technique evaluation study in the laboratory of Dr. Carlton Evans.

2001    Summer research fellowship, Smith College. Participated in assessment of ability of soil microbes to degrade polyaromatic hydrocarbons.

2000    Research experience for undergraduates, the Whitney Laboratory, University of Florida. Used cloning strategies to search for novel ion channels in marine invertebrates.


1999    Summer research fellowship, Smith College. Surveyed coliform bacteria in freshwater sites in western Massachusetts.

TEACHING EXPERIENCE

2004    Teaching assistant for Microbial Genetics course, University of California, San Diego.

2000-2001    Teaching assistant for Introductory Biology and Molecular Biology laboratory sections, Smith College.
HONORS, AWARDS AND FELLOWSHIPS

2007  HHMI Med into Grad fellowship.

2003-2006  NIH/NIGMS Genetics training grant

2001  Elected to Phi Beta Kappa honor society

2001  Elected to Sigma Xi Scientific Research Society

1998-2001  Dean’s list, Smith College

1999  First Group Scholar, Smith College

PUBLICATIONS


*Joint first authors.


POSTERS AND PRESENTATIONS

*Joint first authors.

mispair-binding domain of Msh3 can be transferred to Msh6 and confer Msh3-like mispair recognition and repair properties onto Msh6.

Shell, S.S., Putnam, C.D., and Kolodner, R.D. 2005. Biomedical Sciences Graduate Program annual retreat, La Jolla, California. Talk: A 131-residue mispair binding domain can be transferred from Msh3 to Msh6 to confer Msh3-like mispair recognition and repair properties onto Msh6


ABSTRACT OF THE DISSERTATION

Mechanisms of Initiation of DNA Mismatch Repair in

Saccharomyces cerevisiae

by

Scarlet Sara Shell

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2008

Professor Richard Kolodner, Chair

DNA mismatch repair (MMR) is a conserved pathway that ensures replication fidelity by repairing errors that occur during replication of the genome. In eukaryotes MMR is initiated by two heterodimeric MutS homologue (MSH) complexes, Msh2-Msh3 and Msh2-Msh6, which recognize and bind mispairs and transmit signals to downstream components that effect repair. In this work we have investigated several
aspects of the mechanisms by which the MSHs locate and recognize mispairs and initiate the downstream events in MMR. We have found that the N-terminal regions (NTRs) of Msh6 and Msh3 in *Saccharomyces cerevisiae* are long unstructured linkers that tether the MSH complexes to the sliding clamp PCNA. A second region of the Msh6 NTR is functionally partially redundant with the PCNA interacting region, and in the absence of the entire NTR MMR is severely impaired. The NTRs of Msh3 and Msh6 can be exchanged to create functional chimeric proteins, and the Msh6 NTR is functional when fused to Msh2.

Here we also describe a chimeric *S. cerevisiae* MSH in which the mispair-binding domain of Msh6 is replaced by that of Msh3. The chimera is functional for MMR and the mispair-binding domain confers Msh3-like mispair recognition specificities. The chimera retains Msh6-like interactions with the MutL homologues, indicating that the mispair-binding domains and the type of mispair recognized do not contribute to the specificity of subsequent interactions. These findings suggest that the mispair-binding domains of Msh3 and Msh6 are modules that confer distinct mispair specificities but communicate signals of recognition to other domains in the same fashion.

Additional studies described here demonstrate that a dominant negative mutation in the ATPase domain of Msh6 is sufficient to cause cancer in mice. However, cells from these mice do not show the increased resistance to drugs that is seen in MMR knock-out cells. Experiments in *S. cerevisiae* demonstrate that the mutation in these mice is equivalent to a previously described yeast mutation and a mutation in a human cancer case.
CHAPTER 1

Overview of DNA mismatch repair
1.1 SUPPRESSION OF MUTATION AND HUMAN DISEASE

Suppression of mutations is important for most organisms. Mutations can arise from a variety of sources of DNA damage, both endogenous and exogenous. Endogenous sources of mutations include replication errors and damage to nucleotides or bases by reactive oxygen species that are present as by-products of metabolism. Exogenous sources of DNA damage such as methylating agents and radiation can also result in mutations. Numerous DNA repair pathways have evolved to repair various types of DNA damage and thereby suppress mutation rates. This work will focus on the DNA mismatch repair (MMR) pathway, which suppresses mutation rates by up to several orders of magnitude by correcting replication errors that result in misincorporated bases (Kolodner, 1996; Kolodner and Marsischky, 1999).

The importance of suppressing mutation rates is clear from the clinical consequences that result from the absence of MMR. Humans with germline mutations that result in the absence of non-redundant components of the MMR pathway have a cancer-predisposition syndrome called Lynch syndrome (also called hereditary non-polyposis colorectal cancer [HNPCC]). Patients with Lynch syndrome develop early-onset cancers, most frequently in the gastrointestinal tract and endometrium, that are characterized by the presence of frequent insertions and deletions in di- and tri-nucleotide repeat sequences, termed microsatellite instability (Bronner et al., 1994; Fishel et al., 1993; Leach et al., 1993; Peltomaki, 2003). Germline mutations that cause partial defects in MMR may result in later-onset cancer susceptibility syndromes.
(Akiyama et al., 1997; Berends et al., 2002; Kolodner et al., 1999; Miyaki et al., 1997). Additionally, somatic inactivation of MMR genes is a common feature of many different types of sporadic cancers (Kane et al., 1997; Veigl et al., 1998).

In addition to contributing to cancer susceptibility and development, defects in the MMR pathway appear to reduce sensitivity of tumors to certain chemotherapeutic agents (Carethers et al., 2004; Ribic et al., 2003). MMR status may therefore have implications for prognosis and appropriate treatment of some cancers. More work is needed to understand the mechanisms by which MMR affects sensitivity to chemotherapeutics, and how treatment strategies can be optimized based on MMR status.

There may also be implications for loss of MMR in human pathogens. Variants of *Pseudomonas aeruginosa* that have lost key components of the MMR pathway have been found to be associated with persistent infection and drug resistance in cystic fibrosis patients, and the high mutation rates in these strains have been proposed to give them an advantage in immune evasion and allow for accelerated development of drug resistance (Oliver et al., 2000).
1.2 OVERVIEW OF THE PROKARYOTIC MISMATCH REPAIR PATHWAY

MMR was first described in prokaryotes, and subsequently found to be conserved in all eukaryotic species examined (Iyer et al., 2006). The prokaryotic pathway is simpler, and after over 20 years of research is relatively well understood and can be reconstituted in vitro from purified components (Lahue et al., 1989).

Prokaryotic MMR is initiated by recognition of mispaired DNA sequences by MutS, which is active as a homodimer. MutS can recognize base:base mispairs resulting from nucleotide misincorporations by the replicative polymerases, as well as loops of between one and several unpaired nucleotides resulting from insertion or deletion events caused by polymerase slippage (Modrich, 1991). MutS is a multi-domain protein (Figure 1-1, panel A) (Lamers et al., 2000; Obmolova et al., 2000). The most N-terminal domain is responsible for direct interaction with mispairs, while domain IV facilitates DNA binding by forming a clamp that contacts the DNA backbone in a non-sequence-specific manner. Mismatch-binding results in conformation changes that are transmitted throughout the protein and affect nucleotide binding and hydrolysis by domain V (Blackwell et al., 2001; Lamers et al., 2004). The details of mismatch-binding and nucleotide interactions will be considered in greater detail in subsequent sections.

Following recognition of a mispair, MutS becomes competent to bind another homodimer called MutL, forming the so-called ternary complex of MutS, MutL, and mispair-containing DNA (Grilley et al., 1989). MutL then becomes activated and is
able to subsequently activate the endonuclease MutH and DNA helicase II (Au et al., 1992; Welsh et al., 1987; Yamaguchi et al., 1998). MutH recognizes hemi-methylated Dam (GATC) sites that are transiently present following replication and prior to methylation of the nascent strand (Iyer et al., 2006). Upon activation by MutL, MutH selectively cleaves the unmethylated daughter strand. Hemi-methylated Dam sites upstream or downstream of the mispair can be utilized, and unwinding of the daughter strand by DNA helicase II occurs in the appropriate direction toward the mispair in either case (Au et al., 1992; Cooper et al., 1993; Grilley et al., 1993; Lahue et al., 1987). One of several exonucleases then degrades the mispair-containing strand, which is subsequently re-synthesized by DNA polymerase III (Burdett et al., 2001; Iyer et al., 2006; Viswanathan et al., 2001).
Figure 1-1. Structure of *Thermus aquaticus* MutS bound to a single nucleotide insertion mispair (Protein Data Bank ID 1ewq).

A. Domain structure of MutS. The non-mispair-contacting subunit of the homodimer is colored in light blue, while the mispair-contacting subunit is colored by domain, from N-terminus: Domain I (red), Domain II (orange), Domain III (yellow), Domain IV (green), and Domain V (blue).

B. Interaction with the unpaired thymine by stacking of Phe39. For clarity, only Domain I of the mispair-contacting subunit is shown (red). The unpaired thymine is in bold black and the side-chain of Phe39 is shown. The figure was generated with PyMOL (DeLano, 2002).
1.3 RECOGNITION OF MISPAIRS BY THE EUKARYOTIC MUTS HOMOLOGUES

In eukaryotes, three MutS Homologues (MSHs) form two heterodimeric complexes that recognize partially overlapping subsets of mispairs and together are responsible for initiation of MMR (Acharya et al., 1996; Marsischky et al., 1996). Msh2 is the common subunit that is present in both complexes, while the unique subunits are Msh3 and Msh6. The unique subunits are responsible for determining mispair specificity. The Msh2-Msh6 complex, also called MutSα, recognizes most base:base mispairs and small insertion/deletion mispairs, while the Msh2-Msh3 complex, MutSβ, recognizes both large and small insertions and deletions (Marsischky et al., 1996). The specificity of mispair recognition is heavily influenced by sequence context, and in certain contexts MutSα can recognize loops of up to 10 nucleotides (Marsischky and Kolodner, 1999), while in other contexts MutSβ has been shown to initiate repair of base:base mispairs (Harrington and Kolodner, 2007).

The partially overlapping functions of MutSα and MutSβ result in a situation where losses of either Msh3 or Msh6 individually have small to moderate effects on overall mutation rates, while loss of Msh2 causes a much more dramatic increase in the accumulation of mutations (Marsischky et al., 1996). Mutations in MSH2 are therefore a common cause of HNPCC and result in tumors with a high degree of microsatellite instability, while mutations in MSH6 appear to cause milder cancer-predisposition syndromes, and inherited MSH3 mutations have not been found in
association with cancer susceptibility (Fishel et al., 1993; Huang et al., 2001; Kolodner et al., 1999; Peltomaki, 2003).

Initial clues to the mechanism of mispair recognition were given by cross-linking and mutagenesis studies that implicated a conserved phenylalanine residue in domain I of MutS, which is also present in eukaryotic Msh6. Mutation of this Phe39 to alanine in *Thermus aquaticus* was found to dramatically decrease mispair specificity *in vitro* (Malkov et al., 1997), and mutation of the equivalent Phe36 residue in *Escherichia coli* to alanine was also shown to decrease mispair binding *in vitro* as well as abolish MMR activity *in vivo* (Yamamoto et al., 2000). Crystal structures of *E. coli* MutS bound to a G:T mispair and *T. aquaticus* MutS bound to an insertion mispair subsequently demonstrated that the aromatic ring of Phe36 or 39, respectively, from one of the subunits in each homodimer stacks with the mispaired or unpaired thymine base (Figure 1-1, panel B) (Lamers et al., 2000; Obmolova et al., 2000). Similar stacking of Phe36 was seen with adenine and guanidine bases in subsequent structures solved with substrates containing C:A, A:A, and G:G mispairs (Natrajan et al., 2003).

Another important feature of these crystal structures was that in each case, the MutS homodimer binds DNA in an asymmetric manner such that one subunit makes direct contact with the mispair while the other subunit only contacts the DNA backbone. This asymmetric mode of binding has evolved in eukaryotes into specialized roles performed by different proteins; Msh2 contacts only the DNA backbone, while Msh3 and Msh6 make direct contact with mispairs and thereby determine mispair specificity (Habraken et al., 1996; Marsischky et al., 1996; Warren et al., 2007).
The mispair-binding domain of eukaryotic Msh6 retains the conserved phenylalanine, and mutation of this Phe337 to alanine in *Saccharomyces cerevisiae* causes loss of function *in vivo* and defects in mispair binding *in vitro* (Bowers et al., 1999). Mutation of the equivalent phenylalanine in human Msh6 to alanine causes a substantial MMR defect *in vitro* (Dufner et al., 2000). In contrast, mutation of the phenylalanine or tyrosine present at the equivalent position of human or yeast Msh2, respectively, does not cause any discernible defect, consistent with Msh2 not interacting directly with mispairs (Bowers et al., 1999; Dufner et al., 2000). A crystal structure of human MutSα confirms that Phe432 indeed stacks with mispaired bases in a variety of different types of mispairs (Warren et al., 2007).

In contrast to Msh6, Msh3 lacks the conserved phenylalanine in its mispair-binding domain, and the mechanism by which it mediates mispair recognition is therefore poorly understood. The recent finding that domain I of Msh2 is dispensable for activity of MutSα but necessary for function of MutSβ further implies that there are significant differences in how mispairs are recognized by these two complexes (Goldfarb and Alani, 2005; Lee et al., 2007). Mechanisms of interaction of prokaryotic MutS and eukaryotic MutSα with loops of greater than one nucleotide are also unknown, as no crystal structures containing such mispairs have been reported. Crystals structures showed that mispair-containing DNA is bent at an angle of approximately 60° when bound by MutS or MutSα, leading to the suggestion that the presence of a mispair makes DNA easier to kink, and that this bendability may be a component of mispair recognition (Wang et al., 2003).
1.4 TRANSMISSION OF SIGNALS FROM THE MUTS HOMOLOGUES TO DOWNSTREAM FACTORS

Mispair binding by MutS and its eukaryotic homologues causes conformational changes throughout the dimer that affect its activities in a number of ways. Interaction with mispairs stabilizes the structure of the mispair-binding domain, and is in fact required for this domain to be ordered in reported crystal structures (Obmolova et al., 2000). Mispair-binding also affects the affinity of MutS and its homologues for adenine nucleotides as well as rates of ATP hydrolysis (Antony and Hingorani, 2003; Bjornson et al., 2000; Gradia et al., 1997; Hess et al., 2002; Junop et al., 2001; Lamers et al., 2003). Finally, the changes in conformation and nucleotide binding state of the MutS proteins are required for interaction with the MutL proteins and the subsequent steps of the MMR pathway (Grilley et al., 1989; Habraken et al., 1998; Hess et al., 2006; Mendillo et al., 2005).

A notable feature of the crystal structures of prokaryotic MutS is their modularity. Each of the five domains forms a discrete unit that resembles another known structure, and domains I and IV fold independently and are connected to the rest of the protein mainly by flexible peptide linkers (Obmolova et al 2000). The structure of *T. aquaticus* MutS demonstrated that while domains II, III, and V maintain similar conformations in the presence and absence of mispaired DNA, domains I and IV, which represent the mispair-binding domain and DNA backbone-clamping domains, respectively, are too flexible in the absence of DNA to be
discerned crystallographically (Figure 1-1). Mispair-binding therefore induces stabilization of the domains of MutS that are directly involved in interaction with DNA.

Transmission of the mispair-induced conformational changes throughout the MutS dimer is indicated by changes in nucleotide affinity and hydrolysis in domain V at the C-terminal end of the protein. Domain V is a member of the ABC superfamily of ATPases and hydrolyzes ATP by a Walker A motif (Haber et al., 1988). This region is highly conserved between prokaryotic MutS and its eukaryotic homologues, and mutations that disrupt nucleotide binding and/or hydrolysis cause dramatic MMR defects (Alani et al., 1997; Das Gupta and Kolodner, 2000; Dufner et al., 2000; Fishel and Wilson, 1997; Haber and Walker, 1991; Wu and Marinus, 1994). Residues from both subunits contribute to each of the two ATP-binding sites (Lamers et al., 2000; Obmolova et al., 2000).

The asymmetry of the MutS homodimer is also evident in the ATPase domains, where one nucleotide-binding site has higher affinity for ATP than the other (Antony and Hingorani, 2004; Lamers et al., 2003). This asymmetry is also present in the eukaryotic homologues; the nucleotide-binding site that is composed mainly of Msh6 residues has higher affinity for ATP than the nucleotide-binding site that is composed mainly of Msh2 residues, while the Msh2 site has higher affinity for ADP (Mazur et al., 2006). Furthermore, nucleotide binding seems to induce conformational changes that allow communication between the two sites, since ATP-binding by the Msh6 site causes a decrease in the affinity of the Msh2 site for ADP (Mazur et al., 2006).
The presence of mispair-containing DNA has been shown to increase the ATPase activity of MutS and its homologues, while detailed kinetic studies have indicated that an initial burst phase of activity is inhibited by interaction with mispairs (Antony and Hingorani, 2003; Bjornson et al., 2000; Gradia et al., 1997; Hess et al., 2002). As interaction with DNA affects the ATPase domain, ATP-binding subsequently affects interaction with DNA. In the presence of ATP, affinity for mispairs appears to be lower (Alani, 1996; Gradia et al., 1997). Studies over the past several years have revealed that ATP-binding converts the MutS proteins into sliding clamps that diffuse freely along the DNA while remaining closely associated with it, and dissociating from the DNA only at free ends (Acharya et al., 2003; Gradia et al., 1999; Mendillo et al., 2005). In this sliding clamp state, MutS and its homologues do not remain associated with mispairs.

Recent studies have provided a model for the relationship between mispair binding and events at the ATPase domains of MutSα (Mazur et al., 2006). Mispair binding seems to occur when ADP is present in the Msh2 site or when both sites are empty. Upon mispair binding, the Msh6 site binds ATP and in the mispair-binding-induced conformation hydrolysis is inhibited. The stable presence of ATP in this site induces a conformational change that decreases the affinity of the Msh2 site for ADP and allows it to bind ATP. In the dual-ATP-bound state, the complex is competent to form a sliding clamp and to interact with the MutL homologues.
1.5 THE ROLES OF THE MULTIPLE MUTL HOMOLOGUE COMPLEXES IN EUKARYOTIC MISMATCH REPAIR

In eukaryotes, three MutL Homologue (MLH) proteins form two heterodimeric complexes that are functionally equivalent to the MutL homodimer in prokaryotes. The Mlh1-Pms1 complex in *S. cerevisiae* (called Mlh1-Pms2 in mammalian systems) seems to be the predominant MutL homologue complex in MMR, as evidenced by high mutation rates in *pms1* mutants that are indistinguishable from those of *msh2* mutants (Li and Modrich, 1995; Prolla et al., 1994a; Prolla et al., 1994b). The Mlh1-Mlh3 complex plays a role that is smaller but may be highly significant, particularly in mammals (Chen et al., 2005; Flores-Rozas and Kolodner, 1998).

Prokaryotic MutL is functional as a dimer. It stably dimerizes through interactions between globular C-terminal domains (Drotschmann et al., 1998; Guarne et al., 2004; Kosinski et al., 2005). A central region of approximately 100 residues is an unstructured linker joining the C-terminal dimerization domain to the globular N-terminal domain, which contains a nucleotide-binding site and forms a second site of dimerization upon ATP binding (Ban and Yang, 1998; Guarne et al., 2004). The nucleotide-binding sites are very weak ATPases but are required for MMR (Aronshtam and Marinus, 1996; Ban and Yang, 1998; Spampinato and Modrich, 2000). This structure is generally conserved in eukaryotes, which have globular N-and C-terminal domains connected by linkers of between 100 and 300 residues in length that are predicted to be unstructured (Pang et al., 1997). The C-terminal domain of
Pms1 has been recently shown to contain endonucleolytic activity that is required for MMR (Kadyrov et al., 2006). The active site residues are also conserved in Mlh3, but are not present in MutL sequences from MutH-containing prokaryotes.

All biochemical characterization of eukaryotic MutL homologues reported to date has been performed with Mlh1-Pms1 (or the equivalent human Mlh1-Pms2). Like MutL, this complex interacts with MutS homologue complexes upon mispair binding in the presence of ATP (Habraken et al., 1998; Mendillo et al., 2005). Mutations in Msh2 and Msh6 that prevent dual occupancy by ATP also fail to interact with Mlh1-Pms1 (Hess et al., 2006). Interaction with Mlh1-Pms1 appears to moderate the sliding of Msh2-Msh6 along DNA, but studies have been confounded by the affinity of Mlh1-Pms1 for DNA ends (Mendillo et al., 2005). The N-terminal region of Mlh1 has been implicated in interaction with MutSα, but the site of interaction on the MutSα complex is unknown (Plotz et al., 2003; Plotz et al., 2006).

The events following ternary complex formation in eukaryotes therefore remain obscure. In MMR reactions performed with purified human proteins in vitro, a mispair-containing circular DNA substrate that also contains a single nick either 3’ or 5’ to the mispair can be repaired by the combined activities of MutSα, Mlh1-Pms2, ExoI, RPA, PCNA, RFC, and DNA polymerase δ (Constantin et al., 2005; Zhang et al., 2005). Excision of the mispair-containing strand is performed by the 5’ to 3’ exonuclease activity of ExoI, starting either at the pre-existing nick if it is located 5’ to the mismatch, or at a nick introduced by the endonuclease activity of Mlh1-Pms1 at a position 5’ to the mismatch when the pre-existing nick is located 3’ to the mismatch (Kadyrov et al., 2006). Closed circular DNAs that do not contain pre-introduced nicks
are refractory to MMR in these reactions. It is not known how the nuclease activities of Mlh1-Pms1 and ExoI are regulated to produce appropriately directed degradation of the mispair-containing strand from the nick to the mispair. Furthermore, ExoI is not essential for MMR in yeast or mice and EXOI mutations have not been clearly linked to cancer susceptibility in humans, indicating that other nucleases perform partially redundant functions (Peltomaki, 2003; Tishkoff et al., 1997; Wei et al., 2003). While nicks provide effective strand-discrimination signals in vitro, the mechanism of discrimination in vivo remains unknown. The transient presence of nicks following replication of the lagging strand could direct MMR to the nascent strand in vivo, but it is not clear whether nicks occur frequently enough on the leading strand for this to be the sole mechanism of discrimination.

The Mlh1-Mlh3 complex was initially shown to have a role in MMR by genetic studies of S. cerevisiae (Flores-Rozas and Kolodner, 1998). Despite the fact that pms1 mutants have mutation rates as high as those of mlh1 mutants, mlh3 mutants do have small but significant increases in mutation rate that are equivalent to rates seen in msh3 mutants. An mlh3 msh3 double mutant is not any more defective than either of the single mutants, but an mlh3 msh6 double mutant shows a synergistic increase in mutation rate. This finding indicated that the Mlh1-Mlh3 complex may be required for repair of some mispairs that are detected by Msh2-Msh3, but does not play a significant role in repair of mispairs detected by Msh2-Msh6. The determinants of use of the Mlh1-Pms1 complex versus the Mlh1-Mlh3 complex by Msh2-Msh3 were unclear.
Experiments in mice have confirmed a role for Mlh3 in MMR in mammals. Mice with defects in Pms2 have different tumor spectra than mice deficient in Mlh1 (Prolla et al., 1998). Mice with mlh3 mutations have milder but still significant increases in tumorigenesis and microsatellite instability, and pms2 mlh3 double mutant mice are phenotypically comparable to mlh1 mice (Chen et al., 2005). A role for Mlh3 in suppression of mutation in humans is also suggested by the relative lack of HNPCC cases found to have defects in Pms2, although detection of Pms2 defects is complicated by a large number of related pseudogenes (Peltomaki, 2003).
1.6 ROLES OF PCNA AND THE N-TERMINAL REGIONS OF MSH6 AND MSH3 IN MISMATCH REPAIR

Proliferating Cell Nuclear Antigen (PCNA) is a homotrimeric sliding clamp involved in a number of processes related to DNA replication and repair. It is a processivity factor required for efficient activity of the replicative DNA polymerases, and also interacts with DNA ligase, helicases, the flap endonuclease Fen1/Rad27, chromatin assembly factors, and components of nucleotide excision repair, base excision repair, and mismatch repair pathways (Burgers, 1991; Chen et al., 1996; Clark et al., 2000; Flores-Rozas et al., 2000; Gary et al., 1997; Haracska et al., 2001; Johnson and O'Donnell, 2005; Li et al., 1995; Otterlei et al., 1999; Prelich et al., 1987; Schmidt et al., 2002; Shibahara and Stillman, 1999; Subramanian et al., 2005; Unk et al., 2002; Zhang et al., 2000). PCNA is required for the completion of MMR in eukaryotes because of its role in helping the polymerase to perform the resynthesis step, and it plays a second role earlier in the pathway as well (Gu et al., 1998; Umar et al., 1996).

Most of PCNA’s binding partners interact through a conserved sequence motif called the PIP-box, which associates with a groove created by the loop that connects the two domains of each PCNA subunit (Gulbis et al., 1996; Warbrick, 2000). Canonical PIP-boxes were found located near the N-termini of eukaryotic Msh3 and Msh6 (Clark et al., 2000; Flores-Rozas et al., 2000). Both Msh3 and Msh6 have additional regions of sequence that lie N-terminally to their mispair-binding domains,
while in Msh2 and prokaryotic MutS the mispair-binding domain is the most N-terminal portion of the protein. The N-terminal regions of the known eukaryotic Msh3 and Msh6 sequences are between approximately 100 and 400 residues in length and share relatively little sequence homology aside from the PIP-boxes located near their N-termini.

Mutations replacing two key phenylalanine residues of the PIP-box on \textit{S. cerevisiae} Msh6 with alanines have been shown to cause a mild defect in MMR \textit{in vivo} and eliminate interaction with PCNA \textit{in vitro}, and similar effects have been reported for Msh3 (Clark et al., 2000; Flores-Rozas et al., 2000). Studies of human cells expressing an Msh6 deletion mutant missing the PIP-box suggested that interaction with PCNA might facilitate sub-nuclear localization to replication foci, and found that extracts from these cells had a modest defect in MMR reactions \textit{in vitro} (Kleczkowska et al., 2001). It has been reported that interaction with mispaired DNA may disrupt the Msh6-PCNA interaction, although this finding was seemingly contradicted by another report that instead indicated that PCNA disrupts interaction between MutS\(\alpha\) and Mlh1-Pms1 (Bowers et al., 2001; Lau and Kolodner, 2003). A direct interaction between PCNA and Mlh1 has also been reported, although it appears to be much weaker than the Msh6-PCNA interaction and to occur through a different mechanism (Lee and Alani, 2006). The significance of this proposed interaction \textit{in vivo} is unclear.

Aside from their role in facilitating interaction with PCNA, little is known about the N-terminal regions of eukaryotic Msh3 and Msh6. Because no homologous regions exist in prokaryotic MutS, the reported MutS crystal structures did not provide
any information on the eukaryotic N-terminal regions. The recently reported crystal structure of human MutSα also failed to shed light on the N-terminal region of Msh6 because the crystallized protein was a truncated form missing the first 340 amino acids (Warren et al., 2007).
1.7 MISMATCH REPAIR AND APOPTOSIS

The earliest suggestions that mismatch repair played a role in drug sensitivity came from studies of bacteria that showed increased sensitivity to the alkylating agent N-methyl-N’-nitro-N-nitrosoguanidine (MNNG) in a dam methylation mutant, and rescue of the phenotype when MutS or MutL were deleted (Jones and Wagner, 1981; Karran and Marinus, 1982). Studies with human cell lines later suggested that loss of MMR resulted in low-level resistance to cisplatin (Aebi et al., 1996; Fink et al., 1996) and significantly increased tolerance to MNNG (Koi et al., 1994; Umar et al., 1997). In *S. cerevisiae*, loss of MMR has no effect on the sensitivity of otherwise wild-type strains to MNNG but rescues the increased sensitivity of strains that have defects in both methylguanine methyltransferase activity (*mgt1*) homologous recombination (*rad52*) as well (Cejka et al., 2005).

These findings led to two different models to explain how MMR could lead to drug sensitivity. Since alkylating agents create adducts such as O^6^-methylguanine that cause nucleotide misincorporations, one model suggests that futile cycles of MMR that could ensue and ultimately result in strand breaks and cell death (Karran and Bignami, 1994; Modrich, 1997). Another model suggests that components of the MMR pathway could have direct signaling functions that activate cell cycle checkpoints and lead directly to cell death by apoptosis when the DNA damage is too extensive to be repaired (Duckett et al., 1996; Modrich, 1997).

These concepts also led to questions about how MMR defects lead to the development of cancer. An early proposal suggested that increased rates of
accumulation of mutations could underlie cancer by leading to mutations in genes such as tumor suppressor genes and proto-oncogenes, and that these secondary mutations are ultimately the cause of tumorigenesis; the subsequent finding that MMR mutations are the cause of HNPCC was highly consistent with this theory (Loeb, 1991; Loeb, 2001). However, the idea that MMR components could also have checkpoint signaling functions led to a competing model in which the loss of these activities rather than increased mutagenesis per se is the cause of cancer susceptibility (Fishel, 2001). In the second model, loss of MMR would provide a selective advantage to pre-cancerous cells in of itself, prior to facilitating the acquisition of further mutations.
1.8 SUMMARY AND AIMS

The work described in this dissertation addresses several aspects of the initiation of MMR. Chapter 2 is an investigation of the effects of a dominant-negative mutation in the ATPase domain of Msh6 in mice, with additional experiments in *S. cerevisiae* that relate the mouse work with a mutation found in a human cancer case. The findings have implications for both the mechanism of cancer development in mammals with MMR defects and the mechanism by which MMR activity affects drug sensitivity. Chapter 3 describes a thorough analysis of the structure and function of the N-terminal region of Msh6, and also provides insight into the role of the N-terminal region of Msh3 and the role of PCNA interaction with MutSα and MutSβ. In Chapter 4 we describe a chimeric MutS homologue protein that contains the mispair-binding domain of Msh3 placed within Msh6. The chimera demonstrates the modularity of the MutS homologues and is also used to examine the determinants of differential use of the MutL homologues and the role of domain I of Msh2.
1.9 REFERENCES


CHAPTER 2

Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility
2.1 SUMMARY

Mutations in DNA mismatch repair (MMR) genes cause hereditary nonpolyposis colorectal cancer (HNPCC), and MMR defects are associated with a significant proportion of sporadic cancers. MMR maintains genome stability and suppresses tumor formation by preventing the accumulation of mutations and by mediating an apoptotic response to DNA damage. We describe the analysis of a dominant \textit{MSH6} missense mutation in yeast and mice that causes loss of DNA repair function while having no effect on the apoptotic response to DNA damaging agents. Our results demonstrate that \textit{MSH6} missense mutations can effectively separate the two functions, and that increased mutation rates associated with the loss of DNA repair are sufficient to drive tumorigenesis in MMR-defective tumors.
2.2 SIGNIFICANCE

Tumors that develop in HNPCC patients or sporadic cancers associated with MMR deficiency show increased genomic instability, which can be diagnosed by microsatellite instability (MSI) analysis, and also display resistance to a wide variety of chemotherapeutic agents. A substantial number of HNPCC cases result from missense mutations in MMR genes; however, the functional consequences of such mutations in normal and tumor cells are not well understood. By modeling a human MSH6 missense mutation in yeast and mice, we demonstrate that MMR missense mutations can differentially affect individual MMR functions. Our results provide new molecular insights into the anticancer functions of MMR and indicate that genotype-phenotype correlations can provide valuable information for the diagnosis and treatment of tumors with MMR mutations.
3.3 INTRODUCTION

DNA mismatch repair (MMR) functions to repair misincorporation errors that occur during DNA replication and acts on mispaired bases formed in recombination intermediates (Buermeyer et al., 1999; Kolodner, 1996; Kolodner and Marsischky, 1999; Modrich and Lahue, 1996). As a consequence of the role of MMR in DNA replication, MMR-defective cells have high mutation rates. MMR acts to prevent recombination between divergent DNA sequences, and as a consequence, MMR may also act to prevent some types of genome rearrangements (Evans and Alani, 2000; Harfe and Jinks-Robertson, 2000; Myung et al., 2001). MMR proteins also play a role in some cellular responses to DNA damage, and consequently, MMR-defective mammalian cells are resistant to killing by DNA-damaging agents, including some types of chemotherapeutic agents (Aebi et al., 1996; Bignami et al., 2003; D'Atri et al., 1998; Li, 1999; Modrich, 1997). MMR defects can underlie the development of cancer. HNPCC is an inherited cancer susceptibility syndrome where, in many cases, the tumors show microsatellite instability (MSI) indicative of a MMR defect (Peltomaki, 2003). The majority of these HNPCC cases are due to inherited mutations in two MMR genes, MSH2 and MLH1 (Peltomaki, 2003; Peltomaki and Vasen, 1997; Yan et al., 2000). A small number of HNPCC cases are due to inherited mutations in MSH6, whereas a larger number of later onset familial atypical-HNPCC cases are due to mutations in MSH6 (Akiyama et al., 1997; Berends et al., 2002; Kolodner et al., 1999; Miyaki et al., 1997; Wagner et al., 2001). Sporadic MMR-defective cancers are largely due to somatic silencing of the MLH1 gene (Kane et al., 1997; Kuismanen et
al., 2000; Veigl et al., 1998). Defects in other MMR genes have been implicated in no more than a small proportion of inherited and sporadic cancers. These differences between MMR genes have been attributed to the fact that some MMR proteins have unique functions in MMR, whereas others play partially redundant roles (Harfe and Jinks-Robertson, 2000; Kolodner and Marsischky, 1999; Marsischky et al., 1996). However, some MMR proteins could have different functions, such as, for example, a unique role in DNA damage responses.

Mouse models have provided insight into the relationship between MMR defects and cancer susceptibility (Buermeyer et al., 1999; Wei et al., 2002). Mice that are homozygous for null mutations in Msh2 or Mlh1 develop a broad spectrum of tumor types at early ages, and these tumors show MSI; in many regards, these mice mimic HNPCC caused by MSH2 and MLH1 mutations (de Wind et al., 1998; Edelmann et al., 1999; Prolla et al., 1998; Reitmair et al., 1996). Mice that are homozygous for Msh6 mutations are similar to Msh2 and Mlh1 mutant mice in regard to tumor spectrum. However, the Msh6 mutant mice develop cancers at later ages, and the tumors show little or no MSI (Edelmann et al., 1997). Studies of Msh6 mutant mice were the first to implicate MSH6 as a late onset cancer susceptibility gene, a result that has been extended to human populations. Pms2 defects in mice cause the development of tumors with MSI at early ages, although the tumor spectrum is distinct from that of Msh2, Msh6 or Mlh1 mutant mice (Baker et al., 1995). This difference in tumor spectrum correlates with the limited involvement of PMS2 defects in HNPCC. Mutations in Msh3 and Exo1 only cause the development of tumors at very old ages, which parallels the limited involvement of these genes in HNPCC. However, Msh3
mutations accelerate the development of tumors when combined with $Msh6$ mutations, similar to the interaction between $msh6$ and $msh3$ mutations in $S. cerevisiae$ (de Wind et al., 1999; Edelmann et al., 2000; Kolodner et al., 1999; Marsischky et al., 1996; Sia et al., 1997; Wei et al., 2003).

Mismatch recognition in eukaryotic MMR involves three proteins, MSH2, MSH6 and MSH3, which are homologs of the bacterial MutS protein. These three proteins form two heterodimeric complexes, the MSH2-MSH6 and MSH2-MSH3 complexes, that each have different mismatch recognition properties (Acharya et al., 1996; Drummond et al., 1995; Genschel et al., 1998; Guerrette et al., 1998; Marsischky et al., 1996; Palombo et al., 1995). The MSH2-MSH6 complex is the major mismatch recognition complex and recognizes both base:base and single base insertion/deletion mismatches, whereas the MSH2-MSH3 complex appears to primarily recognize insertion/deletion mismatches. This difference is also reflected in their relative abundance with MSH2-MSH6 complexes occurring in excess of MSH2-MSH3 complexes in yeast, human, and mouse (Marsischky and Kolodner 1996; Genschel et al. 1998; deWind et al. 1999). It is this difference in mismatch recognition spectrum that has been suggested to underlie the genetic differences between MSH2, MSH6 and MSH3 (Marsischky et al., 1996; Umar et al., 1998). When the MSH complexes bind a mismatch, they form a ring around the DNA, and upon ATP binding, this ring can move along the DNA away from the mismatch (Blackwell et al., 1998; Gradia et al., 1999; Hess et al., 2002). The importance of the ATP binding-mediated release from mispairs has been in part established by the isolation of dominant mutations in $MSH6$ in yeast (Das Gupta and Kolodner, 2000).
These mutations strongly interfere with MMR when present on a single copy plasmid in a wild-type strain. At the chromosomal locus, they cause high rates of base substitution and frameshift mutations, in contrast to only high rates of base substitution mutations caused by \textit{MSH6} deletion mutations, and the increased rate of frameshift mutations is suppressed by overexpression of \textit{MSH3}. The mutant MSH2-MSH6 complexes hydrolyze ATP in the absence of DNA, but have a much longer half-life bound to a mispair in the presence of ATP than the wild-type complex; the reduced turnover from mispairs may be responsible for preventing MSH2-MSH3 dependent MMR from acting on insertion/deletion mispairs and accounts for the dominant phenotypes caused by these \textit{msh6} mutations (Hess et al., 2002).

Currently, there are two views on how MMR defects lead to increased development of cancer. In one view, increased mutation rates increase the rate of accumulation of mutations in tumor suppressor genes and proto-oncogenes, which results in the development of cancer (Kinzler and Vogelstein, 1996; Loeb, 2001). In the other, defects in DNA damage responses and DNA damage induced apoptosis provides a selective advantage to tumor cells (Fishel, 2001; Nowell, 1976; Tomlinson and Bodmer, 1999). To gain insight into these two possible mechanisms, we have investigated the properties of specific single nucleotide changes in \textit{MSH6} in \textit{S. cerevisiae} and mouse models. We describe a dominant mouse \textit{Msh6} mutation that causes a strong mutator phenotype and increased cancer development but no defect in DNA damage induced apoptosis and provide evidence for such mutations in humans. These results indicate that the increased mutation rates resulting from MMR defects are capable of driving the development of cancer, and that certain heterozygous MMR
missense mutations can act in a dominant negative manner to increase mutation rates in normal tissue.
2.4 RESULTS

*Analysis of MSH6 missense mutations in S. cerevisiae.*

Previous studies described three dominant mutations in the *S. cerevisiae MSH6* gene that cause high rates of both base substitution and frameshift mutations, in contrast to an *msh6* deletion mutation, which only causes high rates of base substitution mutations. Two of the affected amino acids (*S. cerevisiae* codons Ser1036 and Gly1142) are conserved in the human and mouse MSH6 proteins whereas in one case, Gly1067Asp, the amino acid at the equivalent residue in the human and mouse protein is a Thr (residues 1219 and 1217, respectively). Although the Gly1067Asp substitution is not located in the P-loop domain, which is required for ATP binding and processing, molecular modeling indicates that it is situated at the MSH2-MSH6 heterodimer interface close to the P-loop of the MSH2 protein (Hess et al., 2002). A recent study reported an MSH6 mutation causing a Thr to Ile change at codon 1219 in an HNPCC case with a tumor with MSI, although little information was presented supporting the pathogenic effects of this mutation (Berends et al., 2002). To study the effect of the Thr and Ile residues on MSH6 function, mutations altering *S. cerevisiae MSH6* codon 1067 to either a Thr or Ile codon were studied on both an ARS CEN plasmid and at the chromosomal *MSH6* locus. On an ARS CEN plasmid in a wild-type strain, the wild-type *MSH6* control and the 1067Thr substitution did not alter the rate of reversion of the *lys2-A10* frameshift mutation (detects deletions in a run of 10 As) whereas the 1067Ile substitution caused a large increase in mutation rate seen in a patch test (Figure 1A). Similar results were obtained using the *CAN1*
forward mutation assay (detects gene-inactivating frameshift and base substitution mutations) and the \textit{hom3-10} frameshift reversion assay (detects deletions in a run of 7 Ts) (data not shown). When these plasmids were tested for their ability to complement an \textit{msh6 msh3} double mutant, the wild-type MSH6 control and the MSH6 1067Thr substitution reduced the rate of reversion of the \textit{hom3-10} frameshift mutation similarly, whereas the vector and the MSH6 1067Ile substitution had no effect (Figure 1B). Similar results were obtained using the \textit{CAN1} forward mutation assay and the \textit{lys2-A10} frameshift reversion assay (data not shown). Finally, the MSH6 1067Asp, Thr and Ile substitutions were introduced at the chromosomal \textit{MSH6} locus and tested for their effects on the \textit{CAN1} forward mutation rate and the \textit{hom3-10} frameshift mutation reversion rate (Figure 1C). The MSH6 1067Thr substitution did not significantly increase the mutation rate in either assay compared to the wild type strain. In contrast, the MSH6 1067Asp and MSH6 1067Ile substitutions had essentially the same effect and increased the \textit{hom3-10} frameshift reversion rate 8 to 10-fold more than the \textit{msh6} deletion mutation and increased the \textit{CAN1} mutation rate about 2-fold more than the \textit{msh6} deletion mutation; these mutation rates are about 50\% of that seen in an \textit{msh2} deletion mutant. These results demonstrate that substituting a Thr for Gly1067 has no effect on MSH6 function, whereas substitution of an Ile at this residue results in loss of MSH6 function and a strong dominant defect in the repair of frameshift mutations as observed for the original dominant \textit{msh6} mutations.
Figure 2-1. Genetic analysis of msh6 mutations in yeast.

A: The indicated plasmids were transformed into the wild type strain RDKY 3686 and the presence of a dominant mutator phenotype was detected by determining the relative rate of reversion of the lys2::InsE-A10 allele by replica plating the strains onto SD plates lacking lysine.

B: The indicated plasmids were transformed into the RDKY 4234 msh6Δ::hisG msh3Δ::hisG strain and complementation of the msh6Δ::hisG mutation was detected by determining the relative rate of reversion of the hom3-10 allele by replica plating the strains onto SD plates lacking threonine.

C: Fluctuation analysis was used to determine the hom3-10 reversion rate and the CAN1 forward mutation rate of the indicated strains. The numbers given are the mutation rate followed by the fold increase in mutation rate relative to the wild type rate in parentheses.
Generation of Msh6\textsuperscript{T1217D} mutant mice.

We previously generated a mouse line carrying an Msh6 null allele (Msh6\textsuperscript{-/-}) (Edelmann et al., 1997). Msh6\textsuperscript{-/-} mice are viable and highly predisposed to the development of tumors. The tumor spectrum in these mice resembles that of MMR-deficient human patients and includes lymphoid tumors, gastrointestinal tumors and tumors of the skin. Msh6 deficiency causes a defect in the repair of base substitution mutations but not of 2 bp insertion/deletion mutations. Consequently, the tumors derived from these mice do not display an MSI phenotype. However, a large proportion of MSH6 mutations in human patients represent missense mutations, whose functional consequences remain unknown (Berends et al., 2002; Kolodner et al., 1999; Wijnen et al., 1999; Wu et al., 1999). We therefore created a mouse line carrying the Msh6\textsuperscript{T1217D} mutation to assess the effect of an Msh6 missense mutation on MMR function and cancer susceptibility (Figures 2A and 2B); the Msh6\textsuperscript{T1217D} allele is referred to as Msh6\textsuperscript{TD} throughout the manuscript. The Msh6\textsuperscript{TD} allele is transmitted in a normal Mendelian ratio, and both heterozygous and homozygous mutant mice develop normally. RT-PCR and Western blot analysis showed that the Msh6\textsuperscript{TD} mutation did not interfere with normal gene expression or the stability of the mutant protein (Figures 2C and 2D). In human cells, MSH6 forms a complex with MSH2 to initiate the repair of mispairs and loss of MSH6 affects the stability of the MSH2 protein (de Leeuw et al., 2000; Palombo et al., 1995). Consistent with this we detected a reduction of Msh2 protein in Msh6\textsuperscript{-/-} MEF cells. However, we did not detect a similar reduction in the Msh2 protein levels in Msh6\textsuperscript{TD+} or Msh6\textsuperscript{TD/TD} cell extracts indicating that the mutation had no effect on the stability of Msh2 protein (Figure 2D).
Figure 2-2. Generation of $Msh6^{T1217D}$ mice

A: Gene targeting strategy for the generation of $Msh6^{T1217D}$ mice showing the pMsh6T1217D knock-in targeting vector, the wild-type and targeted genomic $Msh6$ locus before and after Cre-loxP mediated deletion of the resistance cassette.

B: Representative Southern-blot analysis of NdeI-digested tail genomic DNA from $Msh6$ wild-type (WT), heterozygous (TD/+) homozygote (TD/TD) and heterozygote undeleted (TDhyg+) mice.

C: RT-PCR analysis of total RNA isolated from $Msh6$ wild-type (WT), heterozygous (TD+/+) and homozygous (TD/TD) mutant ES cells. The TD mutation creates an SfaNI site in the mutant transcript. The RNA transcript in WT mice is resistant to SfaNI digestion (246bp fragment) while the mutant transcripts in TD/+ and TD/TD mice can be digested into 221bp and 25 bp (not shown) restriction fragments. The presence of the TD mutation in the RT-PCR products of $Msh6^{TD/+}$ and $Msh6^{TD/TD}$ mice was further verified by DNA sequencing.

D: Western blot analysis shows expression of stable Msh2 and Msh6 protein in WT, heterozygous (TD/+) and homozygous (TD/TD) mutant MEF cell extracts, while no Msh6 protein and reduced amounts of Msh2 is detected in Msh6-deficient (-/-) extracts.
Immunohistochemical analysis also showed that the subcellular distribution of the mutant Msh6^{TD} or Msh2 protein was not affected (data not shown).

Mismatch repair activities in Msh6^{TD} mutant extracts.

We next analyzed the effect that the mutant Msh6^{TD} protein had on mismatch recognition and repair in ES cell extracts. Using gel mobility shift assays, we found that extracts of Msh6^{TD+/+} and Msh6^{TD/TD} embryonic stem cells contained a G•T mismatch binding activity that was similar to that observed in Msh6^{+/+} extracts (Figure 3A); the activity present in all of the extracts also bound homoduplex DNA, albeit to a lower extent. However, in contrast to Msh6^{+/+} extracts the G•T binding activity present in both Msh6^{TD+/+} and Msh6^{TD/TD} extracts was resistant to ATP-dependent mismatch release, even at high ATP concentrations (see Figure 3A). These results are consistent with previous studies of the corresponding msh6 mutation in yeast and indicate that the mutant Msh2-Msh6^{TD} complex displays a defect in the modulation of mispair binding and dissociation by ATP (Hess et al., 2002).

To determine the impact of the Msh6^{TD} mutation on DNA repair, we measured the in vitro DNA mismatch repair activity in ES cell extracts using a variety of substrates containing G•G mismatches, single-base insertion or two-base insertion mismatches with a nick either 5' or 3' to the mismatched base (Figure 3B). The Msh6^{+/+} and Msh6^{TD+/+} extracts were proficient in the repair of all substrates tested whereas the Msh6^{TD/TD} extracts did not repair any of these substrates. The repair defect in the Msh6^{TD/TD} extracts differs from the defect previously found in Msh6^{-/-} extracts, which are defective in the repair of base-base mismatches but not single-base
Figure 2-3. Mismatch repair activities in Msh6 mutant cell extracts.

A: Mismatch binding activities of Msh6+/+, Msh6TD/+ and Msh6TD/TD nuclear extracts. Gel mobility shift assay in Msh6 mutant nuclear extracts was performed by incubating nuclear protein extracts isolated from Msh6+/+ (+/+), Msh6TD/+ (TD/+) and Msh6TD/TD (TD/TD) ES cells with either GC- containing homoduplex or G/T-containing heteroduplex oligonucleotides. The position of the MutSα-DNA complex and the unbound free probe are indicated by the arrows. Unlabeled (cold) G/T-containing heteroduplex competitor oligonucleotide or ATP was included in the reaction mixture in the molar ratio or concentration indicated.

B: MMR defects in Msh6TD mutant cell extracts. DNA mismatch repair activity was assayed in Msh6TD/TD, Msh6TD/+ and Msh6+-/− cell extracts as described (Thomas et al., 1995). Substrates designated with a Ω contain the number of extra nucleotides that accompany the symbol. Substrates are designated 3’ or 5’ indicating the position of the nick relative to the mismatch. The 3’ nick is in the (-) strand at the Ava II site (position -264). The 5’ nick is in the (-) strand at the Bsu36 I (position +276). The nucleotide position of the mismatch or unpaired bases in the lacZ complementation gene is indicated after the @, where position +1 is the first transcribed base of the lacZ complementation gene. The (-) sign designates the strand containing the extra nucleotide(s). The results are averages based on counting over 500 plaques per variable in at least three independent experiments.
insertion and dinucleotide insertion mispairs (Edelmann et al. 1997, Edelmann et al. 2000). Furthermore, addition of increasing amounts of Msh6\(^{TD/TD}\) extracts to Msh6\(^{+/+}\) extracts inhibited the repair of dinucleotide insertion mutations compared to control assays in which the same amounts of Msh6\(^{-/-}\) extracts were added to Msh6\(^{+/+}\) extracts (data not shown). These results demonstrate that in homozygous mutant extracts the mutant Msh6\(^{TD}\) protein interferes with the ability of Msh2-Msh3 complexes to act in repair of dinucleotide insertion/deletion mutations.

Microsatellite Instability in Msh6\(^{TD}\) mutant mice.

The repair defect observed in Msh6\(^{TD/TD}\) cell extracts suggested that the in vivo mutator phenotype in Msh6\(^{TD}\) mutant mice would differ from wild-type and also Msh6\(^{-/-}\) mice. To assess this, the MSI phenotype was analyzed in tail genomic DNA of Msh6\(^{+/+}\), Msh6\(^{-/-}\), Msh6\(^{TD/+}\) and Msh6\(^{TD/TD}\) mice at a mononucleotide repeat locus (U12235) and two dinucleotide repeat loci (D7Mit91 and D17Mit123) (Figures 4A and 4B). The Msh6\(^{TD/TD}\) mice displayed a significant increase in MSI at all three loci as compared to Msh6\(^{+/+}\) mice (Figure 4B). The increase in MSI in the Msh6\(^{TD/TD}\) was also significant compared to the MSI observed in Msh6\(^{-/-}\) mice. In addition, the MSI at these loci in Msh6\(^{TD/TD}\) was comparable to the MSI previously found in Mlh1\(^{-/-}\) mice (Wei et al., 2003). Interestingly, we also noted a significant increase in MSI in Msh6\(^{TD/+}\) mice at the U12235 locus as compared to wild-type mice. The increase in MSI in Msh6\(^{TD/+}\) mice at the D7Mit91 locus was of borderline significance and not
Figure 2-4. MSI in genomic DNA of Msh6 mutant mice.

A: Genomic DNA from wild-type (Msh6\(^{+/+}\)), Msh6 null (Msh6\(^{-/-}\)) and Msh6\(^{TD/TD}\) homozygous mutant mice (Msh6\(^{TD/TD}\)) was analyzed for MSI at marker D7Mit91. Asterisks indicate unstable alleles.

B: Comparison of MSI in genic DNA of Msh6 mutant mice

<table>
<thead>
<tr>
<th>Marker</th>
<th>Msh6(^{+/+})</th>
<th>Msh6(^{-/-})</th>
<th>Msh6(^{TD/+})</th>
<th>Msh6(^{TD/TD})</th>
<th>Mlh1(^{+/a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>U12335 (A)(_h)</td>
<td>5% (5/102)</td>
<td>9% (6/68)</td>
<td>15% (16/108)(^1)</td>
<td>22% (24/108)(^2,3,4)</td>
<td>16% (12/73)</td>
</tr>
<tr>
<td>D7Mit91 (CA)(_h)</td>
<td>2% (3/124)(^a)</td>
<td>3% (3/100)</td>
<td>8% (14/182)(^5)</td>
<td>12% (19/158)(^2,3,4)</td>
<td>15% (19/123)</td>
</tr>
<tr>
<td>D17Mit123 (CA)</td>
<td>4% (4/99)</td>
<td>6% (6/132)</td>
<td>10% (11/116)</td>
<td>14% (19/141)(^2,3,4)</td>
<td>28% (19/69)</td>
</tr>
</tbody>
</table>

Mutation frequency was determined as the number of unstable alleles divided by the number of total alleles scored. \(^a\) data from Wei et al. 2003. Msh6\(^{TD/+}\) vs. Msh6\(^{+/+}\): \(^1p < 0.05, \(^2p = 0.073, \) Msh6\(^{TD/TD}\) vs. Msh6\(^{-/-}\): \(^3p < 0.05; Msh6\(^{TD/TD}\) vs. Msh6\(^{-/-}\): \(^4p < 0.05.\) Fisher’s exact test.
significant at the D17Mit123 locus as compared to wild-type mice. These results demonstrate that the Msh6TD mutation results in a significant MSI phenotype.

In vivo mutation frequencies and spectra in Msh6 mutant mice.

The analysis of MSI only allows the detection of insertion/deletion mutations and might not reflect all of the features of the mutator phenotype in the Msh6 mutant mice. Therefore, we analyzed the mutation frequency at the cII locus in spleenocytes using the Big Blue transgene system, which allows the analysis of insertion/deletion mutations and base substitutions (Andrew et al., 2000; Kohler et al., 1991). There was a significant increase in mutation frequency in the Msh6TD/ TD mice and also to a slightly lesser extent in the Msh6+/- mice (Figure 5A). The difference between Msh6TD/ TD and Msh6+/- was significant (P < 0.0001). Interestingly, this analysis also revealed a 2- to 3-fold increase in the mutation frequency in Msh6TD/+ mice as compared to wild type mice (P < 0.0001), which was not observed in the Msh6+/- mice.

The analysis of the mutation spectra showed that the mutations in Msh6+/+, Msh6+/− and Msh6−/− mice were predominantly base substitutions with a majority of mutations being transitions and only some transversions (Figure 5B). The mutation spectrum in Msh6TD/+ mice largely resembled the spectrum found in wild type or Msh6 null mutant mice. In contrast, in Msh6TD/ TD mice almost half of the mutations were single base deletions or insertions. The difference between the mutation spectra in Msh6TD/ TD and Msh6+/+ or Msh6−/− mice is highly significant (P values < 0.0001) and
Figure 2-5. Increased mutation frequencies in Msh6 mutant mice.

A: Comparison of mutation frequencies at the cII locus in Msh6 mutant mice.

B: Comparison of cII mutation spectra in Msh6 mutant mice.
indicates that the Msh6<sup>TD</sup> mutation interferes with the repair of insertion/deletion mispairs.

**Survival and cancer phenotype in Msh6<sup>TD</sup> mutant mice.**

To study survival and cancer susceptibility, we followed a cohort of Msh6<sup>TD</sup> mutant mice for a period of up to 24 months. The Msh6<sup>TD/ TD</sup> mice had a significantly reduced survival compared to Msh6<sup>TD/+</sup> or wild-type mice (P values < 0.0001; log rank test) (Figure 6). The 50% survival of Msh6<sup>TD/TD</sup> mice was at 12 months of age and all of the mice died by 20 months of age. Although the overall reduction in survival of Msh6<sup>TD/TD</sup> mice was comparable to Msh6<sup>−/−</sup> mice (P = 0.91), there was a significant difference in survival during the first 10 months of life. At this age more than 80% of Msh6<sup>TD/TD</sup> mice were still alive compared to only 60% of Msh6<sup>−/−</sup> mice (P < 0.05).

There was also a significant reduction in the survival of Msh6<sup>TD/+</sup> mice compared to their wild-type littermates (P < 0.004). While 50% of the Msh6<sup>TD/+</sup> mice died by 20 months of age, more than 80% of wild-type mice were still alive.

The reduction in survival in Msh6<sup>TD/TD</sup> mice was caused by increased cancer susceptibility. To examine the types of tumors present in the Msh6<sup>TD/TD</sup> mutant mice, we analyzed the tumor spectrum in 12 mice that were between 7 and 14 months of age. This showed that similar to Msh6<sup>−/−</sup> mice, the majority of Msh6<sup>TD/TD</sup> mice had developed B- or T-cell non-Hodgkin’s lymphomas (11/12 mice) (Figures 7A and 7B) (Edelmann et al., 1997). We also found tumors in the small intestines of the mice, including tubular adenomas (2/12 mice), a tubulovillous adenoma (1/12 mice), and
Figure 2-6. Reduced survival of Msh6<sup>TD</sup> mutant mice.
The time of death or the time when the mice became moribund was recorded. The survival curves were generated by using GraphPad Prism 3.0 software. Statistical analysis was performed according to log-rank test. Solid green line: Msh6<sup>+/+</sup> mice (n=25); solid blue line: Msh6<sup>TD/+</sup> mice (n=29); solid red line: Msh6<sup>TDTD</sup> mice (n=25); solid black line: Msh6<sup>−/−</sup> mice (n=25).
focal areas of dysplasia (4/12 mice) (Figure 7C). Some of the mice also developed basal cell carcinomas of the skin (2/12). We analyzed several of the tumors for the presence of Msh6TD protein either by immunohistochemical analysis (Figure 7D) or Western blot analysis of micro-dissected tumor tissue (not shown). We found that in all tumors (n=4) that were analyzed, Msh6TD protein could readily be detected. Similarly, immunohistochemical analysis of intestinal tumors showed that while Msh3 was absent in the tumor cells in Msh3/- mice (Figure 7E), Msh3 could readily be detected at normal levels in the tumor cells in Msh6TD/TD mice (Figure 7F).

To assess whether the reduction in survival of Msh6TD/+ mice was associated with increased tumorigenesis, we sacrificed 11 Msh6TD/+ mice and 10 Msh6+/+ littermates that were still alive between 17 and 22 months of age. A majority of Msh6TD/+ mice had developed non-Hodgkin’s lymphoma (7/11 mice) at this age while none of the wild-type mice had detectable tumors. This analysis did not include the Msh6TD/+ mice that had died previously and therefore is probably an underestimate of the tumor phenotype. Some of the tumors in the heterozygous mice were also analyzed for the status of the wild-type Msh6 allele by loss of heterozygosity (LOH) analysis. This showed that in all tumors (n=5) tested the wild-type allele was retained indicating that tumor formation in Msh6TD/+ mice is not associated with LOH at the wild-type allele (data not shown).

The MMR defect observed in Msh6TD/TP extracts and the mutator phenotype observed in normal tissue of Msh6TD mutant mice suggested the tumors in Msh6TD
Figure 2-7. Tumors in Msh6TD mutant mice.

A: Section of mesenteric lymphoma in an Msh6TD/ mouse. Histologically it was classified as diffuse non-Hodgkin’s lymphoma, mixed large non-cleaved and cleaved cell type. (200 X).

B: Section of mesenteric lymphoma stained with a B cell-specific antibody (400 X).

C: Adenoma in the small intestine of an Msh6TD/ mouse. Histological features showed that the tumor was composed of glandular structures (40 X).

D: Section of intestinal tumor in an Msh6TD/ mouse stained with an Msh6 specific monoclonal antibody. The Msh6 protein expression was mostly detected in nuclei of tumor cells (400 X).

E: Section of intestinal tumor in an Msh3+/ mouse stained with an Msh3 specific monoclonal antibody. The Msh3 protein expression was absent in tumor cells (400 X).

F: Section of intestinal tumor in an Msh6TD/ mouse stained with an Msh3 specific monoclonal antibody. The Msh3 protein expression was mostly detected in the nuclei of tumor cells (400 X).

G – H: MSI at dinucleotide marker D7Mit91 in an Msh6TD/ GI adenoma (G) and an Msh6TD/+ lymphoma (H).

I – J: MSI at mononucleotide marker U12235 in an Msh6TD/ GI adenoma (I) and an Msh6TD/+ lymphoma (J).
mice would show MSI. Indeed, we found that tumors in both \textit{Msh6}^{TD/TD} homozygous and \textit{Msh6}^{TD/+} heterozygous mice displayed an MSI phenotype (Figures 7G-J). In \textit{Msh6}^{TD/TD} mice 5 of 9 (5/9) tumors analyzed were unstable at \textit{D7Mit91}, 3/9 tumors were unstable at \textit{D17Mit123} and 7/9 tumors were unstable at \textit{U12235}. A similar analysis in \textit{Msh6}^{TD/+} mice showed that 2/5 tumors were unstable at \textit{D7Mit91}, 1/5 tumors at \textit{D17Mit123} and 1/5 tumors at \textit{U12235}.

\textit{DNA damage response in Msh6}^{TD/TD} \textit{MEF cells.}

Numerous studies have indicated that MMR not only repairs mismatched bases, but also mediates an apoptotic response after exposure of cells to DNA damaging agents (Buermeyer et al., 1999; Li, 1999; Modrich, 1997). In addition, tumor cell lines or mouse cell lines that are deficient in either MSH2 or MSH6 display increased resistance to the cytotoxic effect of a variety of genotoxic agents (de Wind et al., 1999; Drummond et al., 1995; Fink et al., 1997; Umar et al., 1997; Zhang et al., 1999). To study the effect of the Msh6\textit{TD} mutation on this MMR function, we established \textit{Msh6}^{+/+}, \textit{Msh6}^{-/-} and \textit{Msh6}^{TD/TD} MEF strains and analyzed their response to exposure to several DNA damaging agents. \textit{Msh6}^{-/-} cells were resistant to the genotoxic effects of cisplatin at the drug levels tested. In contrast, \textit{Msh6}^{TD/TD} cells were as sensitive to cisplatin treatment as \textit{Msh6}^{+/+} cells (Figures 8A and 8B). The difference in cisplatin sensitivity between \textit{Msh6}^{-/-} and \textit{Msh6}^{+/+} or \textit{Msh6}^{TD/TD} cells was highly significant (\textit{P} < 0.0001). Similarly, \textit{Msh6}^{-/-} cells were largely resistant to
Figure 2-8. Drug sensitivity and apoptosis in Msh6 mutant cells.
MEF strains of the different Msh6 genotypes were exposed to cisplatin, MNNG and 6-TG at varying concentrations and for different time periods.
A: Survival of cells after 48h exposure at different cisplatin concentrations.
B: Survival of cells after exposure to 80 µM cisplatin at different time intervals.
C: Apoptotic response to cisplatin treatment (20 µM cisplatin for 24h).
D: Survival of cells after 48h exposure at different MNNG concentrations.
E: Survival of cells after exposure with 1.0 µg/ml MNNG at different time intervals.
F: Apoptotic response to MNNG treatment (1.0 µg/ml MNNG for 24h).
G: Survival of cells after 72h exposure at different 6-TG concentrations.
H: Survival of cells after exposure with 20 µM 6-TG at different time intervals.
I: Apoptotic response to 6-TG treatment (10 µM 6-TG for 24h).
treatment with MNNG or 6-TG, while Msh6\textsuperscript{TD/TD} cells remained as sensitive to
treatment with these agents as Msh6\textsuperscript{+/+} cells (Figures 8D–E and Figures 8G–H,
respectively). Furthermore, the sensitivity of the Msh6\textsuperscript{+/+} and Msh6\textsuperscript{TD/TD} cells to all
three DNA damaging agents was associated with a significant increase in the number
of apoptotic cells as compared to untreated cells (P values < 0.0001) (Figures 8C, 8F
and 8I). In contrast, no significant increase in the number of apoptotic cells was seen
with Msh6\textsuperscript{-/-} cells under these conditions. These results indicate that although the
Msh6\textsuperscript{TD} mutation causes DNA repair deficiency, it does not affect the apoptotic
response to DNA damage.
2.5 DISCUSSION

Previous studies in yeast have defined a number of novel dominant missense mutations in *MSH6* that cause a stronger mutator phenotype than deletion of *MSH6* (Das Gupta and Kolodner, 2000). We have investigated whether the equivalent of one such mutation, resulting in the amino acid substitution Thr1217Asp, would be dominant in the mouse and what the phenotypic consequences of such a mutation might be. Extensive analysis of both ES cells and normal mouse tissues containing this mutation indicate that it causes dominant defects in MMR similar to those caused by the yeast mutation. The mouse Thr1217Asp amino acid substitution resulted in an MSH2-MSH6 complex that could bind mispaired bases but was resistant to ATP-induced release. *In vitro*, the homozygous mutation resulted in a defect in the repair of dinucleotide insertion/deletion mispairs in contrast to an *Msh6* deletion mutation, which does not cause such defects. In addition, in normal mouse tissues, the homozygous *Msh6<sup>TD</sup>* mutation caused both dinucleotide repeat instability and increased frameshift mutations in a reporter transgene in contrast to the weaker mutator phenotype caused by an *Msh6* deletion mutation. Normal tissues that were heterozygous for the *Msh6<sup>TD</sup>* mutation also showed increased MSI and accumulation of mutations in the reporter transgene. The homozygous mutant *Msh6<sup>TD</sup>* mice had a cancer susceptibility phenotype which was similar to that caused by an *Msh6* deletion mutation, except that the tumors showed dinucleotide repeat instability in contrast to an absence of dinucleotide repeat instability in tumors from *Msh6* deletion mutant
animals (Edelmann et al., 1997). Importantly, the heterozygous \(Msh6^{TD}\) missense mutant mice had significantly increased cancer susceptibility predicted to result from a dominant mutation. Consistent with this, a human \(MSH6\) mutation resulting in the substitution of Ile at the equivalent amino acid reported in a suspected HNPCC case associated with dinucleotide repeat instability was a dominant mutation when modeled in yeast. These results illustrate the complementary insights that can be obtained through the analysis of yeast, mouse and human systems.

The analysis of microsatellite instability in human tumors indicates that in contrast to tumors with \(MSH2\) or \(MLH1\) mutations, mutations in \(MSH6\) are often associated with a low microsatellite instability phenotype (MSI-L) at dinucleotide markers (Akiyama et al., 1997; Kolodner et al., 1999; Verma et al., 1999; Wijnen et al., 1999; Wu et al., 1999). However, in recent studies of suspected HNPCC cases with \(MSH6\) missense mutations, several tumors have been identified that display high microsatellite instability (MSI-H) (Berends et al., 2002). Our studies with \(Msh6^{-/-}\) and \(Msh6^{TD/TD}\) mutant mice provide an explanation for these differences and indicate a genotype–phenotype correlation for this type of genetic instability in \(MSH6\) mutant tumor cells. In \(Msh6^{-/-}\) tumors, the lack of Msh6 protein does not interfere with the function of the Msh2-Msh3 heterodimer, while the presence of the mutant Msh6\(^{TD}\) protein in tumor cells apparently interferes with the Msh2-Msh3 mediated repair of dinucleotide insertion/deletion mutations, paralleling the impairment of the repair of insertion/deletions in nuclear extracts. This predicts that in \(Msh6\) mutant tumors that display an MSI-H phenotype, the mutant Msh6 protein will be retained. In addition, loss of MSH3 protein is not required for the MSI phenotype. Indeed, we observed that
all of the tumors tested in the mouse remained positive for Msh6 and Msh3 immunostaining. Our results are consistent with a recent study of MSH6 tumors in human HNPCC and HNPCC-like patients that documented MSI-H colorectal and endometrial tumors that were positive for MSH6 staining and MSI-L tumors that were negative for MSH6 staining (Berends et al., 2002). Furthermore, the human MSH6 Thr1219Ile mutation found to be a dominant mutation in yeast was identified in a suspect HNPCC case with an MSI-H tumor. It is therefore likely that the MSI phenotype in MSH6 mutant tumors is a function of the type of MSH6 mutation and the stability of the mutant MSH6 protein present in the tumor cells. It should be noted that in cases where MSH6 is not expressed, an MSI-H phenotype can occur due to a secondary mutation in MSH3, a phenomenon seen in both yeast and a small number of suspected HNPCC cases (Akiyama et al., 1997; Marsischky et al., 1996; Sia et al., 1997).

Other important roles for MMR in addition to the correction of mispaired bases have been recognized. One of these functions is the processing of DNA damage and the subsequent induction of an apoptotic response. As a consequence, mammalian cells deficient in MSH2, MSH6, MLH1 and PMS2 display resistance to DNA-damaging agents such as 6-TG, cisplatin and MNNG (Buermeyer et al., 1999; de Wind et al., 1999; Drummond et al., 1995; Duckett et al., 1996; Fink et al., 1997; Fishel, 2001; Koi et al., 1994; Li, 1999; Modrich, 1997; Umar et al., 1997). Complementation by chromosome transfer of MSH2, MSH6 and MLH1 reversed the resistance phenotype in mammalian cells and clearly implicated the MMR system in this response (Koi et al., 1994; Umar et al., 1997). Because the Msh2-Msh6 complex
recognizes DNA adducts such as O\textsuperscript{6}-methylguanine and cisplatin adducts, it has been suggested that MMR proteins may act in sensing DNA damage and also directly participate in the processing of damaged DNA residues (Fishel, 2001; Li, 1999; Modrich, 1997). Several different models have been developed to explain the role of MMR proteins in DNA damage responses. In one model, DNA repair-competent cells engage in futile repair cycles after treatment with alkylating agents, and these futile repair cycles lead to the formation of double strand breaks that signal cell cycle arrest and apoptosis (Karran and Bignami, 1994). Alternatively, it has been suggested that MMR proteins may function as damage sensors directly linked to apoptotic responses via a signal transduction cascade or indirectly linked by binding to damaged bases, which subsequently blocks DNA replication or other processes, such as transcription and repair, leading to cell death (Fishel, 2001; Li, 1999; Modrich, 1997). The finding that the mutant Msh2-Msh6\textsuperscript{TD} complex does not function in MMR but is still capable of mismatch binding and of initiating apoptosis in response to cisplatin, MNNG and 6-TG exposure supports the idea that Msh2-Msh6 complexes can function as DNA damage sensors and suggest that excision of DNA lesions is not required for the DNA damage response function of MMR.

While it is generally accepted that an increase in mutation rate due to MMR defects is an important factor underlying tumorigenesis, the importance of defects in MMR-mediated DNA damage-induced apoptosis remains unclear. It has been proposed that mutations in MMR genes, which result in resistance to DNA damage induced apoptosis, provide a selective advantage in the initial stages of tumorigenesis (Fishel, 2001). In this model, inactivation of MMR and the resulting increased
mutation rates accelerate tumorigenesis but are not the origin of it. The observation that Msh6\textsuperscript{+/−} ES cells, but not Msh3\textsuperscript{+/−} ES cells, display resistance to MNNG is consistent with the unique involvement of Msh2-Msh6 (but not Msh2-Msh3) in DNA damage response (de Wind et al., 1999), and it was suggested that this difference in function may explain the prevalence of mutations in MSH2 and MSH6 but not MSH3 in HNPCC tumors (Heinen et al., 2002). In addition, it was suggested that the tissue selectivity of HNPCC tumors might be traced to exposure to certain types of DNA damaging processes such as alkylation or oxidation in target tissues (Fishel, 2001).

The analysis of Msh6\textsuperscript{+/−} and Msh6\textsuperscript{TD/TD} mice indicates that the apparently normal DNA damage response in Msh6\textsuperscript{TD/TD} mutant mice did delay the onset of tumorigenesis during the first 10 months of life. However, it did not alter either the overall survival or tumor spectrum of Msh6\textsuperscript{TD/TD} mice as compared to Msh6\textsuperscript{+/−} mice. This suggests that the increased mutator phenotype in the mice is sufficient to drive tumorigenesis regardless of the status in DNA damage response function. In a similar analysis of mice with an Msh2\textsuperscript{G674A} missense mutation, which were MMR-deficient but retained normal apoptotic signaling, we also observed a strong cancer phenotype (Lin et al., 2004). Like the Msh6\textsuperscript{TD/TD} mice, the Msh2\textsuperscript{G674A/G674A} mice also showed a delayed tumor onset as compared to Msh2\textsuperscript{+/−} mice, suggesting that the DNA damage response function in these mice could inhibit tumorigenesis at an early stage. However, the age of tumor onset of both Msh6\textsuperscript{TD/TD} and Msh6\textsuperscript{+/−} mice is delayed compared to Msh2\textsuperscript{+/−} and Msh2\textsuperscript{G674A/G674A} mice, which is possibly due to a stronger mutator phenotype in the Msh2 mutant mice. A similar relative mutator phenotype is seen in yeast (Das Gupta and Kolodner, 2000). Based on these studies in mice, it seems likely that the MMR
functions in DNA repair and apoptosis cooperate in tumor suppression but that the increase in mutation rates resulting from MMR defects is sufficient to drive tumorigenesis. Unfortunately, it has not yet been possible to identify mutations in MMR genes that inactivate apoptosis but not DNA repair to evaluate whether defects in MMR-induced apoptosis also can drive tumorigenesis. One implication of the observation that MMR missense mutations can differentially affect DNA repair and MMR-induced apoptosis is that the clinical characteristics of HNPCC caused by missense mutations may be more heterogeneous than that caused by complete loss of function mutations. In addition, human tumors with certain \textit{MSH6} missense mutations will display MSI and may remain responsive to treatment with chemotherapeutic agents. Therefore, determining the genotype/phenotype correlations of \textit{MSH6} missense mutations may provide valuable information for prevention, treatment and prognosis of individuals with such mutations.
2.6 EXPERIMENTAL PROCEDURES

Yeast procedures.

All yeast genetics methods have been described in detail (Amin et al., 2001; Das Gupta and Kolodner, 2000; Marsischky et al., 1996). Yeast cells were grown in YEPD or synthetic drop-out (SD) medium with or without 2% bactoagar. SD medium was supplemented with the appropriate drop out mix of amino acids (BIO101, Inc., Vista, CA) and canavanine plates were SD arginine medium supplemented with 60 mg/L of canavanine (SIGMA, St Louis, MO). Transformations were performed using standard procedures. Mutations were placed at the chromosomal MSH6 locus using standard pop-in, pop-out procedures and then the MSH6 gene present in each strain was amplified by PCR and sequenced to verify the presence of the desired mutation and the lack of additional mutations. Patch tests to identify mutator phenotypes were performed as described (Amin et al., 2001) and mutation rates were determined by fluctuation analysis using at least 14 independent cultures as described (Amin et al., 2001; Das Gupta and Kolodner, 2000; Lea and Coulson, 1948; Marsischky et al., 1996).

The yeast strains used in patch tests were isogenic derivatives of S288C (Amin et al., 2001). The wild type strain was RDKY 3686 MATα ura3-52 leu2Δ1 trp1Δ63 his3Δ200 hom3-10 lys2::InsE-A10 and the msh3 msh6 double mutant was a derivative RDKY 4234 that also contained the msh6Δ::hisG msh3Δ::hisG mutations (Marsischky et al., 1996). The strain background used for analysis of the effect of
chromosomal $msh6$ mutations on mutation rates is RDKY 2311 $MATa trp1 ura3-52 ade2-1 hom3-10 leu2-3112$. RDKY 2311 and the mutant derivatives RDKY 3519 $msh6Δ::hisG$ and RDKY 3660 $msh6-G1067D$ have been described (Das Gupta and Kolodner, 2000). The derivatives RDKY 5054 $msh6-G1067T$ and RDKY 5022 $msh6-G1067I$ were constructed for this study using $URA3$ integrative plasmids containing the $msh6-G1067T$ and $msh6-G1067I$ mutations, respectively.

The plasmids used were constructed as follows. pRS315 is a standard ARS $CEN$ $LEU2$ cloning vector and pRDK439 is a derivative from our lab collection containing a genomic $MSH6$ fragment spanning the BamHI site upstream of the $MSH6$ promoter to the first HindIII site downstream of $MSH6$ inserted into BamHI, HindIII-cut pRS315 (Das Gupta and Kolodner, 2000). The $msh6-G1067T$ and $msh6-G1067I$ mutations were introduced with the QuickChange™ Site Directed Mutagenesis kit (Stratagene) and the resulting plasmids were fully sequenced. $URA3$ integrative plasmid derivatives were then constructed by subcloning the BamHI, HindIII fragment into pRS306.

**Generation of $Msh6^{TD}$ mutant mice.**

A 10 kb EcoRI fragment containing the $Msh6$ region surrounding exon 8 was isolated from a 129SvEv BAC genomic library and subcloned into pBluescript. A mutation was introduced that changed codon 1217 from threonine (ACC) to aspartic acid (GAT) by site directed mutagenesis (Stratagene Quick Change Kit) and verified by sequencing. A 5.0 kb NotI fragment containing two loxP sites flanking a neomycin-PGKhygromycin resistance cassette was subcloned into a single Bsu36I site
in intron 7. The targeting vector was linearized by NotI restriction digestion and electroporated into WW6 embryonic stem cells (Edelmann et al., 1997). Four correctly targeted ES cell lines were injected into C57BL/6J blastocysts. Male chimeras that resulted form these injections were mated to C57BL/6J females and transmitted the mutant allele through their germ line. F1 males carrying the mutant allele were mated to Zp3Cre transgenic females (C57BL/6J) to remove the resistance cassette by LoxP-mediated recombination. Male and female mice carrying the modified allele were intercrossed to generate Msh6+/+, Msh6TD/+ and Msh6TD/TD mutant mice.

**RT-PCR analysis.**

Total RNA was isolated from Msh6TD mutant ES cell lines using the RNeasy purification Kit (Qiagen). RT-PCR was performed using forward primer: 5’-AAGACAGGCTGGTCTGTTGG-3’ and reverse primer: 5’-GCTGTCCCATCAAAA GTTGC-3’ using the Titan One Tube-RT-PCR reaction kit (Roche) according to the manufacturers instruction. The resulting 246 bp fragment was digested with SfaNI to detect the presence of the mutant RNA transcript.

**Western blot analysis.**

MEF cell extracts were prepared according to standard procedures and 50µg protein of each cell lysate was separated on a 10% SDS-PAGE gel. Protein was transferred onto PROTRAN membranes and the membranes were subsequently incubated with monoclonal antibodies directed against Msh2 (Ab-2, Oncogene), Msh6
(clone 44, BD Biosciences) and rabbit polyclonal antibody directed against β-tubulin (H-235, Santa Cruz).

_Gel mobility shift assays._

Nuclear extracts were prepared as described (Dignam et al., 1983). The invariant sense oligonucleotide 5'GGGAAGCTGCCAGGCCAGCCCATG GTCAGCCTCCTATGCTC-3' was end-labeled with γ<sup>32</sup>-ATP and annealed in 1 x DNA binding buffer (12 % glycerol, 20 mM HEPES pH 7.9, 100 mM KCl, 1 mM DTT, and 5 mM MgCl<sub>2</sub>) with 3 x molar ratios of antisense oligonucleotide: 5'-GAGCATAGGAG GCTGACACTGGGGCCTGGCAGCTTCCC-3' to form a GC homoduplex probe or with 3 x molar ratios of antisense oligonucleotide: 5'-GAGCATAGGAGGCTGACATTGGGGCCTGGCAGCTTCCC-3' to form a GT mismatch-containing heteroduplex probe. Ten µg of nuclear extract was pre-incubated in 1 x DNA binding buffer, 1 µg dIdC and 20 ng unlabeled homoduplex for 5 minutes on ice in a total volume of 19 µl. Twenty ng of radio-labeled DNA probe was subsequently added and the binding mixture was incubated on ice for 30 minutes. For reaction using cold probe competition, cold competitor (100 ng) was included in the preincubation mixture. For adenine nucleotide exchange experiments, ATP (0.1 to 60 mM) was added 15 minutes after addition of the DNA probe. The reaction mixture was then subjected to electrophoresis in a 5 % polyacrylamide gel in 0.5 x TBE buffer containing 2.5% glycerol. The gels were dried and autoradiographed.
Cell-free extracts and mismatch repair assay.

Procedures for extract and heteroduplex substrate preparation and for measuring repair activity were as described (Thomas et al., 1995). The substrates used are described in the legend to Figure 3B.

MSI Analysis.

Mutations in microsatellite sequences were assayed by PCR of single target molecules. Equal amounts of tail DNA from ten mice per mouse strain (Msh6+/+, Msh6−/−, Msh6TD+/ and Msh6TD/TD) were pooled and diluted to 0.5 to 1.0 genome equivalents. Cycling reactions for the three markers analyzed, U12235, D7Mit91 and D17Mit123 were performed as described previously (Wei et al., 2003).

In vivo mutation analysis.

The in vivo mutation frequency in spleenocytes of WT and Msh6 mutant mice was assessed using the target cII transgene in the Big Blue Transgenic Rodent Mutagenesis Assay System (Stratagene) according to the manufacturer’s guidelines. Three mice each from Msh6+/+Big Blue, Msh6−/+BigBlue, Msh6−/−Big Blue, Msh6TD+/Big Blue and Msh6TD/TD/Big Blue mouse strains were sacrificed at 10 weeks of age and analyzed. To characterize the cII locus in mutant phage particles, the entire cII gene was PCR amplified and sequenced. Statistical analysis was performed using the Fisher’s exact.
Analysis of tumors.

Tumors from sacrificed mice were removed and fixed in 10% neutral buffered formalin. All tumors were processed for paraffin embedding and sections were prepared for staining with Hematoxylin and Eosin according to standard procedures. For immunohistochemical analysis the tumor sections were deparaffinized and stained with antibodies for immunotyping (B lymphocyte CD45R/B220 (Pharmingen) and T lymphocyte CD3 (Vector and Zymed)) or antibodies directed against Msh6 (clone 44, BD Biosciences) and Msh3 (clone 52, BD Transduction Laboratories). Statistical analyses of tumor incidence was performed using the Fisher’s exact.

MEF survival analysis.

MEF cells (5 x 10⁴) of each Msh6 genotype were exposed to cisplatin, 6-TG or MNNG at different drug concentrations or for different time periods. After drug exposure the cells were washed once with PBS, once with PBS:Methanol (1:1), fixed in 0.5 ml 100% methanol. The cells were air dried and stained with 0.1 % crystal violet, extensively washed with PBS and the dye extracted in 10 % acetic acid. The dye concentration was determined by measuring absorption at A₆₀₀nm and the percentage cell survival was calculated as (treated cells/untreated cells) x 100. The apoptotic response to drug exposure was measured by TUNEL (DeadEnd Fluorometric TUNEL System, Promega). All experiments were performed for three different MEF strains for each Msh6 genotype and repeated at least three times for each strain. Cisplatin (Bedford Laboratories), 6-TG (Sigma) and MNNG (Sigma)
dilutions in culture medium were prepared fresh before use. For the exposure to MNNG, 20 µM O\textsuperscript{6}-benzylguanine (Sigma) was added to the medium.
2.7 ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health (CA76329 and CA93484 to W.E., GM50006 to R.D.K., ES11040 and CA84301 to R.K. and center grant CA13330 to the Albert Einstein College of Medicine), a Deutsche Krebshilfe fellowship (to S.J.S.) and an Irma T. Hirschl Career Scientist Award (to W.E.).

Chapter 2, in full, is a reprint of the material as it appears in Yang, G., Scherer, S. J., Shell, S. S., Yang, K., Kim, M., Lipkin, M., Kucherlapati, R., Kolodner, R. D., and Edelmann, W. (2004). Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. Cancer Cell 6, 139-150. The dissertation author was a secondary investigator and author of this paper and was responsible the work displayed in Figure 2-1 and described in the corresponding sections of the text.
2.8 REFERENCES


CHAPTER 3

The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA
3.1 SUMMARY

The eukaryotic MutS homolog complexes, Msh2-Msh6 and Msh2-Msh3, recognize mismatched bases in DNA during mismatch repair (MMR). The eukaryote-specific N-terminal regions (NTR) of Msh6 and Msh3 have not been characterized other than by demonstrating that they contain an N-terminal PCNA-interacting motif. Here we have demonstrated genetically that the NTR of Msh6 has an important role in MMR that is partially redundant with PCNA binding. Small angle X-ray scattering (SAXS) was used to determine the solution structure of the complex of PCNA with Msh2-Msh6 and with the isolated Msh6 NTR, revealing that the Msh6 NTR is a natively disordered domain that forms an extended tether between Msh6 and PCNA. Moreover, computational analysis of PCNA-interacting motifs in the S. cerevisiae proteome indicated that flexible linkers are a common theme for PCNA interacting proteins that may serve to localize these binding partners without tightly restraining them to the immediate vicinity of PCNA.
3.2 INTRODUCTION

Mismatch repair (MMR) is a conserved DNA repair pathway that increases the fidelity of DNA replication by repairing mispairs that occur during DNA replication (Harfe and Jinks-Robertson, 2000a; Iyer et al., 2006; Kolodner, 1996; Kolodner and Marsischky, 1999). Defects that inactivate this pathway result in mutation rate increases of up to several thousand-fold (Harfe and Jinks-Robertson, 2000b; Marsischky et al., 1996). In humans, germline MMR defects lead to hereditary non-polyposis colorectal cancer (HNPCC), and somatic defects in MMR are observed in sporadic tumors as well (Jacob and Praz, 2002).

MMR is best understood in Escherichia coli, where both the early steps in mispair detection and later steps in discrimination between the newly synthesized and template DNAs, strand nicking and resynthesis are understood (Iyer et al., 2006). The bacterial MutS homodimer that binds and recognizes mismatches is conserved in eukaryotes in the form of two partially redundant heterodimeric MutS homolog complexes with overlapping substrate specificities: Msh2-Msh6 and Msh2-Msh3 (Harfe and Jinks-Robertson, 2000a; Kolodner and Marsischky, 1999). The bacterial MutL homodimer, which subsequently binds to the MutS/mismatch complex, is also conserved as multiple heterodimeric complexes of MutL homologs (Harfe and Jinks-Robertson, 2000a; Kolodner and Marsischky, 1999). Later steps in eukaryotes are not as well understood; however, partial MMR reactions can be reconstituted in vitro using purified human proteins and mispaired DNA substrates with pre-introduced
nicks to direct mispair excision and resynthesis (Constantin et al., 2005; Zhang et al., 2005). How the nascent strand is recognized in eukaryotes in vivo is not understood.

PCNA was first implicated in eukaryotic MMR through genetic analysis of PCNA mutants in S. cerevisiae (Johnson et al., 1996), which had mutator phenotypes due to both replication and MMR defects (Chen et al., 1999; Merrill and Holm, 1998). PCNA is a conserved homotrimeric sliding clamp controlling processivity of DNA polymerases δ and ε (Burgers, 1991; Johnson and O'Donnell, 2005). PCNA functions in MMR prior to mispair excision as well as during resynthesis of the excised strand (Gu et al., 1998; Umar et al., 1996). PCNA is now implicated in many aspects of DNA replication and repair and interacts with translesion polymerases, the flap endonuclease Rad27/Fen1, DNA ligase, the chromatin assembly factor Cac1, DNA helicases, base-excision repair proteins Ung1/Ung2 and Apn2, the nucleotide-excision repair factor XPG, and the MMR proteins Msh6 and Msh3 (Burgers, 1991; Chen et al., 1996; Clark et al., 2000; Flores-Rozas et al., 2000; Gary et al., 1997; Haracska et al., 2001; Johnson and O'Donnell, 2005; Li et al., 1995; Otterlei et al., 1999; Prelich et al., 1987; Schmidt et al., 2002; Shibahara and Stillman, 1999; Subramanian et al., 2005; Unk et al., 2002; Zhang et al., 2000). Most PCNA-interacting proteins bind PCNA through a conserved sequence motif termed the PIP box (Warbrick, 2000).

The early role for PCNA in MMR may be related to interactions with Msh6 and Msh3. Mutations affecting the PIP box binding site on PCNA or the PIP box in Msh6 disrupt PCNA-Msh6 interactions in vitro and cause varying degrees of MMR defects in vivo (Chen et al., 1999; Clark et al., 2000; Flores-Rozas et al., 2000). Deletion of the PIP box in human Msh6 causes partial defects in MMR reactions in
vitro, and PCNA has been proposed to localize MMR factors to regions of newly replicated DNA (Kleczkowska et al., 2001). Additionally, PCNA may hand off Msh2-Msh6 to mispairs in newly replicated DNA (Lau and Kolodner, 2003); however, the \textit{in vivo} effects of mutations that disrupt \textit{in vitro} Msh6-PCNA interactions are small, and the genetic analysis performed to date does not fit with simple models whereby these mutations result in more MMR substrates or cause general MMR defects (Lau et al., 2002).

To understand the Msh2-Msh6-PCNA interaction, we have analyzed the role of the uncharacterized N-terminal region (NTR) of Msh6, which contains the PIP box and precedes the MutS homology (Sup. Fig. 1). Genetic analysis demonstrates that the NTR consists of at least two functional regions, one involved in PCNA interaction and one of unknown function, either one of which is sufficient for most MMR. The NTRs of Msh3 and Msh6 were interchangeable, and the Msh6 NTR was functional when located on Msh2. In addition we determined the solution structure of PCNA in complex with the Msh6 NTR as well as full-length and truncated versions of Msh2-Msh6 by small-angle X-ray scattering (SAXS). These data show that the NTR is a long, flexible linker even in the Msh2-Msh6-PCNA complex, and genetic analysis indicated that a minimum length of unstructured linker is required for functional interaction between PCNA and Msh2-Msh6. Furthermore, flexible linkers may represent a common theme for PCNA-interacting regions, supporting a scenario in which PCNA helps localize repair proteins to newly replicated DNA.
3.3 RESULTS

*The Msh6 NTR is hypersensitive to proteolysis.*

We initially identified the NTR/mispair-binding domain boundary by partial trypsin digestion. In titration experiments, the 140 kDa Msh6 was digested to a 112 kDa fragment under conditions where Msh2 was resistant to proteolysis (Fig. 1, upper arrow). Mass spectrometric analysis revealed that this Msh6 product was an N-terminal truncation; the most N-terminal peptide consistently identified by mass spectrometry included residues 252-265, whereas peptides including the C-terminal residue of Msh6 were recovered. No smaller proteolysis products were visible in the gel, indicating extensive digestion of the NTR. The next stable Msh6 proteolysis product (Fig. 1, lower arrow) was a N-terminal truncation of the first 441 residues, corresponding to loss of the mispair-binding domain. Thus, the Msh6 NTR is hypersensitive to cleavage by trypsin.

*N-terminal deletions of Msh6 define two mutation classes.*

The Msh6 NTR contains a canonical PIP box (27-QSSLLSFF-34). Mutation of Phe33 and 34 to Ala (*msh6 F33F34AA*) eliminates co-immunoprecipitation of PCNA but only causes a moderate MMR defect (Flores-Rozas et al., 2000). As residues surrounding the PIP box also contact PCNA in other proteins (Bruning and Shamoo, 2004; Chapados et al., 2004; Gulbis et al., 1996; Kontopidis et al., 2005; Matsumiya et al., 2002; Sakurai et al., 2005), we constructed a series of N-terminal
Figure 3-1. Partial proteolysis of Msh2-Msh6 and the Msh6 NTR.
Increasing amounts of trypsin were incubated with 1 µg purified Msh2-Msh6 or Msh6 NTR and analyzed by SDS-PAGE and silver staining or western blotting with antibodies against Msh2 or Msh6. Both Msh6 in the Msh2-Msh6 complex and the isolated Msh6 NTR were sensitive to levels of exposure to trypsin that did not digest Msh2.
deletions to investigate the contribution of residues surrounding the PIP box as well as the role of the rest of the NTR in vivo. Deletions were tested for complementation of the high mutation rates of an MMR-defective msh3Δ msh6Δ double mutant strain by expression from the native MSH6 promoter on a low copy-number plasmid.

Short N-terminal deletions (msh6Δ2-50 and msh6Δ2-123) caused modestly increased mutation rates equivalent to the PIP box mutation msh6 F33F34AA (Table 1; p > 0.1 for all pair-wise comparisons, Mann-Whitney test). In contrast, the msh6Δ2-294 allele had a null phenotype indistinguishable from the empty vector control (p > 0.2); GFP-tagging this mutant at its C-terminus to assess expression by western blotting revealed that it was not present at detectable levels (data not shown). Intermediate deletions (msh6Δ2-143, msh6Δ2-163, msh6Δ2-183, msh6Δ2-201, msh6Δ2-251, and msh6Δ2-271) caused mutation rates substantially higher than msh6 F33F34AA, msh6Δ2-50 and msh6Δ2-123 (p < 0.0001 for all comparisons between the two groups) but lower than the empty vector in at least one assay. Western blotting analysis of a GFP-tagged version of msh6Δ2-251 showed that it was present at higher levels than tagged wild-type MSH6, indicating that the defect in this mutant was not due to protein expression or stability problems (data not shown). Thus, mutants fell into two groups: the msh6 F33F34AA, msh6Δ2-50 and msh6Δ2-123 mutants had minor MMR defects, while deletions of 143 or more codons caused substantial MMR defects.

The increased mutation rates caused by deletions extending past codon 123 suggested that the region between residues 123 and 251 has a role in Msh6 function
Table 3-1. Mutation rates of Msh6 mutants.

<table>
<thead>
<tr>
<th>Plasmid allele&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Mutation rate (fold increase)&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thr&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>None</td>
<td>4.2 [3.6-5.3] x 10&lt;sup&gt;-6&lt;/sup&gt; (129)</td>
</tr>
<tr>
<td>MSH6</td>
<td>3.2 [2.4-4.2] x 10&lt;sup&gt;-8&lt;/sup&gt; (1)</td>
</tr>
<tr>
<td>msh6 F33F34AA</td>
<td>3.8 [2.0-8.6] x 10&lt;sup&gt;-7&lt;/sup&gt; (12)</td>
</tr>
<tr>
<td>msh6Δ2-50</td>
<td>4.7 [3.5-6.1] x 10&lt;sup&gt;-7&lt;/sup&gt; (14)</td>
</tr>
<tr>
<td>msh6Δ2-123</td>
<td>4.0 [2.4-7.2] x 10&lt;sup&gt;-7&lt;/sup&gt; (12)</td>
</tr>
<tr>
<td>msh6Δ2-143</td>
<td>2.4 [1.5-3.1] x 10&lt;sup&gt;-5&lt;/sup&gt; (74)</td>
</tr>
<tr>
<td>msh6Δ2-163</td>
<td>4.7 [3.9-5.4] x 10&lt;sup&gt;-6&lt;/sup&gt; (147)</td>
</tr>
<tr>
<td>msh6Δ2-183</td>
<td>2.3 [1.5-2.8] x 10&lt;sup&gt;-6&lt;/sup&gt; (70)</td>
</tr>
<tr>
<td>msh6Δ2-201</td>
<td>1.7 [1.3-2.5] x 10&lt;sup&gt;-6&lt;/sup&gt; (54)</td>
</tr>
<tr>
<td>msh6Δ2-251</td>
<td>4.4 [3.3-5.8] x 10&lt;sup&gt;-6&lt;/sup&gt; (136)</td>
</tr>
<tr>
<td>msh6Δ2-271</td>
<td>4.6 [3.9-6.4] x 10&lt;sup&gt;-6&lt;/sup&gt; (142)</td>
</tr>
<tr>
<td>msh6Δ2-294</td>
<td>4.2 [3.5-6.4] x 10&lt;sup&gt;-6&lt;/sup&gt; (129)</td>
</tr>
<tr>
<td>msh6Δ51-251</td>
<td>2.6 [2.0-3.7] x 10&lt;sup&gt;-7&lt;/sup&gt; (8)</td>
</tr>
<tr>
<td>msh6Δ51-251 F33F34AA</td>
<td>1.2 [0.74-1.5] x 10&lt;sup&gt;-6&lt;/sup&gt; (37)</td>
</tr>
<tr>
<td>msh6Δ51-271</td>
<td>6.0 [3.4-10.5] x 10&lt;sup&gt;-6&lt;/sup&gt; (185)</td>
</tr>
<tr>
<td>msh6Δ2-271</td>
<td>1.3 [0.9-3.2] x 10&lt;sup&gt;-6&lt;/sup&gt; (4)</td>
</tr>
<tr>
<td>msh6Δ2-284</td>
<td>4.3 [1.9-32.9] x 10&lt;sup&gt;-9&lt;/sup&gt; (132)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Median rates of hom3-10 reversion, lys2-10A reversion and CAN1 inactivation with 95% C.I. in square brackets and fold increase over wild-type MSH6 in parentheses.

<sup>b</sup>Gene present on low copy-number plasmid in msh3 msh6 strain RDKY 4234.
(Table 1). The internal deletion allele \textit{msh6}\textsubscript{Δ}51-251, in which PIP box is retained, complemented the \textit{msh3}\textsubscript{Δ} \textit{msh6}\textsubscript{Δ} strain similarly to wild-type \textit{MSH6}. Thus, the 51-251 region is only important in the absence of residues 2-50. Interestingly, combining the internal deletion with the PIP box mutations resulted in an allele (\textit{msh6}\textsubscript{Δ}51-251 \textit{F33F34AA}) that had a more serious defect than either \textit{msh6}\textsubscript{Δ}51-251 or \textit{msh6} \textit{F33F34AA} alone, but still complemented better than the N-terminal deletions that extend past residue 123 (p < 0.006 for all comparisons). Given that there are residues in addition to F33 and F34 that participate in PCNA-PIP box interactions (Gulbis et al., 1996), we suspect that the functionality of this mutant may be due to residual interaction with PCNA (see below).

In contrast to the mostly functional \textit{msh6}\textsubscript{Δ}51-251 allele, deletion of codons 51-271 (\textit{msh6}\textsubscript{Δ}51-271) caused a dramatic increase in mutation rate. To assess whether there was an important region between residues 251 and 271, we made an \textit{msh6}\textsubscript{Δ}252-271 mutant, which was similar to wild-type \textit{MSH6} (Table 1). Together with results described below, this suggests that the defect caused by \textit{msh6}\textsubscript{Δ}51-271 is due to the NTR being too short for productive interaction with PCNA. In contrast to the \textit{msh6}\textsubscript{Δ}252-271 allele, deletion of codons 252-284, 252-294, or 252-304 caused severe defects (Table 1 and Sup. Fig 2); taken together with the NTR-swap experiment described below, this suggests that the boundary between the NTR and the mispair-binding domain probably lies between residues 272 and 284.
Deletion of the N-terminal 123 residues of Msh6 completely abrogates interactions with PCNA.

To assess the possibility that the mildness of the defects caused by msh6 F33F34AA, msh6Δ2-50 and msh6Δ2-123 was due to residual interaction of the encoded proteins with PCNA, we tested the binding of purified Msh2-Msh6 F33F34AA and Msh2-Msh6Δ2-123 to immobilized PCNA by surface plasmon resonance. Wild-type Msh2-Msh6 bound PCNA with a $K_D$ of 2.5 nM (Sup. Fig. 3), consistent with published reports (Lee and Alani, 2006). In contrast, Msh2-Msh6 F33F34AA had greatly reduced binding, and Msh2-Msh6Δ2-123 lacked measurable interaction. Thus, complementation by the alleles encoding these proteins is not due to significant interactions with PCNA, although the low level of interaction seen with the F33F34AA mutant may allow partial function in the context of the msh6Δ51-251 F33F34AA mutant described above.

The Msh6 NTR is not required for nuclear localization, DNA binding, or mispair discrimination.

To ensure that MMR-defective NTR deletions do not disrupt Msh6 localization, we examined an msh3Δ msh6Δ strain expressing GFP-tagged wild-type Msh6 or Msh6Δ2-251 by fluorescence microscopy. The tagged wild-type Msh6 was functional for MMR in vivo and both proteins were localized to the nucleus, in contrast to GFP expressed alone from the Msh6 promoter, which was cytoplasmic (data not shown); the defect of Msh6Δ2-251 is therefore not due to loss of nuclear localization. Furthermore, purified wild-type Msh2-Msh6 and Msh2-Msh6Δ2-251
complexes bound similarly to DNA substrates containing a GT mispair or a +1 insertion, and both complexes discriminated mispairs from fully base-paired DNA (Sup. Fig. 4) and dissociated from DNA upon addition of ATP. The defects caused by deletion of residues 2-251 are therefore not due to failure to recognize DNA or discriminate between mispairs and homoduplex DNA.

*The Msh3 and Msh6 NTRs are interchangeable and the Msh6 NTR can function when placed on Msh2.*

Both Msh3 and Msh6 have NTRs with N-terminal PIP boxes. A chimeric Msh3 in which the first 154 residues were replaced with residues 1-304 of Msh6 complemented the msh3Δ msh6Δ strain nearly as well as wild-type MSH3 (Table 2). Similarly, a chimeric Msh6 in which the first 251 residues were replaced with residues 1-154 of Msh3 was functional, as was a version in which the first 271 residues were replaced with 1-154 of Msh3. However, shifting the fusion position in the Msh6 chimera to residue 285, 295 or 305 resulted in strongly MMR defective phenotypes (Table 2 and Sup. Fig. 2). Combined with our partial proteolysis and deletion results, the properties of these chimeric constructs suggest that Msh6 residues between 272 and 304 may stabilize Msh6 or be part of the mispair-binding domain. The functional chimeras demonstrate that the Msh3 and Msh6 NTRs perform similar functions, consistent with the NTRs lacking (or functionally replacing) interactions with the rest of the Msh2-Msh3 and Msh2-Msh6 heterodimers.
<table>
<thead>
<tr>
<th>Chromosomal alleles: $\text{MSH2 msh3 msh6}^a$</th>
<th>Mutation rate (fold increase)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid allele$^c$</td>
<td>Thr$^a$</td>
</tr>
<tr>
<td>None</td>
<td>4.2 [3.6-5.3] x 10$^6$ (129)</td>
</tr>
<tr>
<td>$\text{MSH6}$</td>
<td>3.2 [2.4-4.2] x 10$^4$ (1)</td>
</tr>
<tr>
<td>$\text{MSH3}$</td>
<td>1.8 [0.8-2.2] x 10$^4$ (1)</td>
</tr>
<tr>
<td>$\text{MSH6}(1-304)-\text{MSH3}(155-1047)$</td>
<td>6.3 [3.2-12] x 10$^4$ (2)</td>
</tr>
<tr>
<td>$\text{MSH3}(1-154)-\text{MSH6}(252-1242)$</td>
<td>1.9 [1.4-4.5] x 10$^7$ (6)</td>
</tr>
<tr>
<td>$\text{MSH3}(1-154)-\text{MSH6}(272-1242)$</td>
<td>2.9 [1.8-4.0] x 10$^4$ (1)</td>
</tr>
<tr>
<td>$\text{MSH3}(1-154)-\text{MSH6}(285-1242)$</td>
<td>3.4 [2.1-19.7] x 10$^6$ (105)</td>
</tr>
<tr>
<td>$\text{MSH3}(1-154)-\text{MSH6}(305-1242)$</td>
<td>2.7 [2.1-6.4] x 10$^6$ (83)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chromosomal alleles: $\text{msh2 msh3 msh6}^d$</th>
<th>Mutation rate (fold increase)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{MSH2 plasmid allele}^c$</td>
<td>$\text{MSH6 plasmid allele}^c$</td>
</tr>
<tr>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>$\text{MSH2}$</td>
<td>$\text{MSH6}$</td>
</tr>
<tr>
<td>$\text{MSH2}$</td>
<td>msh6Δ2-251</td>
</tr>
<tr>
<td>$\text{MSH6}(1-251)-\text{MSH2}(2-964)$</td>
<td>$\text{MSH6}$</td>
</tr>
<tr>
<td>$\text{MSH6}(1-251)-\text{MSH2}(2-964)$</td>
<td>msh6Δ2-251</td>
</tr>
</tbody>
</table>

$^a$RDKY 4234

$^b$Median rates for hom3-10 reversion, lys2-10A reversion and CAN1 inactivation with 95% C.I. in square brackets and fold increase over strain containing wild-type plasmid(s) in parentheses.

$^c$Gene present on low copy-number plasmid.

$^d$RDKY 5963
If the primary role of the Msh6 NTR is to localize Msh2-Msh6 within the nucleus, an Msh6 NTR-Msh2 fusion might restore function to a complex containing Msh6Δ2-251. A chimeric *MSH2* gene encoding residues 1-251 of Msh6 fused onto its N-terminus caused a slight increase in mutation rate relative to wild-type *MSH2* when expressed from a plasmid along with wild-type *MSH6* in an *msh2Δ msh3Δ msh6Δ* strain, but significantly decreased the mutation rate of a strain expressing Msh6Δ2-251 (Table 2). Thus the Msh6 NTR is functional when fused to either Msh2 or Msh6.

*The Msh6 NTR is an unstructured peptide that can interact with PCNA trimers at a 3:1 ratio.*

We cloned and expressed the first 304 residues of Msh6 for biochemical and biophysical characterization. A similar Msh3 construct (residues 1-154) expressed poorly and was not studied further. The isolated Msh6 NTR (34 kDa) showed similar trypsin sensitivity to the NTR in the Msh2-Msh6 heterodimer (Fig. 1). It also eluted in size-exclusion chromatography earlier (Stokes radius 45.7 Å) than predicted for a compact-globular protein (Stokes radius 26.7 Å; Fig. 2A), consistent with oligomerization or an extended conformation. The NTR bound PCNA and formed complexes that eluted at one of three retention volumes prior to PCNA or NTR alone (Fig. 2B), with the earlier eluting species forming at the higher molar ratios of NTR to PCNA. Thus, each PCNA trimer bound one, two, or three Msh6 NTR molecules.

Statistical analysis of the amino acid composition of the NTR showed hydrophilic Ser, Thr, and Asn residues were overrepresented and hydrophobic Val, Ile,
Figure 3-2. Size exclusion chromatography of Msh6 NTR, PCNA, and their complex.

A. The Msh6 NTR eluted from a Superose 6 column more rapidly than PCNA trimers, indicating that the NTR either oligomerizes or has an extended structure. B. Various molar ratios of PCNA monomers to NTR monomers resulted in the formation of complexes that migrated at three separate positions. A PCNA to NTR ratio of 3:1 or 3:2 gave a peak at the first position with some free NTR. Ratios of 3:2 and 3:3 gave a peak at the second position with some components at the first position as well, and ratios of 3:4.5, 3:6, and 3:9 gave a third peak with increasing amounts of unbound NTR. C. Heating PCNA at 100 C for 10 min caused the protein to form a high molecular weight aggregate. D. Similar heat treatment of the NTR did not change its elution profile.
and Leu residues were underrepresented (in the top and bottom 95% quantiles of all *S. cerevisiae* proteins, respectively; (Brendel et al., 1992)). This amino acid bias is characteristic of intrinsic disorder (Romero et al., 2001; Tompa, 2002), consistent with the protease accessibility and aberrant size exclusion chromatography. Additionally, boiling PCNA for ten minutes resulted in formation of a large aggregate (Fig. 2C), but the elution profile of NTR was unchanged by boiling (Fig. 2D).

SAXS can directly monitor protein folding and solution conformations, so we collected x-ray scattering data for the NTR, PCNA, and the PCNA-NTR complex (Fig. 3A). Consistent with gel filtration results, the radius of gyration ($R_g$) for the NTR was much larger than that observed for similarly sized globular proteins and the NTR-PCNA complex was larger than PCNA or the NTR alone (Fig. 3B).

The Kratky plot measures the overall compactness of proteins. Folded proteins follow Porod’s law ($I(S)$ is proportional to $S^{-4}$, where $S$ is reciprocal resolution) and have a bell-shaped curve as seen for both PCNA and the complex (Fig. 3C). In contrast, the NTR curve was flat, similar to proteins denatured by chemicals or heat (Doniach, 2001). Thus, the NTR does not fold into a compact domain.

SAXS can provide structural information for unfolded polymers by fitting to equation 1 (see methods; Fig 3D). The polypeptide length of 1078 Å matched the expected length of 1091 Å (304 residues x 3.78 Å/residue x 0.95, where 0.95 is a factor due to non-linearity of an extended polypeptide chain (Cantor, 1980)). The fitted parameter $b=18.7$ Å agrees well with the observed range of 19-25 Å (Perez et al., 2001), corresponding to a persistence length, the length for which the polypeptide
**Figure 3-3. SAXS of Msh6 NTR, PCNA, and their complex.**

A. Scattering curves for 10 mg/mL Msh6 NTR, PCNA, and the PCNA-MSH6 NTR complex. The NTR and complex curves were scaled by 100 and 0.001 to eliminate overlap. The solid line through PCNA represents the fitting of the S. cerevisiae PCNA trimer (PDB id 1PLQ; (Krishna et al., 1994)) to the experimental data with the program CRY SOL (Svergun, 1995), and the dashed line represents the fit corresponding to the ab initio PCNA model in panel F. B. Rg of PCNA, NTR, and the PCNA-NTR complex were calculated at various concentrations with average values of 34.1±0.1 Å, 56±2 Å, and 70.2±0.4 Å respectively shown by horizontal lines. The linear relationship of concentration to I(0) and consistency of Rg indicate the samples did not aggregate or dissociate during the experiment. C. Kratky plots of the samples show that, unlike the PCNA-containing samples, the NTR shows no globular fold. D. Plot illustrating the fits of the functional form for an unfolded polymer (equation 1, methods) to the experimental scattering curves for 2.5, 5, and 10 mg/mL NTR. The global parameters used for all of the curves are L=1078 Å, b=18.7 Å, Rc=6.99 Å. I(0) was fit independently for each curve and was found to be 461, 242, and 106 intensity units for the 10, 5, and 2.5 mg/mL samples, respectively. E. Pair-distribution functions for PCNA, NTR, and the PCNA-NTR complex. F. PCNA and NTR scattering curves were fitted with GASBOR (Svergun et al., 2001). Representative solutions from multiple runs are displayed as spheres at the positions of the "dummy residues" from the fitting. The reconstruction of PCNA (blue) was robust, whereas the shape of the extended NTR polypeptide (red) varied dramatically from run to run.
tends to act as a rigid unit, of 9.35 Å or roughly 3 amino acids. The $R_g$ of 57.2 Å calculated from these parameters (equation 2, see methods) is quite close to the observed $R_g$ of 56±2 Å calculated by the inverse Fourier transform method (Fig.3B).

**PCNA does not induce substantial structure in the Msh6 NTR.**

X-ray scattering of the PCNA-NTR complex suggests that PCNA binding does not induce substantial structure, as can be inferred by comparing the pair-distribution functions ($P(r)$’s) calculated from the scattering curves for all of the samples (Fig. 3E). The $P(r)$ function is a histogram of all distances within the scattering particle and can be calculated from atomic models. The overall shape of the PCNA $P(r)$ function was similar to those observed for well-folded proteins, and the double peak was due to the presence of a hole in the center of the PCNA ring, which could be reconstructed from the SAXS curve by *ab initio* calculations (Fig. 3F). The “tail” of the $P(r)$ function for the PCNA-NTR complex ($r>100$ Å) was consistent with the presence of a long unstructured addition to PCNA. The maximum intra-particle distance in the $P(r)$ function (D$_{\text{max}}$, where the $P(r)$ function returns to zero) of the PCNA-NTR was 250 Å. The free NTR and free PCNA, on the other hand, had maxima of only 185 and 89 Å, respectively. The difference between the D$_{\text{max}}$ of the complex and the D$_{\text{max}}$ of the NTR, 70 Å, was almost the same as the D$_{\text{max}}$ of PCNA alone. In fact, the $P(r)$ function of the complex closely resembled that of the NTR once the contributions from PCNA, including the double peak with the maximum at 55 Å and the shift of the scattering from the NTR tail by 70 Å, were excluded (Fig. 3E). These data indicate that binding
PCNA did not dramatically change the overall unfolded nature of most of the length of the NTR. These SAXS results were supported by partial proteolysis experiments in which trypsin was added to NTR-PCNA complexes or NTR mixed with equivalent amounts of BSA (Fig. 4A). The NTR was cleaved by a concentration of trypsin too low to have any visible effect on PCNA, consistent with the results seen for the NTR of full-length Msh6 described above (Fig. 1) and with the observed lack of secondary structure. Furthermore, PCNA did not affect the cleavage of the NTR, consistent with the NTR binding PCNA primarily via the PIP box.

*Solution structure of the Msh2-Msh6-PCNA complex*

Since the Msh6 NTR lacks structure alone, when bound to PCNA, or in the context of full-length Msh2-Msh6, we tested the protease sensitivity of the NTR in the Msh2-Msh6-PCNA complex (Fig. 4B). PCNA conferred only a very slight degree of protection to the NTR relative to BSA controls, and did not alter the proteolytic products produced. The lack of protection is consistent with the NTR remaining unstructured, proteolytically accessible, and not having significant regions of interaction with PCNA outside of the PIP box.

An unstructured NTR could be consistent with Msh2-Msh6-PCNA complexes where folded domains of Msh2-Msh6 and PCNA make direct contacts and the NTR is looped out. To assess this possibility, we collected SAXS data on full-length Msh2-
Figure 3-4. Partial proteolysis of the NTR and full length Msh2-Msh6 complexed to PCNA and solution structure of the Msh2-Msh6 complex with PCNA.

A. Tryptic digestion of Msh6 NTR was essentially unaffected by the addition of PCNA to the reaction at a 1:1 monomer ratio when compared to a BSA control as observed in coomassie blue stained gels and western blots using antibodies against Msh6. 

B. Similarly, proteolysis of Msh6 in the Msh2-Msh6 heterodimer showed no substantial change when PCNA was added at a 1:1 monomer ratio as compared to the BSA control. 

C. Pair distribution functions for PCNA, Msh2-Msh6, and the ternary complex. 

D. Pair distribution functions for Msh2-Msh6Δ51-251 and Msh2-Msh6Δ51-251 F33F34AA with and without PCNA. 

E. P(r) functions calculated for randomly generated models of MutS dimers linked to PCNA via random peptides reveals that no single conformer can account for the observed P(r) curve of the Msh2-Msh6-PCNA complex.
Msh6 alone and in complex with PCNA (Fig. 4C). The Msh2-Msh6-PCNA complex was detected by size-exclusion chromatography (data not shown); however, the P(r) function of Msh2-Msh6-PCNA complex differed only subtly from Msh2-Msh6 alone, with the long extended tail (r>100 Å) in the P(r) function of the complex having a stronger signal. These results are inconsistent with a substantial population of complexes having close contacts between the folded domains of Msh2-Msh6 and PCNA. However, these results could be explained if the domains are connected by with a flexible, ~1000 Å long NTR, which would average out the positions of the intermolecular vectors between the folded domains of PCNA and Msh2-Msh6.

To confirm this hypothesis, we used the deletion mutant Msh6Δ51-251 to shorten the tether to PCNA and decrease the NTR flexibility, thereby decreasing the conformations that average out the inter-molecular components in the P(r) function. We collected SAXS data on Msh2-Msh6Δ51-251 alone and in complex with PCNA (Fig. 4D). As predicted, the shorter tethering of the Msh2-Msh6Δ51-251-PCNA complex resulted in larger differences in the tail of the P(r) function when comparing the complex to Msh2-Msh6Δ51-251 alone. Moreover, SAXS data collected on the Msh2-Msh6Δ51-251 F33F34AA complex with and without PCNA demonstrated that these changes were lost when the Msh6-PCNA interaction was eliminated (Fig. 4D). Calculation of the P(r) curves using atomic models of the complex confirmed that no single conformation could account for the observed scattering (Fig. 4E), while the average of the scattering from many complexes might approximate the observed data. Thus in the absence of DNA and MutL homologs, the NTR is a flexible molecular leash between Msh2-Msh6 and PCNA.
**Figure 3-5. Disorder predictions for PIP-box containing proteins.**

**A.** IUPRED long-range disorder predictions for *S. cerevisiae* Msh2, which lacks an extended NTR region, Msh6, and Msh3. Black boxes indicate the positions of the PIP-box consensus. Note the disorder tendency parameter is substantially lower in the regions with homology to Msh2 and MutS.

**B.** Averaged disorder tendency for consensus PIP-box residues and 10 residues on either side for all *S. cerevisiae* proteins containing a PIP-box consensus. Individual proteins are plotted as points in groups based on whether they bind PCNA, do not bind PCNA, or are not known to bind PCNA. Horizontal lines represent the median disorder tendency.
Most functional PIP boxes in S. cerevisiae proteins are located in disordered regions.

Our results show the PIP box-containing NTR of Msh6 is unstructured. Long-range disorder predictions by IUPRED (Dosztanyi et al., 2005) for yeast (Fig. 5A) and human (Sup. Fig 5) Msh2, Msh3, and Msh6 are consistent with our experimental results and indicate that Msh6 and Msh3 have unstructured NTRs whereas Msh2 does not. To determine how representative these results are of other functional PIP boxes, we investigated the predicted flexibility of PIP box sequences (Fig. 5B). We calculated the IUPRED long-range disorder parameters for all S. cerevisiae proteins with the consensus PIP box sequence Qxx[LIV]xx[FY][FY] as well as known PCNA interacting proteins with non-canonical PIP boxes such as Rfc1. A PIP box disorder score was assigned as the average disorder tendency of the PIP box and ten residues surrounding it on both sides (if present). The scores were divided into three groups: proteins that bind PCNA, proteins that do not bind PCNA, and proteins with unknown binding status (Burgers, 1991; Clark et al., 2000; Flores-Rozas et al., 2000; Gary et al., 1997; Haracska et al., 2001; Li et al., 1995; Otterlei et al., 1999; Schmidt et al., 2002; Subramanian et al., 2005; Unk et al., 2002; Zhang et al., 2000). The median disorder tendency for the 13 known PCNA binding proteins was high at 0.4886. This IUPRED prediction suggests that the PIP boxes of many PCNA binding proteins lie in flexible tails, sometimes of substantial length, and are strongly reminiscent of Msh6 and Msh3 (Sup. Fig 6). These disorder tendencies were different (p<0.001, Mann-Whitney test) from those of PIP box regions of proteins known not to bind to PCNA (median 0.0827) and those with unknown binding status (median 0.0836), which were indistinguishable from each other (p>0.75).
3.4 DISCUSSION

Here we have shown that the N-terminal region (NTR) of Msh6 comprises an unstructured peptide of approximately 270 residues that behaves as a flexible tether in the context of the Msh2-Msh6-PCNA complex. *In vivo*, the Msh6 NTR could be functionally exchanged with a similar N-terminal region of Msh3 or fused to Msh2 to restore MMR proficiency to an *msh6*Δ2-251 strain. N-terminal truncations of up to 123 residues caused only modest MMR defects and were indistinguishable from point mutations in the PIP box. However, N-terminal deletions of 143 or more residues caused more substantial MMR defects that approached the null phenotype in some assays. The Msh6Δ2-251 protein was defective for MMR in spite of the fact that it was expressed stably, entered the nucleus, showed normal mispair binding and discrimination, and dissociated from mispairs when challenged with ATP. A deletion of residues 51-251 was only slightly defective, indicating that the region between residues 123 and 251 was only important for MMR in the absence of normal PCNA interactions. These data strongly suggest that the region between residues 123 and 251 of the NTR plays a second role in MMR in addition to contributing to the length of the PCNA tether, and that this second role is functionally partially redundant with PCNA interaction.

Analysis of the deletion and chimeric fusion proteins defined the probable junction between the NTR and the mispair-binding domain (MBD) of Msh6 and indicated a minimal length requirement for the NTR. The NTR/MBD junction likely lies between residues 272 and 284, as the internal deletion *msh6*Δ252-271 and the
chimeric allele $msh3(1-154)-msh6(272-1242)$ were largely functional for MMR, while the $msh6\Delta 252-284$ internal deletion and chimeric Msh3-NTR fusion to residue 285 of Msh6 resulted in strong MMR defects. Analysis of fusions that do not encroach upon the MBD allowed us to define the minimal distance between the PIP box and the NTR/MBD junction required for efficient MMR function. The $msh6\Delta 51-251$ allele encodes the most truncated protein that retained most of its function, indicating that a length of approximately 37-49 residues between the PIP box and the domain boundary is sufficient for productive interaction with PCNA; however, the $msh6\Delta 51-271$ allele was severely defective, indicating that a distance of approximately 17-29 residues is not long enough for efficient interaction. The minimum length required for NTR function therefore seems to be between approximately 80 and 150 Å when fully extended.

Our findings indicate that the NTR of Msh6 is intrinsically unstructured (or natively unfolded) and does not fold into a single conformation for its function (Dyson and Wright, 2005). The genetically defined minimal length and apparent lack of sequence requirements, as well as the functionality of the NTR when located on Msh2 instead of Msh6, is consistent with the NTR functioning as a tether. One possibility is that this tether recruits Msh2-Msh6 and Msh2-Msh3 to the general region of newly replicated DNA by interaction with PCNA. Remarkably, the 80-150 Å minimum tether length, which is ~1000 Å in wild-type Msh6 and even larger in human Msh6, may allow Msh2-Msh6 complexes to be recruited to PCNA trimers and then reach around complexes involved in replication or Okazaki fragment processing in order to bind mispairs.
The intriguing finding that there is a second region of the NTR that is functionally redundant with the PCNA-interacting site suggests that the unstructured sequence between residues 123 and 251 may mediate additional interactions with other components of the replication machinery, thereby providing an alternative mechanism for localizing MMR factors to sites of active replication. This redundancy would be consistent with the finding that the PCNA-Msh6 interaction is not strictly required for most Msh6-mediated MMR (Table 1; (Clark et al., 2000; Flores-Rozas et al., 2000; Kleczkowska et al., 2001)); however, the possibility for redundancy of the two NTR regions in MMR steps after mispair recognition cannot be ruled out by the experiments presented here. Future studies will be required to identify the putative interacting protein(s) and conclusively determine the functional significance of the interaction.

Although not all S. cerevisiae PIP boxes are likely linked to their cognate proteins through tethers as long as those found in Msh6, the use of a flexible linker to associate with PCNA appears to be quite common. This finding is consistent with the recently reported solution structure of the Sulfolobus solfataricus DNA ligase-PCNA complex, in which the interaction occurs via a PIP box located in a flexible loop (Pascal et al., 2006). This type of interaction not only allows the PIP box access to its binding site on PCNA, but an adjacent linker may be required to allow for simultaneous interactions with PCNA and nearby DNA substrates. Furthermore, flexible linkers could allow multiple proteins to interact simultaneously with a single PCNA trimer to coordinate various aspects of replication and repair. This type of interaction fits well with the established role of PCNA as a sliding clamp that interacts
with many different factors involved in DNA transactions, and suggests that the function of PCNA interaction with some of these other DNA-processing proteins may be also be related to localization to specific regions within the nucleus.
3.5 EXPERIMENTAL PROCEDURES

Strains and plasmids.

The msh2Δ msh3Δ msh6Δ S. cerevisiae strain RDKY5963 was made by disrupting MSH2 with HIS3 in RDKY4234 (MATa ura3-52 leu2Δ1 trp1Δ63 his3Δ200 hom3-10 lys2-10A msh3::hisG msh6::hisG). Strains were propagated in YPD or complete synthetic medium (CSM) lacking specified amino acids to select for plasmid markers. Deletion and domain swap alleles were made from MSH2-, MSH3-, or MSH6-bearing plasmids by recombination of linear fragments in yeast or by PCR with primers diverging from the region to be deleted and containing homology tails to allow reassembly in E. coli (Sup. Table 1). Point mutations were made by site-directed mutagenesis.

Genetic assays.

MSH2, MSH3 and MSH6 alleles were tested for complementation of RDKY4234 or RDKY5963 by fluctuation analysis (Lea, 1948) to determine rates of hom3-10 and lys2-10A reversion (Marsischky et al., 1996; Tran et al., 1997) and CAN1 inactivation. For each strain, at least 14 independent 10 mL cultures were grown in CSM –leu or CSM –leu –ura for 24-30 hours, harvested by centrifugation, washed, resuspended in water, serially diluted, and plated on CSM missing leu or leu and ura for plasmid selection and thr, lys or arg +60 µg/ml canavanine (Sigma, St. Louis, MO). P-values were calculated by two-tailed Mann-Whitney test (http://faculty.vassar.edu/lowry/utest.html).
Protein overexpression and purification.

The his-tagged NTR expression vector pRDK1136 was constructed by inserting codons 1-304 of S. cerevisiae MSH6 between the NcoI and XhoI sites of pET28b resulting in the addition of Ala-Leu-Glu-6His after residue 304. Protein was expressed in the E. coli Rosetta(DE3) strain (Novagen, Madison, WI) and purified by sequential chromatography on Ni-NTA (Qiagen, Valencia, CA) and Source 30S (Amersham, Piscataway, NJ) columns (Supplement). Wild-type and mutant Msh2-Msh6 complexes and PCNA were purified using modifications of previously published protocols (Supplement).

Analytical gel filtration.

Analytical gel filtration was performed on a SMART machine (Pharmacia) using a Superose 6 column and buffer containing 25 mM Tris pH 8, 0.1 mM EDTA, 1 mM DTT and 300 mM NaCl. PCNA-NTR complexes were generated by mixing various molar ratios of the proteins followed by incubation on ice for one hr. Stokes radii were calculated with a linear regression of retention time versus Stokes radius using standards: bovine serum albumin (35.5 Å), aldolase (48.1 Å), ferritin (61.0 Å), and thyroglobulin (85.0 Å).

Partial proteolysis and mass spectrometry.

Partial proteolysis was carried out in 10 µL reactions containing 1 µg Msh6 NTR or Msh2-Msh6 in 25 mM Tris pH 8, 100 mM NaCl, 10% glycerol, and 5 mM
MgCl$_2$. When indicated, Msh6 NTR and Msh2-Msh6 were premixed with 2.5 µg and 1 µg, respectively, of PCNA or BSA (Sigma) and incubated 1 hr on ice before adding trypsin as indicated and shifting the reactions to room temperature for 1 hr before stopping them by adding PMSF to 10 mM. Western blotting was performed with rabbit polyclonal antibodies to *S. cerevisiae* Msh2 or Msh6 raised in this laboratory. For mass spectrometry analysis, individual bands were excised from silver-stained gels, cut into pieces and soaked in acetonitrile for 30 min before drying. Gel pieces were rehydrated in 50 mM NH$_4$HCO$_3$ containing 7 ng/µL trypsin, incubated overnight at 37 C and then washed 3 times for 30 min each with 2% acetic acid 50% acetonitrile. The wash liquid was dried and peptides resuspended in 2% acetic acid and purified by ZipTip (Millipore, Billerica, MA) before analysis by mass spectrometry (Zhou et al., 2004).

**SAXS data collection and processing.**

SAXS data were collected at beamline 12.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratories using a silver behenate calibration standard (diffraction maxima 0.107633 and 0.215266 Å$^{-1}$). Buffer and sample were collected alternately with short exposures bracketing a longer exposure, typically 10 and 100 sec in length. Short exposures were compared to ensure no radiation damage had occurred, and data were merged with the program PRIMUS. Msh2-Msh6 and Msh2-Msh6-PCNA complexes were dialyzed into the following buffer for all SAXS experiments: 50 mM Tris pH 8.0, 0.5 mM EDTA, 1% glycerol, 5 mM DTT, and 200 mM NaCl. Msh6 NTR and NTR-PCNA complexes were dialyzed into buffer
consisting of 25 mM Tris pH 8, 0.1 mM EDTA, 1 mM DTT, and 100 mM NaCl. For each sample, data were collected at several different protein concentrations and the scattering was fit by the indirect Fourier transform method (Moore, 1980) as implemented in GNOM (Svergun, 1988).

For unfolded polymers with a persistence length of $b/2$ and a length of $L$, the observed scattering in the range $S<3/b$ has the following functional form (Brulet, 1996):

\[
I(S)/I(0) = \left[ \frac{2}{y^2} (y-1+e^{-y}) + \frac{b}{L}(4/15+7/(15y)-(11/15+7/(15y)))e^{-y} \right] \exp\left(-\frac{S^2 R_c^2}{6}\right)
\]

[equation 1]

where $y = S^2 L b/6$, $I(0)$ = the scattering intensity at zero angle (forward transfer momentum), $S$ = the reciprocal resolution, and $R_c$ = radius of gyration of the cross section. Scattering curves from Msh6 NTR were fit to this equation in the $S$ range of $[0 \text{ Å}^{-1}, 0.115 \text{ Å}^{-1}]$ using gnuplot (Williams, 1990).

Using these parameters, the expected $R_g$ for the unfolded polypeptide can be calculated with the equation:

\[
(R_g)^2 = b^2 \left[ \frac{y}{6} - \frac{1}{4} - \frac{1}{4y} + \frac{1}{8y^2} \right], \quad \text{where } y=L/b
\]

[equation 2]

**Computational Msh2-Msh6-PCNA complex modeling.**

Models of bacterial MutS bound via a flexible tether to the PIP box binding site in PCNA were generated by first generating the Cα positions of 304 residue random polypeptide. The geometry of the Cα trace was constrained through random
sampling of pseudoangles and pseudotorsion angles (Oldfield and Hubbard, 1994) in 5 degree increments weighted by their frequency in Nh3D version 3.0, a non-homologous subset of the protein database (Thiruv et al., 2005). Cα-Cα distances of 3.80 Å in adjacent residues and pseudosteric Cα-Cα clashes of 5.0 Å used during the generation of the random polypeptide were derived from statistics calculated from the Nh3D database. The final residues of the p21 PIP box sequence in the PCNA-p21 structure (PDB id 1axc; (Gulbis et al., 1996)) were superimposed with the first residues of the polypeptide and the first residues of one of the MutS subunits (PDB id 1fw6; (Junop et al., 2001)) were superimposed with the last residues of the polypeptide. The superimposed PCNA and MutS structures were rotated about the overlapping Cα-Cα bond and accepted if conformations with less than 10 Cα-Cα contacts closer than 4.0 Å were obtained.
3.6 ACKNOWLEDGEMENTS

We thank Cecilia Harries for assistance with mass spectrometry, Greg Hura for assistance in SAXS data collection, Mark Berardini for performing initial expression and purification trials with the Msh6 and Msh3 NTRs, Dan Mazur for assistance in protein purification, Jorrit Enserink for assistance in microscopy, Kristina Schmidt for biotinylated PCNA, and Manju Hingorani for expression plasmids. This work was supported by NIH grants GM50006 and CA92584. S.S.S. was supported by the NIH/NIGMS-funded UCSD Genetics Training Program (T32 GM08666), and C.D.P. was supported as a Robert Black fellow of the Damon Runyon Cancer Research Foundation.

Chapter 3, in full, is a reprint of the material as it appears in Shell, S. S., Putnam, C. D., and Kolodner, R. D. (2007). The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA. Mol Cell 26, 565-578. The dissertation author was a primary investigator and author of this paper and performed all experiments and analysis with the exception of the SAXS experiments, which were performed in collaboration with her co-author and analyzed by her co-author, and the analyses displayed in Figure 3-5 and Supplemental Figures 3-1, 3-5, and 3-6, which were performed by her co-author.
3.7 SUPPLEMENTARY MATERIALS
### Supplemental table 3-1. Plasmids used for complementation and expression experiments.

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Allele present</th>
<th>Construction comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pII-2</td>
<td>MS2</td>
<td>In YCp50 (Reenan and Kolodner, 1992)</td>
</tr>
<tr>
<td>pRDK1088</td>
<td>MS3</td>
<td>BamHI-HindIII fragment in pRS315</td>
</tr>
<tr>
<td>pRDK439</td>
<td>MS6</td>
<td>BamHI-HindIII fragment in pRS315</td>
</tr>
<tr>
<td>pRDK827</td>
<td>msh6 F33F34AA</td>
<td>Derived from pRDK439 (Flores-Rozas et al., 2000)</td>
</tr>
<tr>
<td>pRDK1129</td>
<td>msh6A2-50</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1141</td>
<td>msh6A2-123</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1203</td>
<td>msh6A2-143</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1204</td>
<td>msh6A2-163</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1205</td>
<td>msh6A2-183</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1143</td>
<td>msh6A2-201</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1145</td>
<td>msh6A2-251</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1247</td>
<td>msh6A2-271</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1095</td>
<td>msh6A2-294</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1197</td>
<td>msh6A51-251</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1210</td>
<td>msh6A51-251 F33F34AA</td>
<td>Derived from pRDK1197 by QuikChange</td>
</tr>
<tr>
<td>pRDK1248</td>
<td>msh6A252-271</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1256</td>
<td>msh6A252-284</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1257</td>
<td>msh6A252-294</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1259</td>
<td>msh6A252-304</td>
<td>Derived from pRDK439</td>
</tr>
<tr>
<td>pRDK1120</td>
<td>MSH6(1-304)-MSH3(155-1047)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1181</td>
<td>MSH3(1-154)-MSH6(252-1242)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1198</td>
<td>MSH3(1-154)-MSH6(272-1242)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1260</td>
<td>MSH3(1-154)-MSH6(285-1242)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1261</td>
<td>MSH3(1-154)-MSH6(295-1242)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1122</td>
<td>MSH3(1-154)-MSH6(305-1242)</td>
<td>Derived from pRDK439 and pRDK1088</td>
</tr>
<tr>
<td>pRDK1217</td>
<td>GFP</td>
<td>GFP replaces coding sequence of pRDK439</td>
</tr>
<tr>
<td>pRDK1213</td>
<td>MSH6-GFP</td>
<td>C-terminal GFP inserted into pRDK439</td>
</tr>
<tr>
<td>pRDK1215</td>
<td>msh6A2-251-GFP</td>
<td>C-terminal GFP inserted into pRDK1145</td>
</tr>
<tr>
<td>pRDK1255</td>
<td>msh6A2-294-GFP</td>
<td>C-terminal GFP inserted into pRDK1145</td>
</tr>
<tr>
<td>pRDK1221</td>
<td>MSH6(1-251)-MSH2(2-964)</td>
<td>Derived from pII-2</td>
</tr>
</tbody>
</table>

Low copy-number yeast plasmids
### Supplemental table 3-1. Plasmids used for complementation and expression experiments, continued.

<table>
<thead>
<tr>
<th>Bacterial overexpression plasmids</th>
<th>Allele present</th>
<th>Construction comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1135</td>
<td><em>MSH3</em>(1-54)-6X<em>HIS</em></td>
<td>NcoI-XhoI fragment in pET28b</td>
</tr>
<tr>
<td>pRDK1136</td>
<td><em>MSH6</em>(1-304)-6X<em>HIS</em></td>
<td>NcoI-XhoI fragment in pET28b</td>
</tr>
<tr>
<td>pRDK1211</td>
<td>msh6 F33F34AA</td>
<td>Derived from pET11a-Msh6 by QuikChange (Antony and Hingorani, 2003)</td>
</tr>
<tr>
<td>pRDK1222</td>
<td>msh6Δ2-123</td>
<td>Derived from pET11a-Msh6 (Antony and Hingorani, 2003)</td>
</tr>
<tr>
<td>pRDK1224</td>
<td>msh6Δ2-251</td>
<td>Derived from pET11a-Msh6 (Antony and Hingorani, 2003)</td>
</tr>
<tr>
<td>pRDK1225</td>
<td>msh6Δ51-251</td>
<td>Derived from pET11a-Msh6 (Antony and Hingorani, 2003)</td>
</tr>
<tr>
<td>pRDK1212</td>
<td>msh6Δ51-251 F33F34AA</td>
<td>Derived from pRDK1225</td>
</tr>
</tbody>
</table>

*aAll genes retain native promoter; for chimeric alleles the promoter of the N-terminal gene segment is retained, with the exception of pRDK1221, which retains the *MSH2* promoter.  
*bRelevant markers are Amp' ARS-CEN LEU2 for pRS315-derived plasmids (Sikorski and Hieter, 1989), Amp' URA3 for YCp50, Amp' for pET11a-derived plasmids, and Kan' for pET28b-derived plasmids.*
Supplemental Figure 3-1. Alignment of yeast and human MutS homologs.

A. *E. coli* MutS was aligned with human and yeast Msh2, Msh3, and Msh6. Identical amino acids in the alignment are shaded gray. The consensus PIP box at the N-termini of the Msh6 and Msh3 homologs are boxed. The division between the NTRs of the Msh6 and Msh3 homologs and the beginning of the MutS homology are indicated with lines. Sequence alignment performed using SEQUOIA (Bruns et al., 1999).

B. Structure of bacterial MutS [PDB id 1fw6; (Junop et al., 2001)] displayed with the mismatch-recognizing monomer colored from blue (N-terminus) to orange (C-terminus). The N-terminus that is extended in the eukaryotic MutS homologs is indicated. Molecular image generated with PyMOL (DeLano, 2002).
<p>| MutS | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh1 | hMeh1 | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh3 | hMeh3 | yMeh8 | hMeh8 | yMeh9 | hMeh9 | yMeh10 | hMeh10 | yMeh11 | hMeh11 | yMeh12 | hMeh12 | yMeh13 | hMeh13 | yMeh14 | hMeh14 |</p>
<table>
<thead>
<tr>
<th>------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
<td>PKD</td>
</tr>
<tr>
<td>T</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>NTR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**MutS homology**

| MutS | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh1 | hMeh1 | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh3 | hMeh3 | yMeh8 | hMeh8 | yMeh9 | hMeh9 | yMeh10 | hMeh10 | yMeh11 | hMeh11 | yMeh12 | hMeh12 | yMeh13 | hMeh13 | yMeh14 | hMeh14 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 250  | 295   | 295   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   | 305   |

**junction**

<p>| MutS | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh1 | hMeh1 | yMeh2 | hMeh2 | yMeh6 | hMeh6 | yMeh3 | hMeh3 | yMeh8 | hMeh8 | yMeh9 | hMeh9 | yMeh10 | hMeh10 | yMeh11 | hMeh11 | yMeh12 | hMeh12 | yMeh13 | hMeh13 | yMeh14 | hMeh14 |</p>
<table>
<thead>
<tr>
<th>------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
<th>-------</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
<td>ATM</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>NTR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Supplemental Figure 3-1. Alignment of yeast and human MutS homologs, continued.

A. *E. coli* MutS was aligned with human and yeast Msh2, Msh3, and Msh6. Identical amino acids in the alignment are shaded gray. The consensus PIP box at the N-termini of the Msh6 and Msh3 homologs are boxed. The division between the NTRs of the Msh6 and Msh3 homologs and the beginning of the MutS homology are indicated with lines. Sequence alignment performed using SEQUOIA (Bruns et al., 1999). B. Structure of bacterial MutS [PDB id 1fw6; (Junop et al., 2001)] displayed with the mismatch-recognizing monomer colored from blue (N-terminus) to orange (C-terminus). The N-terminus that is extended in the eukaryotic MutS homologs is indicated. Molecular image generated with PyMOL (DeLano, 2002).
Supplemental Figure 3-1. Alignment of yeast and human MutS homologs, continued.

A. E. coli MutS was aligned with human and yeast Msh2, Msh3, and Msh6. Identical amino acids in the alignment are shaded gray. The consensus PIP box at the N-termini of the Msh6 and Msh3 homologs are boxed. The division between the NTRs of the Msh6 and Msh3 homologs and the beginning of the MutS homology are indicated with lines. Sequence alignment performed using SEQUOIA (Bruns et al., 1999).

B. Structure of bacterial MutS [PDB id 1fw6; (Junop et al., 2001)] displayed with the mismatch-recognizing monomer colored from blue (N-terminus) to orange (C-terminus). The N-terminus that is extended in the eukaryotic MutS homologs is indicated. Molecular image generated with PyMOL (DeLano, 2002).
Supplemental Figure 3-2. Complementation of the mutator phenotype of an msh3Δ msh6Δ strain by low-copy number plasmids bearing MSH6 alleles indicated.

A. An internal deletion of residues 51-251 complements well, while deletions of 51-271, 51-284, 51-294, and 51-304 have moderate to severe mutator phenotypes. B. An internal deletion of residues 252-271 is similar to wild-type MSH6, while deletions of 252-284, 252-294, and 252-304 have strong mutator phenotypes. C. Chimeras in which residues 1-154 of Msh3 are fused to different positions of Msh6. Fusion to residue 272 of Msh6 produces a chimera fully proficient for MMR, while fusion to residue 285, 295, or 305 produces chimeras that are increasingly defective.
Supplemental Figure 3-3. PCNA interaction with protein constructs corresponding to complementing MSH6 alleles.

Msh2-Msh6 (red curve), Msh2-Msh6 F33F34AA (green curve), and Msh2-Msh6Δ2-123 (blue curve) were flowed over a PCNA-coated surface at 100 nM and association and dissociation monitored by surface plasmon resonance. A horizontal bar indicates the time during which protein was being injected. The $K_D$ of the Msh2-Msh6 interaction with PCNA is 2.5 nM; affinity constants were not determined for the two mutant complexes due to the weakness of their interactions.
Supplemental Figure 3-4. Mispair binding by Msh2-Msh6 and Msh2-Msh6Δ2-251.

5 nM Msh2-Msh6 (A) or Msh2-Msh6Δ2-251 (B) was flowed over surfaces coated with 236 bp DNAs containing a central GT mispair (black curve), +1 insertion (red curve), or GC base pair (green curve), and association and dissociation monitored by surface plasmon resonance. A horizontal bar indicates the time during which protein was being injected, and addition of 250 µM ATP is indicated by a vertical arrow.
Supplemental Figure 3-5. Disorder predictions indicate that human and yeast NTRs are similar.
Disorder predictions of yeast and human Msh2, Msh3, and Msh6 calculated by IUPRED (Dosztanyi et al., 2005) are roughly aligned on the basis of the start position for the beginning of the MutS homology. PIP boxes (if present) are indicated by a dark box. In each case, the human homologs are quite similar to the yeast homologs; hMsh2 lacks an NTR, hMsh6 has an extensive NTR, and hMsh3 has an intermediate NTR. These graphs suggest that the results from the yeast studies presented here will be applicable to the human homologs as well.
Supplemental Figure 3-6. Disorder predictions for other PIP box-containing yeast proteins.

Long-range IUPRED disorder predictions (Dosztanyi et al., 2005) for Cac1, Rfc1, Apn2, Rrm3, Pol32, Ung1, Rad27, Cdc9, Rad2, Pol2, and Rad30, all of which have PIP box consensus sequences (dark boxes) and are known to interact with PCNA. Like in Msh6 and Msh3, most of these PIP boxes are located in regions of the proteins that are predicted to be disordered, and frequently in disordered N- and C-terminal tails.
Supplemental methods

Purification of Msh2-Msh6 and Msh6 mutant complexes.

Msh2 was overexpressed in *E. coli* BL21-CodonPlus(DE3)-RIPL (Stratagene) from PLANT-2-B-MSH2 (Antony and Hingorani, 2003) for all complexes. Wild-type Msh6 was overexpressed from pET11a-MSH6 (Antony and Hingorani, 2003) and mutants were overexpressed from derivations of this plasmid listed in the supplementary table. Lysis was performed as described (Antony and Hingorani, 2003). All purification steps were performed in buffer A3 (25 mM Tris pH 8, 1 mM EDTA, 10% glycerol, 5 mM DTT, 0.02% igepal) containing inhibitor mix I1 (1 mM PMSF, 1 mM benzamadine, 0.5 mg/L bestatin, and 1 mg/L each chymostatin, pepstatin A, aprotinin, and leupeptin). Cells were lysed in 500 mM NaCl, and supernatants were diluted to 100 mM NaCl and loaded on a 30 mL SP sepharose column. The column was washed with 100 mM NaCl and eluted with a 300 mL linear gradient from 100 mM to 800 mM NaCl. Msh2-Msh6-containing fractions were diluted to 200 mM NaCl and loaded onto a 10 ml ssDNA cellulose column, washed with 200 mM NaCl and step eluted at 600 mM NaCl. Eluted proteins of interest were diluted to 150 mM NaCl and loaded onto a 5 mL Hi-Trap heparin column (Pharmacia), washed with 150 mM NaCl and eluted with a 50 mL linear gradient from 150 to 900 mM NaCl. Msh2-Msh6-containing fractions were diluted to 100 mM NaCl, and loaded onto a 5 mL HP Q column (Pharmacia). The column was washed with 100 mM NaCl and eluted with a 50 mL linear gradient from 100 to 800 mM NaCl.
Purification of PCNA.

PCNA was overexpressed in *E. coli* BL21-CodonPlus(DE3)-RIPL (Stratagene) and purified as described (Ayyagari et al., 1995), with the following modifications. After lysis and ammonium sulfate precipitation, the pellet was resuspended in, and subsequently dialyzed against, buffer A4 (50 mM HEPES pH 7.5, 1mM DTT, 1 mM EDTA, 10% glycerol and inhibitor mix I1), which was used in all subsequent steps. The dialyzed sample was loaded onto a 5 mL HiTrap DEAE FF column (Pharmacia), washed with 20 mM NaCl, and eluted with a 100 mL linear gradient from 20 mM to 600 mM NaCl. Fractions of interest were then flowed through two 5 ml HiTrap SP HP columns (Pharmacia) at 100 mM NaCl and loaded onto a 5 mL HiTrap Q HP column, where it was eluted with a 50 mL gradient from 100 mM to 1M NaCl.

Surface plasmon resonance.

200 response units of biotin-labeled 236 bp DNA substrates containing a central GT mispair, a central +1 insertion, or no mispair were bound to an SA Biacore chip coated with immobilized avidin. Substrate construction has been described (Lin et al., 1998; Mendillo et al., 2005). 5 nM wild-type Msh2-Msh6 and Msh2-Msh6Δ2-251 were flowed over the chip for 2 min. at a rate of 30 µl/min. in running buffer consisting of 25 mM Tris pH 8, 2% glycerol, 4 mM MgCl₂, 110 mM NaCl, 0.01% igepal and 2 mM DTT. ATP-induced disassociation was measured by flowing 60 µl of running buffer containing 250 µM ATP over the chip. The DNA surface was regenerated with a 30 sec. pulse of 3 M NaCl. PCNA interaction was assessed by
binding 20 ng of biotinylated PCNA to an SA chip and flowing 100 nM Msh2-Msh6, Msh2-Msh6 F33F34AA, or Msh2-Msh6Δ2-123 over the surface at 30 µl/min. in running buffer. The PCNA surface was regenerated by a 20 sec. pulse of running buffer containing 0.05% SDS. For \( K_D \) determination a series of Msh2-Msh6 concentrations were used and the results analyzed in BIAevaluation version 3.1. Reference subtraction was made from an unmodified flow cell on the SA chip for all Biacore experiments.
3.8 REFERENCES


replication complex. Proc Natl Acad Sci U S A 93, 11597-11602.


CHAPTER 4

A chimeric *Saccharomyces cerevisiae* Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins.
4.1 ABSTRACT

Msh2-Msh3 and Msh2-Msh6 are two partially redundant mispair-recognition complexes that initiate mismatch repair (MMR) in eukaryotes. Crystal structures of the prokaryotic homolog MutS suggest the mechanism by which Msh6 interacts with mispairs, since key mispair-contacting residues are conserved in these two proteins. As Msh3 lacks these conserved residues, we constructed a series of mutants to investigate the requirements for mispair interaction by Msh3. We found that a chimeric protein in which the mispair-binding domain (MBD) of Msh6 was replaced by the equivalent domain of Msh3 was functional for MMR. This chimera possessed the mispair-binding specificity of Msh3 and revealed that communication between the MBD and the ATPase domain is conserved between Msh2-Msh3 and Msh2-Msh6. Further, the chimeric protein retained Msh6-like properties with respect to genetic interactions with the MutL homologues and an Msh2 MBD deletion mutant, indicating that Msh3-like behaviors beyond mispair specificity are not features controlled by the MBD.
4.2 INTRODUCTION

Mismatch repair (MMR) is an important mechanism for repair of base misincorporations and insertions or deletions of one or more nucleotides that occur during DNA replication. MMR has been conserved from bacteria to humans, and MMR defects cause increased mutation rates and contribute to cancer susceptibility and development (Fishel et al., 1993; Iyer et al., 2006; Kane et al., 1997; Marsischky et al., 1996). While many of the components of MMR have been identified, gaps remain in our understanding of how mispairs are recognized and how this recognition is converted into signals for repair.

Recognition of mispairs by MutS or its eukaryotic homologs is required for recruitment and activation of downstream factors and thereby initiates MMR. In bacteria, both base-base mispairs and insertion/deletions are recognized by the MutS homodimer (Iyer et al., 2006). Crystal structures of MutS in complex with mispaired DNA have shown that the homodimer interacts with DNA in an asymmetric manner, such that only one subunit directly contacts the mispair while the other contacts only the backbone of the DNA (Lamers et al., 2000; Natrajan et al., 2003; Obmolova et al., 2000). In eukaryotes, three MutS Homologs (MSH) form two heterodimers that participate in MMR. The Msh2-Msh6 complex recognizes both base-base and small insertion/deletion mispairs, while the Msh2-Msh3 complex primarily recognizes both small and large insertion/deletion mispairs (Acharya et al., 1996; Marsischky and Kolodner, 1999); the two complexes are thus partially redundant. Specificity in the eukaryotic heterodimers is mediated by the Msh6 and Msh3 subunits (Marsischky et
al., 1996), which correspond to the mispair-contacting subunit in the asymmetric MutS homodimer (Iyer et al., 2006).

An N-terminally located mispair-binding domain (MBD) mediates mispair recognition. Structures of MutS complexed with DNA containing G/T, G/A, A/A or G/G mispairs or an insertion of a single thymidine demonstrate a single mechanism for mispair recognition by the MBD (Lamers et al., 2000; Natraj an et al., 2003; Obmolova et al., 2000). In all cases a highly conserved phenylalanine residue (Phe36 in E. coli, Phe39 in T. aquaticus) stacks with a mispaired base (Figure 1B), consistent with the key role indicated for this residue by genetic studies (Drotschmann et al., 2001; Schofield et al., 2001). The stacking Phe, a glutamate residue two positions downstream that hydrogen-bonds with the mispaired bases in the crystal structures, and multiple residues that appear to contact the DNA backbone in the crystal structures are conserved between MutS and Msh6 sequences from a number of eukaryotes, suggesting a conserved mechanism of mispair interaction. Mutation of the stacking Phe creates a non-functional protein that is unable to bind mispairs (Bowers et al., 1999; Dufner et al., 2000; Lamers et al., 2000; Natraj an et al., 2003; Obmolova et al., 2000). In contrast, while several of the putative backbone-contacting residues are conserved in eukaryotic Msh3 sequences, all known Msh3 sequences lack the two conserved residues implicated in mispair recognition by MutS and Msh6. Moreover, individual mutations altering some of the conserved backbone-contacting residues cause more severe defects in Msh6 than in Msh3 (Drotschmann et al., 2001; Lee et al., 2007). Consequently, the interaction of Msh2-Msh3 with insertions and deletions likely has important differences from that of both Msh2-Msh6 and MutS (Lee et al.,
Remarkably, Msh3 is non-functional when complexed with a Msh2 deletion that lacks the domain homologous to the MutS MBD (Msh2ΔMBD), whereas the Msh2ΔMBD-Msh6 heterodimer is almost fully functional (Goldfarb and Alani, 2005; Lee et al., 2007).

Mispair recognition by MutS as well as the Msh2-Msh6 and Msh2-Msh3 complexes induces exchange of ADP for ATP at the distant ATPase domains, likely as a result of coupled conformational changes within the mispair recognition proteins (Iyer et al., 2006). ATP binding by MutS is required for interaction with the MutL homodimer, which induces downstream factors to cleave, unwind, and degrade the nascent DNA strand (Iyer et al., 2006). In eukaryotes, MutL is conserved in the form of multiple heterodimers containing common and unique subunits. The Mlh1-Pms1 complex in *S. cerevisiae* (Mlh1-Pms2 in higher eukaryotes) is required for the majority of MMR, as *pms1Δ* mutants have high mutation rates like those of *mlh1Δ* or *msh2Δ* mutants (Flores-Rozas and Kolodner, 1998). The Mlh1-Mlh3 complex, on the other hand, appears to be required for a subset of Msh3-dependent MMR, as evidenced by the finding that *mlh3Δ msh6Δ* double mutants have a considerably higher mutation rate than *mlh3Δ* or *msh6Δ* single mutants, while *mlh3Δ msh3Δ* double mutants have a mutation rate similar to *mlh3Δ* or *msh3Δ* single mutants (Flores-Rozas and Kolodner, 1998). It is not clear whether the ability of Mlh1-Mlh3 to participate in Msh3-dependent MMR depends on the types of mispairs being recognized or on other factors. Furthermore, it is not known which domains of MutS or its eukaryotic homologs interact with the MutL complexes or how mispair and ATP binding facilitate the interaction.
Since mispair recognition by the Msh2-Msh3 and Msh2-Msh6 complexes clearly differs but the complexes seem to have similar functions overall, we investigated whether point mutations or swapping portions of the Msh3 sequence homologous to the MutS MBD into Msh6 could produce functional chimeric alleles. We found that a version of Msh6 in which the MBD was replaced by the MBD of Msh3 was functional for MMR, and genetic and biochemical characterization demonstrated that the chimeric protein had the mispair-recognition specificities of Msh3. The finding that the chimera was fully functional for MMR indicated that transmission of mismatch recognition into signals for downstream repair events were not compromised, suggesting that Msh3 and Msh6 retain conserved mechanisms mediating communication between domains. Like Msh6, the chimeric protein did not appear to significantly interact with Mlh1-Mlh3, indicating that the regions of the MSHs dictating specificity toward the MutL homologs lie outside of the MBD, and that recruitment of MutL homologs is not strongly influenced by the types of lesions being recognized. Surprisingly, the chimeric allele was functional in combination with the Msh2 ΔMBD mutant, indicating that the MBD of Msh3 is in itself competent to recognize mispairs without the function of the MBD of Msh2.
4.3 RESULTS

*Generation of chimeric MSH6-MSH3 alleles that function* in vivo.

We sought to understand how Msh3 recognizes mispairs by investigating whether selective mutations in the MBD could convert the normal mispair recognition properties of Msh6 into those of Msh3 or *vice versa*. Alignments of Msh6 and Msh3 sequences from a variety of organisms show a number of residues in the MBD that are strongly conserved amongst the Msh6 sequences or amongst the Msh3 sequences, but are different between the Msh6 and Msh3 sequences (data not shown). These residues likely contribute to mispair specificity and include the Msh6-conserved mispair-contacting Phe residue as well as a number of residues nearby in the linear sequence. We therefore made *MSH3* mutants in which sequence encoding two, three, or seven conserved residues was changed to encode the equivalent Msh6 residues (Y186K and K187F; Y186K, K187F and K189E; Y186K, K187F, K189E, C190L, F191Y, A192E, and E193K) as well as the reciprocal Msh6 mutants to test the effect of exchanging these homolog-specific residues between Msh6 and Msh3 (Fig. 1). The engineered alleles were tested by patch test for their ability to complement the high mutation rate of an *msh3Δ msh6Δ* strain when present on low copy-number plasmids, and we none were able to complement (data not shown).

The failure of the point mutations to yield Msh3 and Msh6 derivatives that were functional for MMR *in vivo* might have been due to defects altering the stability, affinity or mispair discrimination of the resulting chimeric MBD-containing proteins.
Figure 4-1. Mispair-binding domain structure and sequence alignments.

A. Ribbon diagrams illustrating the overall fold and recognition of a DNA substrate with a G:T mispair (DNA backbone in black, bases in grey) by the *E. coli* MutS homodimer (PDB id 1w7a; Lamers et al., 2004). The Msh2-like MutS subunit that does not contact the mispair is yellow, while Msh6-like mispair-interacting subunit is depicted in orange with the MBD in red.

B. Interactions of the MBD with the DNA mispair. The mispaired thymidine (blue) stacks with Phe36. Small spheres and changes in ribbon coloring (green at the N-terminus and red at the C-terminus) indicate Cα positions of residues at junctions in the subdomain swap constructs. Panels A and B generated with PyMOL (DeLano, 2002).

C. Sequence alignment of *E. coli* and *T. aquaticus* MutS and human and yeast Msh6 and Msh3. Secondary structural elements derived from the *E. coli* MutS structure are shown above the alignment. Absolutily conserved positions are indicated with a star, Phe36 is boxed and the seven residues examined by site-directed mutagenesis are underlined. Vertical lines indicate the beginning and end of the region swapped in Msh6(3-MBD). Alignments generated by SEQUOIA (Bruns et al., 1999).
or possible failure of the chimeric constructs to properly signal mispair recognition to induce subsequent repair steps. To avoid these potential problems, we created an

*MSH6* allele termed *msh6(3-MBD)* in which Msh6 residues 305 through 421, which correspond by homology to the MBD of MutS, were replaced by the homologous Msh3 sequence, residues 155 through 285 (Fig 1C & 2). Patch tests showed that the plasmid-borne *msh6(3-MBD)* allele complemented the *msh3Δ msh6Δ* strain to a degree similar to that seen for wild-type *MSH3* and *MSH6* (Sup. Fig. 1). The exchanged sequence could be successfully extended into the non-conserved N-terminal region by replacing residues 1-421 of Msh6 with residues 1-285 of Msh3 in the *msh6(3-1-285)* allele (Fig. 2). However, minimizing the size of the replaced MBD region by creating several *MSH6* alleles containing portions of the Msh3 MBD did not yield chimeric genes capable of complementing the *msh3Δ msh6Δ* mutator phenotype (Fig. 2). Breakpoints for the sub-domain swaps were chosen based on analysis of the MutS crystal structure in order to exchange portions of the MBD directly involved in mispair recognition with junctions at conserved boundaries between secondary structural elements within the domain (Fig. 1). We conclude that either the mode of mispair binding found in Msh3 requires the entire Msh3 MBD, or the sub-domain swaps disrupted the overall structure or function of the chimeric MBDs.

In contrast to the successful swap of the MBD of Msh3 into Msh6, an Msh3 construct encoding the reverse swap (codons 155-285 of *MSH3* replaced with codons 305-421 of *MSH6*) failed to complement the *msh3Δ msh6Δ* strain (Fig. 2). Since the *msh6(3-1-285)* allele that swapped both the N-terminal region and the MBD was functional, we also tested *MSH3* alleles that encoded variants containing both the N-
Figure 4-2. Properties of point mutants and chimeric alleles.
Black bars indicate sequence derived from Msh3 and gray bars indicate Msh6 sequence. Numbering refers to amino acid positions of the wild-type proteins. “Complementation” indicates ability (+) or failure (-) of a plasmid bearing the indicated allele to decrease the high mutation rate of an *msh3Δ msh6Δ* strain in patch tests.
terminal region and MBD of Msh6 to eliminate the possibility of an incorrect chimeric junction between the MBD and the N-terminal region. Neither the msh3(6-1-421) allele nor others with different junctions between the MBD of Msh6 and the C-terminal portion of Msh3 [msh3(6-1-421), msh3(6-1-455), msh3(6-1-466), msh3(6-1-483), msh3(6-1-370), msh3(6-1-441a) and msh3(6-1-441b)] were able to complement the msh3Δ msh6Δ strain (Fig. 2). These results suggest either that the chimeras disrupted key aspects of the Msh3 protein, that the MBD of Msh6 fails to make appropriate contacts with other domains of Msh3 for signaling, or that subtle differences exist in the function of the Msh3 and Msh6 MBDs such that the Msh6 MBD cannot functionally replace the Msh3 MBD.

The mutations suppressed by the msh6(3-MBD) allele are similar to those suppressed by MSH3.

To extend these above results, the msh6(3-MBD) allele was placed at the MSH6 chromosomal locus of wild-type and msh3Δ strains. Fluctuation analysis was performed to determine the rates of reversion of the hom3-10 and lys2-10A alleles, which primarily result from single nucleotide deletions, and the rates of inactivation of the CAN1 gene, which reflect the accumulation of both frameshift and base substitution mutations (Marsischky et al., 1996; Tran et al., 1997).

The reversion rates of the msh6(3-MBD) msh3Δ strain were similar to those of the msh6Δ MSH3 strain in the hom3-10 and lys2-10A assays (Table 1), suggesting that the msh6(3-MBD) allele was similar to wild-type MSH3 in the repair of frameshift mispairs. Consistent with this, the Can’ mutation rate of the msh6(3-MBD)
Table 4-1. Mutation rate analysis of *msh6*(3-MBD).

<table>
<thead>
<tr>
<th>Relevant genotype</th>
<th>Strain</th>
<th>Mutation rate (fold increase)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Thr$^+$</td>
</tr>
<tr>
<td><strong>MSH3 MSH6</strong></td>
<td>3686</td>
<td>2.0 [1.2-3.7] x 10$^{-9}$ (1)</td>
</tr>
<tr>
<td><strong>MSH3 msh6</strong></td>
<td>4151</td>
<td>3.0 [1.9-3.9] x 10$^{-8}$ (15)</td>
</tr>
<tr>
<td><strong>msh3 MSH6</strong></td>
<td>4149</td>
<td>2.2 [1.6-3.5] x 10$^{-8}$ (11)</td>
</tr>
<tr>
<td><strong>msh3 msh6</strong></td>
<td>4234</td>
<td>3.1 [1.9-4.8] x 10$^{-6}$ (1548)</td>
</tr>
<tr>
<td><strong>MSH3 msh6(3-MBD)</strong></td>
<td>5153</td>
<td>2.7 [1.7-4.0] x 10$^{-9}$ (1)</td>
</tr>
<tr>
<td><strong>msh3 msh6(3-MBD)</strong></td>
<td>5249</td>
<td>6.6 [5.7-8.5] x 10$^{-4}$ (33)</td>
</tr>
</tbody>
</table>

*aMedian rate of 18-35 cultures per strain, with 95% confidence intervals in square brackets and fold increase in rate relative to the wild-type strain in parentheses.
msh3Δ strain was approximately 13-fold higher than wild-type, similar to both the msh6Δ MSH3 and the msh6(3-MBD) MSH3 strains (Table 1). These Can’ rates are consistent with the idea that the msh6(3-MBD) allele, like wild-type MSH3, can support the repair of frameshifts in the CAN1 gene, but not the base substitutions normally recognized by the wild-type Msh2-Msh6 complex. Thus, substitution of the Msh3 MBD for the Msh6 MBD in the msh6(3-MBD) allele likely changed the mispair-recognition specificity from that of MSH6 to that of MSH3. Intriguingly, the msh6(3-MBD) MSH3 strain had hom3-10 and lys2-10A reversion rates that were similar to those of the wild-type strain, indicating that a functional copy of MSH3 combined with the chimeric allele was sufficient for wild-type levels of repair of frameshift mispairs; this result suggests that the small defect seen in msh3Δ and msh6Δ single mutants in these assays may be due to decreased gene dosage rather than loss of non-redundant activities.

To confirm the results of the mutation rate analysis, we analyzed the spectra of mutations in the CAN1 gene in the same six strains. Base substitutions made up the majority of mutations found in the wild-type strain, and the rate at which base substitutions accumulated increased approximately 20-fold in all strains lacking wild-type MSH6 (Table 2). The msh6(3-MBD) allele, like MSH3, was unable to suppress these high base substitution rates, consistent with the absence of normal MSH6-like recognition of base-base mispairs (Table 2). In contrast, MSH3, MSH6, and msh6(3-MBD) all suppressed the accumulation of ±1 frameshift mutations to a similar extent. We did not observe sufficient numbers of larger insertion/deletion events to reliably
### Table 4-2. CAN1 mutation spectra.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Base substitutions</th>
<th>+/- 1 Frameshifts</th>
<th>Larger deletions or duplications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (^a) Rate (fold increase)(^b)</td>
<td>% (^a) Rate (fold increase)</td>
<td>% (^a) Rate (fold increase)</td>
</tr>
<tr>
<td>Wild-type</td>
<td>80 3.9 x 10(^{-8}) (1)</td>
<td>14 6.9 x 10(^{-7}) (1)</td>
<td>6 2.9 x 10(^{-8}) (1)</td>
</tr>
<tr>
<td>msh6(^\Delta)</td>
<td>93 6.2 x 10(^{-7}) (16)</td>
<td>5 3.4 x 10(^{-8}) (5)</td>
<td>0 N/A</td>
</tr>
<tr>
<td>msh3(^\Delta)</td>
<td>51 3.2 x 10(^{-8}) (1)</td>
<td>24 1.5 x 10(^{-8}) (2)</td>
<td>24 1.5 x 10(^{-8}) (5)</td>
</tr>
<tr>
<td>msh6(^\Delta) msh3(^\Delta)</td>
<td>33 6.9 x 10(^{-7}) (18)</td>
<td>65 1.4 x 10(^{-6}) (203)</td>
<td>2 4.2 x 10(^{-8}) (14)</td>
</tr>
<tr>
<td>msh6(3-MBD)</td>
<td>98 8.9 x 10(^{-7}) (24)</td>
<td>2 1.8 x 10(^{-8}) (3)</td>
<td>0 N/A</td>
</tr>
<tr>
<td>msh6(3-MBD) msh3(^\Delta)</td>
<td>92 8.4 x 10(^{-7}) (22)</td>
<td>4 3.6 x 10(^{-8}) (5)</td>
<td>4 3.6 x 10(^{-8}) (12)</td>
</tr>
</tbody>
</table>

\(^a\)Percent of total mutations that fall into the category indicated (n = 49-59 per strain).

\(^b\)Rate of CAN1 inactivation via mutation type indicated. Fold increase over wild-type rate in parentheses.
discern differences between MSH3, MSH6, and msh6(3-MBD) with respect to repair of these types of mispairs. Together, the mutation rate and spectrum data suggest that the mispair recognition properties of the MBD encoded by MSH3 are maintained when this domain is transferred into Msh6.

The Msh2-Msh6(3-MBD) chimera shows specificity for insertion/deletion loops in vitro.

To assess the ability of the Msh2-Msh6(3-MBD) chimera complex to recognize larger insertion/deletion mispairs, we purified the chimeric Msh6(3-MBD) protein as a heterodimer with Msh2 for biochemical analysis. Surface plasmon resonance was used to assess the interaction of Msh2-Msh6(3-MBD) and wild-type Msh2-Msh6 with 236 bp DNAs harboring centrally-located mispairs (Fig 3). The wild-type complex bound strongly to DNA substrates with a GT or +1 insertion mispair, but interacted more weakly with +3 and +4 insertions (Fig. 3A). In contrast, the Msh2-Msh6(3-MBD) chimera did not bind the GT mispair better than the fully base-paired DNA, and bound to the +1, +3, and +4 insertions more strongly than wild-type Msh2-Msh6, with the +3 insertion showing the strongest relative interaction (Fig. 3B). These findings indicate that the Msh6(3-MBD) chimera was compromised relative to Msh6 in its ability to recognize base-base mispairs, but had an increased ability to recognize the larger insertions that have been shown to be recognized by Msh3 (Iyer et al., 2006).
Figure 4-3. Surface plasmon resonance analysis of DNA binding.
10 nM wild-type Msh2-Msh6 or Msh2-Msh6(3-MBD) complexes were flowed over sensor chips pre-conjugated with 75-85 response units of 236 bp DNAs containing central GT, +1, +3, or +4 mispairs as indicated or no mispairs (GC).
The msh6(3-MBD) allele functions with an msh2 MBD deletion mutant.

Recent reports have indicated that Msh3 but not Msh6 requires the Msh2 domain homologous to the MutS MBD for in vivo MMR activity (Goldfarb and Alani, 2005; Lee et al., 2007). To determine whether this requirement was intrinsic to the Msh3 MBD or was due to other features of the Msh2-Msh3 heterodimer, we assessed complementation of an msh2Δ msh3Δ msh6Δ strain by plasmid-encoded MSH3, MSH6, and msh6(3-MBD) alleles in combination with plasmids encoding either wild-type Msh2 or Msh2 missing residues 2-133 (msh2ΔMBD), which correspond to the entire MBD-like domain (Goldfarb and Alani, 2005; Lee et al., 2007). Consistent with previous reports, coexpression of MSH3 and msh2ΔMBD did not complement the triple mutant, but coexpression of MSH6 and msh2ΔMBD complemented well (Table 3). Surprisingly, msh6(3-MBD) in combination with msh2ΔMBD complemented the triple mutant just as well as msh6(3-MBD) in combination with wild-type MSH2, suggesting that the MBD-like domain of Msh2 is not required for mispair recognition by the MBD of Msh3 per se, but rather that the MBD-like domain of Msh2 may be required for some other aspect of Msh3 function.

Unlike Msh3, Msh6(3-MBD) does not mediate significant repair through the Mlh1-Mlh3 complex.

The basis of the specificity of Mlh1-Mlh3 for Msh2-Msh3-bound mispairs is not known. In principle, both protein-protein interactions and protein-mispair interactions could influence which MutL-related complex is recruited. The Msh6(3-MBD) chimera contains the MBD of Msh3 placed within a scaffold of Msh6 and thus
Table 4-3. Complementation of msh2Δ msh3Δ msh6Δ strain by MSH3/6 alleles and msh2ΔMBD.

<table>
<thead>
<tr>
<th>Plasmid alleles⁶</th>
<th>Mutation rate (fold increase)⁷</th>
<th>Thr⁺ rate x 10⁷</th>
<th>Lys⁺ rate x 10⁸</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>55 [32-120] (103)</td>
<td>130 [86-220] (338)</td>
<td></td>
</tr>
<tr>
<td>MSH2 MSH6</td>
<td>0.54 [0.34-1.3] (1)</td>
<td>0.38 [0.17-0.68] (1)</td>
<td></td>
</tr>
<tr>
<td>MSH2 MSH3</td>
<td>0.29 [0.16-0.56] (1)</td>
<td>1.5 [0.92-2.3] (4)</td>
<td></td>
</tr>
<tr>
<td>MSH2 msh6(3-MBD)</td>
<td>1.7 [1.2-5.3] (3)</td>
<td>2.6 [1.8-4.2] (7)</td>
<td></td>
</tr>
<tr>
<td>msh2ΔMBD MSH6</td>
<td>1.4 [0.90-2.3] (3)</td>
<td>0.62 [0.35-0.97] (2)</td>
<td></td>
</tr>
<tr>
<td>msh2ΔMBD MSH3</td>
<td>95 [72-220] (176)</td>
<td>140 [99-190] (377)</td>
<td></td>
</tr>
<tr>
<td>msh2ΔMBD msh6(3-MBD)</td>
<td>1.7 [1.2-4.8] (3)</td>
<td>2.3 [1.7-2.9] (6)</td>
<td></td>
</tr>
</tbody>
</table>

⁶Median rate of 14-20 cultures. 95% confidence intervals in square brackets, fold increase over MSH2 MSH6-complemented strain in parentheses.
⁷On low copy-number plasmids in msh2Δ msh3 Δ msh6 Δ strain.
recognizes the same types of substrates as Msh3; we therefore used the $msh6(3\text{-MBD})$ allele to test the roles of the MBD and mispair recognition in the recruitment of Mlh1-Mlh3. Consistent with previous studies (Flores-Rozas and Kolodner, 1998), deletion of $MLH3$ in an $msh6\Delta$ background, in which MMR is initiated only by Msh2-Msh3, caused a substantial increase in mutation rate, while deletion of $MLH3$ in an $msh3\Delta$ background caused only a small increase in mutation rate, indicating that Mlh1-Mlh3 is not important in repair events initiated by Msh2-Msh6 (Table 4). Interestingly, the small increase in mutation rate in the $msh3\Delta \ mlh3\Delta$ double mutant relative to the $msh3\Delta$ mutant was significant in all three assays ($p < 0.01$, two-tailed Mann-Whitney test), although the mutation rate of the $msh3\Delta \ mlh3\Delta$ double mutant was only significantly higher than that of the $mlh3\Delta$ mutant in the $lys2-10A$ and $CAN1$ assays. These data suggest that an Mlh3-dependent pathway could be involved in a small fraction of Msh2-Msh6-mediated repair events; however, the increases in the mutation rates were small when compared to the effect of an $MLH3$ deletion in an $msh6\Delta$ background. As in the $msh3\Delta$ background, deletion of $MLH3$ in the $msh6(3\text{-MBD})$ $msh3\Delta$ background only resulted in small increases in mutation rate in the $lys2-10A$ and $CAN1$ assays, whereas the $hom3-10$ reversion rate was not significantly changed (Table 4). Thus $msh6(3\text{-MBD})$ behaved like $MSH6$ with respect to genetic interaction with $MLH3$ rather than like $MSH3$. This suggests that interactions between the MutS homologs and the MutL homologs are not specified by the types of mispairs recognized or by physical contacts involving the MBD, but rather by other features of the MutS homologs.
Table 4-4. Effect of MLH3 genotype on mutation rate.

<table>
<thead>
<tr>
<th>Relevant genotype</th>
<th>Strain #</th>
<th>Thr⁺</th>
<th>Lys⁺</th>
<th>Can⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSH3 MSH6 MLH3</td>
<td>3686</td>
<td>2.0</td>
<td>1.5</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.2-3.7] x 10⁻⁹ (1)</td>
<td>[1.0-3.6] x 10⁻⁸ (1)</td>
<td>[4.6-13.5] x 10⁻⁸ (1)</td>
</tr>
<tr>
<td>MSH3 MSH6 mlh3</td>
<td>5296</td>
<td>1.7</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.0-2.7] x 10⁻⁸ (9)</td>
<td>[0.72 –2.9] x 10⁻⁷ (11)</td>
<td>[1.0-1.6] x 10⁻⁷ (2)</td>
</tr>
<tr>
<td>MSH3 msh6 MLH3</td>
<td>4151</td>
<td>3.0</td>
<td>9.8</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.9-3.9] x 10⁻⁸ (15)</td>
<td>[8.4-15.3] x 10⁻⁷ (67)</td>
<td>[6.1-14.9] x 10⁻⁷ (13)</td>
</tr>
<tr>
<td>MSH3 msh6 mlh3</td>
<td>5298</td>
<td>1.9</td>
<td>7.9</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.4-2.8] x 10⁻⁷ (97)</td>
<td>[6.5-9.4] x 10⁻⁶ (540)</td>
<td>[1.3-2.2] x 10⁻⁶ (22)</td>
</tr>
<tr>
<td>msh3 MSH6 MLH3</td>
<td>4149</td>
<td>2.2</td>
<td>1.3</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.6-3.5] x 10⁻⁸ (11)</td>
<td>[0.87-1.9] x 10⁻⁷ (9)</td>
<td>[0.55-1.3] x 10⁻⁷ (2)</td>
</tr>
<tr>
<td>msh3 MSH6 mlh3</td>
<td>5974</td>
<td>3.8</td>
<td>1.9</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.8-5.1] x 10⁻⁸ (19)</td>
<td>[1.7-2.8] x 10⁻⁷ (13)</td>
<td>[1.9-4.6] x 10⁻⁷ (3)</td>
</tr>
<tr>
<td>msh3 msh6(3-MBD) MLH3</td>
<td>5249</td>
<td>6.6</td>
<td>1.0</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[5.7-8.5] x 10⁻⁸ (33)</td>
<td>[0.88-1.3] x 10⁻⁶ (70)</td>
<td>[7.9-14.6] x 10⁻⁷ (13)</td>
</tr>
<tr>
<td>msh3 msh6(3-MBD) mlh3</td>
<td>5882</td>
<td>6.8</td>
<td>1.3</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[5.5-8.6] x 10⁻⁸ (34)</td>
<td>[1.1-1.8] x 10⁻⁶ (86)</td>
<td>[1.1-1.6] x 10⁻⁶ (19)</td>
</tr>
</tbody>
</table>

 Médian rate of 18-35 cultures. 95% confidence interval in square brackets, increase over WT in parentheses.
4.4 DISCUSSION

Here we show that a chimera in which the MBD of Msh6 was replaced by that of Msh3 was functional for MMR and had Msh3-like mispair recognition properties \textit{in vivo} and \textit{in vitro}. However, the chimera retained \textit{MSH6}-like genetic interactions with \textit{MLH3} and retained the ability of Msh6 to function in complex with an Msh2 mutant lacking its MBD.

The ability of the Msh2-Msh6(3-MBD) complex to bind larger insertions \textit{in vitro} and of msh6(3-MBD) to complement an \textit{msh3Δ msh6Δ} strain \textit{in vivo} leads to two interesting conclusions about Msh3 and Msh6. First, these proteins are modular, with a discrete MBD conferring mispair recognition properties without apparent contribution by the other domains of the protein. Second, the mechanisms by which Msh3 and Msh6 communicate mispair binding at the MBD to the C-terminal ATPase domain must be similar, such that the Msh3 MBD can replace functions normally performed by the Msh6 MBD within the Msh2-Msh6 complex. It is not known how mispair binding leads to conformational changes in the ATPase domain, but this result suggests that the underlying molecular mechanisms are highly conserved.

By recognizing the same types of substrates as Msh3, Msh6(3-MBD) allowed us to distinguish between aspects of Msh3 function that are driven by interactions of the Msh3 MBD with mispaired DNA and protein partners and aspects of Msh3 function that are due to the protein scaffold in which the Msh3 MBD is inserted. In addition to a preference for recognizing large insertions and deletions, \textit{MSH3} differs from \textit{MSH6} with respect to interaction with the \textit{msh2ΔMBD} deletion mutant (Goldfarb
and Alani, 2005; Lee et al., 2007), interaction with specific MutL homologs (Flores-Rozas and Kolodner, 1998), suppression of interchromosomal crossover recombination between homeologous substrates (Nicholson et al., 2006), and removal of non-homologous ends during gene conversion (Sugawara et al., 1997) and single-stranded annealing (Sugawara et al., 2000). In the cases of both gene conversion and single-stranded annealing, specific recognition of double-stranded/single-stranded junctions of branched DNA structures by Msh2-Msh3 has been implicated as the basis of MSH3 specificity (Surtees and Alani, 2006).

The inability of Msh3 to function in complex with Msh2ΔMBD (Goldfarb and Alani, 2005; Lee et al., 2007) could be due to defects in recognition of mispairs or defects in downstream steps. Our results showed that the combination of the msh6(3-MBD) and msh2ΔMBD alleles was functional for MMR, indicating that the MBD of Msh3 lacks an inherent requirement for the MBD of Msh2 for mispair recognition. This suggests that the defects in the Msh2ΔMBD-Msh3 complex causing reduced mispair affinity (Lee et al., 2007) result from something subtler about the interaction between the two subunits. We note that insertions of more than one nucleotide modeled into the known MutS structures (Lamers et al., 2000; Obmolova et al., 2000) would probably be positioned such that the unpaired, extrahelical nucleotides could make physical contact with the MBD of the Msh2-like subunit. If this potential interaction were important in activating the Msh2-Msh3 complex, it would explain both the defect of the Msh2ΔMBD-Msh3 complex and the failure of the reverse swap constructs where the Msh6 MBD was placed into the Msh3 protein.
The *msh6(3-MBD)* allele also allowed us to investigate the interactions involved in recruitment of specific MutL homolog complexes during MMR. Deletions of, and mutations in, the N-terminal domain of Mlh1 disrupt the Mlh1-Pms1 interaction with Msh2-Msh3/6; however, the interfaces of the MutL homologs with Msh2, Msh3 and Msh6 have not been defined (Kadyrov et al., 2006; Plotz et al., 2003; Plotz et al., 2006). Previous work indicated that the Mlh1-Mlh3 complex functions primarily in the Msh2-Msh3 pathway, while the Mlh1-Pms1 complex functions through both Msh2-Msh3 and Msh2-Msh6 (Flores-Rozas and Kolodner, 1998). We found that the *msh6(3-MBD)* allele is equivalent to *MSH6*, not *MSH3*, in regard to utilization of the Mlh1-Mlh3 complex, despite the ability of the Msh2-Msh6(3-MBD) complex to recognize the types of mispairs recognized by Msh2-Msh3. These differences suggest that neither the lesion being recognized nor the MBD are sufficient to impart MutL homolog complex specificity and that the interactions between the MutS homologs and the MutL homologs are mediated by other domains of the MutS homologs.
4.5 MATERIALS AND METHODS

Strains, media and plasmids.

*S. cerevisiae* were grown on standard media, either yeast extract/peptone/dextrose (YPD), or complete supplement mixture (CSM) medium (US Biological) lacking specific amino acids. 60 mg/L canavanine was added to CSM –arg medium to select for canavanine resistance.

All yeast strains were isogenic derivatives of S288C. The wild-type strain RDKY 3686 (Amin et al., 2001) has the genotype *MATα ura3-52 leu2Δ1 trp1Δ63 his3Δ200 hom3-10 lys2::InsE-A10*. Derivatives RDKY 4149, 4151, 4234 and 5249 have *msh3::hisG* and/or *msh6::hisG* mutations (Table 1). In RDKY 5153, 5249 and 5882 *MSH6* was replaced by *msh6(3-MBD)* by a popin-popout strategy (Scherer and Davis, 1979). Strains RDKY 5296, 5298, 5882, and 5974 were made from 3686, 4151, 5249, and 4149, respectively, by replacing nucleotides +26 through +2094 of *MLH3* with a PCR fragment of pRS303 containing the *HIS3* gene. RDKY 5963 was made from 4234 by replacement of nucleotides +10 through +2885 of *MSH2* with *HIS3*.

Wild-type *MSH2* was present on the plasmid pII-2 (Reenan and Kolodner, 1992), and *msh2ΔMBD* was present on a derivative of this plasmid missing codons 2-133, pRDK1250. Other plasmids used for complementation assays were based on the *ARS CEN LEU2* vector pRS315 (Sikorski and Hieter, 1989). In pRDK439, *MSH6* is present on a genomic fragment extending from the BamHI site at position –806 relative to the *MSH6* start site to the HindIII site at position +4626, which is 897 bp
downstream of the stop codon, cloned into pRS315. In pRDK1088, *MSH3* is present on a genomic fragment bordered by the same two sites (positions –1632 to +4633) and cloned into pRS315. Other plasmids were made from pRDK439 and pRDK1088 by site-directed mutagenesis or PCR and recombination (Fig. 2). Those with the N-terminal domain of *MSH6* had the native *MSH6* promoter, while in pRDK1103 and pRDK1126 the native *MSH3* promoter was present.

*Qualitative and quantitative evaluation of mutator phenotypes.*

Patch tests were performed by patching colonies onto YPD or CSM –leucine plates and replica-plating onto –threonine, -lysine, –arginine +canavanine, and YPD or –leucine media after 2-3 days of growth. Fluctuation analysis was performed on 14-35 independent cultures for each strain as follows. 10 ml YPD or CSM –leucine cultures were inoculated with whole colonies, grown 12 hours (24 hours for drop-out media), serially diluted, and plated on CSM and CSM –threonine, -lysine, and –arginine +canavanine and grown 3 days before counting colonies. Amino acids were omitted from all media used for plasmid-bearing strains as necessary to maintain plasmid selection in both patch-testing and fluctuation analysis. Yeast was grown at 30° in all experiments. Rates of *hom3-10* reversion, *lys2-10A* reversion and *CAN1* inactivation were determined by the method of the median (Lea, 1948; Marsischky et al., 1996; Tran et al., 1997). P-values were calculated using the two-tailed Mann-Whitney test (http://faculty.vassar.edu/lowry/utest.html). *CAN1* mutation spectra were determined by sequencing the *CAN1* gene in 49-59 independent isolates per strain as described (Marsischky et al., 1996).
**Protein purification.**

Msh2 was coexpressed with wild-type Msh6 or Msh6(3-MBD) in the E. coli strain BL21 CodonPlus (DE3)-RIL (Stratagene) from the plasmids PLANT-2-B-MSH2 (Antony and Hingorani, 2003), pET11a-Msh6 (Antony and Hingorani, 2003), and pRDK1231 which was derived from pET11a-Msh6. Lysis was performed as described (Antony and Hingorani, 2003) and purification was performed by sequential chromatography over SP sepharose, ssDNA cellulose, heparin and HP Q (Pharmacia) columns (details in supplement).

**Surface plasmon resonance analysis**

Surface plasmon resonance experiments were performed with a Biacore 3000 instrument. 236-bp DNA substrates with a biotin conjugated to one end and a centrally-located GT mispair or no mispair were constructed as described (Lin et al., 1998; Mendillo et al., 2005). DNAs harboring insertions of 1, 3, or 4 nucleotides were constructed in the same way. See supplement for details. Between 12 and 14 ng of DNA (75-85 response units) was conjugated to each flow cell of an avidin-coated SA Biacore chip. 10 nM protein was flowed over the chip at 30 µl/min in running buffer (25 mM Tris pH 8, 2% glycerol, 4 mM MgCl₂, 110 mM NaCl, 0.01% igepal, 2 mM DTT). The DNA-coated surface was regenerated with a 15 µl pulse of 3 M NaCl. Data was analyzed with BIAevaluation 3.1 software. Reference subtraction was made from an unmodified flow cell.
4.6 ACKNOWLEDGEMENTS

We thank Dan Mazur for assistance in protein purification, Manju Hingorani for expression plasmids, and Jill Harrington and Marc Mendillo for comments on the manuscript. This work was supported by NIH grants GM50006 and CA92584. S.S.S. was supported by the NIH/NIGMS-funded UCSD Genetics Training Program (T32 GM08666), and C.D.P. was supported as a Robert Black fellow of the Damon Runyon Cancer Research Foundation.

Chapter 4, in full, is a reprint of the material as it appears in Shell, S. S., Putnam, C. D., and Kolodner, R. D. (2007). Chimeric *Saccharomyces cerevisiae* Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. Proc Natl Acad Sci U S A 104, 10956-10961. The dissertation author was the primary investigator in the development and execution of this study, and the principal author of this paper.
4.7 SUPPLEMENTARY MATERIALS
Supplemental Figure 4-1. Example of a patch test experiment to assess functionality of msh3/6 mutants.

The msh3Δ msh6Δ strain RDKY 4234 was transformed with low copy-number plasmids bearing wild-type MSH3, MSH6, or two different MBD-swap mutants as indicated, patched on –leu media and replica-plated onto –thr, -lys, and –leu. MSH3, MSH6, and msh6(3-MBD) complemented the high mutation rate of the strain, while msh6(3-185-233) did not complement any better than the empty vector.
Supplemental Figure 4-2. Plasmids for making heteroduplex DNA substrates.
The polylinker sequences as shown were inserted into BamHI, Apal-digested pBS-SK. The strand displayed on top for each plasmid corresponds to the biotinylated strand in the heteroduplexes (see Sup. Table 1 and methods below). Each polylinker contains a restriction site as shown.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Polylinker sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK505</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK506</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK507</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK508</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK509</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK510</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK511</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK512</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK513</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK514</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK1252</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK1253</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
<tr>
<td>pRDK1254</td>
<td>5' - GATCCCGCCTCCGATGTGATATTGCGGC - 3'</td>
</tr>
<tr>
<td></td>
<td>3' - GCGGGAGCTCTGCTATAC - 5'</td>
</tr>
</tbody>
</table>
## Supplemental Table 4-1. Top and bottom strand combinations for mispaired substrates.

<table>
<thead>
<tr>
<th>Mispair&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Top strand&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Bottom strand&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Enzyme resistance&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>G/T</td>
<td>pRDK505</td>
<td>pRDK506</td>
<td>XhoI, NsiI</td>
</tr>
<tr>
<td>A/C</td>
<td>pRDK506</td>
<td>pRDK505</td>
<td>NsiI, XhoI</td>
</tr>
<tr>
<td>G/G</td>
<td>pRDK507</td>
<td>pRDK508</td>
<td>XhoI, PstI</td>
</tr>
<tr>
<td>C/C</td>
<td>pRDK508</td>
<td>pRDK507</td>
<td>PstI, XhoI</td>
</tr>
<tr>
<td>C/T</td>
<td>pRDK509</td>
<td>pRDK510</td>
<td>SalI, NsiI</td>
</tr>
<tr>
<td>A/G</td>
<td>pRDK510</td>
<td>pRDK509</td>
<td>NsiI, SalI</td>
</tr>
<tr>
<td>A/A</td>
<td>pRDK511</td>
<td>pRDK512</td>
<td>NsiI, Clal</td>
</tr>
<tr>
<td>T/T</td>
<td>pRDK512</td>
<td>pRDK511</td>
<td>Clal, NsiI</td>
</tr>
<tr>
<td>+1 (A)</td>
<td>pRDK1253</td>
<td>pRDK1252</td>
<td>NsiI, PstI</td>
</tr>
<tr>
<td>+2 (CG)</td>
<td>pRDK1254</td>
<td>pRDK1252</td>
<td>FspI, PstI</td>
</tr>
<tr>
<td>+3 (GTC)</td>
<td>pRDK514</td>
<td>pRDK508</td>
<td>AatII, PstI</td>
</tr>
<tr>
<td>+4 (GCGT)</td>
<td>pRDK513</td>
<td>pRDK508</td>
<td>MluI, PstI</td>
</tr>
</tbody>
</table>

<sup>a</sup>Mispair present in product derived from annealing two indicated strands. Sequences of insertions are indicated in parentheses for insertion/deletion mispairs.

<sup>b</sup>Plasmid used for amplification with upstream 5′ biotinylated and downstream 5′ phosphorylated primers.

<sup>c</sup>Plasmid used for amplification with upstream 5′ phosphorylated and downstream unphosphorylated primers.

<sup>d</sup>Mispaired product is resistant to the enzymes indicated while each fully base-paired PCR product is sensitive to one of the enzymes listed.
Supplemental methods

Protein purification

Msh2 was coexpressed with wild-type Msh6 or Msh6(3-MBD) in the E. coli strain BL21 CodonPlus (DE3)-RIL (Stratagene) from the plasmids PLANT-2-B-MSH2 (Antony and Hingorani, 2003), pET11a-Msh6 (Antony and Hingorani, 2003), and pRDK1231, which was derived from pET11a-Msh6. All purification steps were performed in buffer A (25 mM Tris pH 8, 1 mM EDTA, 10% glycerol, 5 mM DTT, 0.02% igepal, 1 mM PMSF, 1 mM benzamadine, 0.5 mg/L bestatin, and 1 mg/L each chymostatin, pepstatin A, aprotinin, and leupeptin) with NaCl as indicated. Cells were lysed in 500 mM NaCl, and supernatants were diluted to 100 mM NaCl and loaded on a 30 mL SP sepharose column. The column was washed with 100 mM NaCl and eluted with a 300 mL linear gradient from 100 mM to 800 mM NaCl. Msh2-Msh6-containing fractions were diluted to 200 mM NaCl and loaded onto a 10 ml ssDNA cellulose column, washed with 200 mM NaCl and step eluted at 600 mM NaCl. Eluted proteins of interest were diluted to 150 mM NaCl and loaded onto a 5 mL Hi-Trap heparin column (Pharmacia), washed with 150 mM NaCl and eluted with a 50 mL linear gradient from 150 to 900 mM NaCl. Msh2-Msh6-containing fractions were diluted to 100 mM NaCl and loaded onto a 5 mL HP Q column (Pharmacia), which was washed with 100 mM NaCl and eluted with a 50 mL linear gradient from 100 to 800 mM NaCl. All columns were run on a Bio Logic DuoFlo instrument (BioRad). Msh2-Msh6(3-MBD) was purified in the same way as the wild-type complex with the
exception that the ssDNA cellulose column was washed with buffer containing 300 mM NaCl.

*Construction of mispaired DNA substrates*

Plasmids as listed in Sup. Fig. 2 were constructed by annealing pairs of oligonucleotides as shown and inserting them into BamHI, ApaI-digested pBS-SK (Lin et al., 1998). Each inserted sequence contains a key restriction site as indicated. pBS-SK is a phagemid, so single-stranded DNA derived from it can be packaged into M13 phage if necessary by infection with helper phage VCSM13. For the present study, 236 bp heteroduplex substrates were made by a PCR-based strategy (Mendillo et al., 2005) as follows. The plasmid indicated in Sup. Table 1 as “top strand” was used as a template for amplification with an upstream primer with a 5’ biotinylation (5’-Biot-ACCATGATTACGCCAAGCTC) and a downstream primer with a 5’ phosphate and a 36 nt Lac operator (5’-Phos-TCACACATCAATTGTTATCCGCTCACAATTTGGGTAAACGCCAGGGTTTTTC). The plasmid indicated in Sup. Table 1 as “bottom strand” was used as a template for amplification by an upstream primer with the sequence shown above and a 5’ phosphate, and an unmodified version of the downstream primer shown above. The PCR products produced were 236 bp, with the variable polylinker sequence shown in Sup. Fig. 2 located roughly in the center. PCR products were then subject to digestion by lambda exonuclease to eliminate the strands with 5’ phosphates. The resulting single-stranded DNAs were mixed as indicated in Sup. Table 1, heated to 95 degrees C for five min and allowed to cool slowly to anneal the complementary strands. To
eliminate contamination by any remaining homoduplex DNA, annealed substrates
were subject to digestion by the two restriction enzymes indicated in Sup. Table 1,
which cleave the two homoduplexes that can potentially be found but are unable to
cleave heteroduplex DNA containing the desired mispair. Full-length substrates were
then separated from the shorter contaminants by HPLC using a Gen-Pak FAX column
(Alani et al., 1997).
4.8 REFERENCES


CHAPTER 5

Summary and future directions
The studies described in this dissertation have addressed several aspects of the mechanisms of initiation of DNA mismatch repair (MMR) in eukaryotic systems. Results presented in Chapter 2 indicate that a dominant mutation in the ATPase domain of Msh6 is sufficient to induce tumorigenesis in mice but that cells from these mice differ from those of Msh6 knock-out mice with respect to drug sensitivities. In Chapter 3 we have shown that the N-terminal region of Msh6 is a module comprising a long unstructured tether to PCNA as well as a second function that is partially redundant with PCNA interaction. In experiments described in Chapter 4 we found that mispair-specificity is conferred to the MutS homologue complexes by a modular mispair-binding domain, and that mispair recognition does not seem to affect the specificity of interaction with MutL homologues. These findings shed light on multiple aspects of the functions of the MutS homologues in the initiation of MMR, and also raise further questions that indicate the need for continued study.
5.1 EFFECT OF A DOMINANT MUTATION IN THE ATPASE DOMAIN OF MSH6

Chapter 2 describes the generation and characterization of a mouse strain harboring a dominant mutation in the ATPase domain of Msh6, and yeast studies to establish the relationship between this residue in different eukaryotes (Yang et al., 2004).

The *Saccharomyces cerevisiae* mutation Gly1067Asp in Msh6 was previously shown to confer a dominant-negative mutator phenotype (Das Gupta and Kolodner, 2000; Hess et al., 2002). A Msh6 mutation of Thr1219Ile was detected in a human patient with hereditary non-polyposis colorectal cancer (HNPCC) who had no other MMR gene mutations, and sequence comparisons suggested that Gly1067 in yeast and Thr1219 in human Msh6 are homologous residues (Berends et al., 2002). To determine whether a missense mutation in Msh6 can be pathological, a mouse strain was created with a mutation of the equivalent residue Thr1217 to Asp.

Mice homozygous for the Thr1217Asp mutation showed increased cancer susceptibility similar to that of Msh6 homozygous knock-out mice. Unlike Msh6 knock-out mice, the Thr1217Asp mice displayed high levels of microsatellite instability, which is consistent with a dominant-negative activity for the mutant protein and was also observed in tumors from the Thr1219Ile patient (Berends et al., 2002). Cell extracts from the Thr1217Asp mice were highly defective for MMR. Surprisingly, Thr1217Asp mouse embryonic fibroblasts (MEFs) were as sensitive at
wild-type MEFs to cisplatin, N-methyl-N’-nitro-N-nitrosoguanidine (MNNG) and 6-thioguanine (6TG), in contrast to the resistance shown to all three agents by Msh6 knock-out MEFs.

In order to correlate previous findings for the *S. cerevisiae* Gly1067Asp mutation with the mouse Thr1217Asp findings and the reported Thr1219Ile human patient (Berends et al., 2002; Das Gupta and Kolodner, 2000; Hess et al., 2002), we made yeast strains bearing Gly1067Thr and Gly1067Ile Msh6 mutations. The Gly1067Thr mutant had only a very small increase in mutation rate, indicating that Gly1067 in yeast Msh6 likely performs a homologous function to Thr1219 and 1217 in humans and mice, respectively. In contrast, the Gly1067Ile mutation conferred a strong dominant-negative phenotype that was indistinguishable from that of the Gly1067Asp mutation.

Taken together, these findings strongly support the contention that Msh6 Thr1219Ile is a true HNPCC-causing mutation. They also indicate that *S. cerevisiae* residue Gly1067 is homologous to the Thr found at the equivalent position in Msh6 in higher eukaryotes, and therefore that the findings of subsequent biochemical characterizations of the msh6 Gly1067Asp mutant can probably be considered indicative of the defects in mouse and human Thr to Asp and Thr to Ile mutants (Hess et al., 2006).

This work also has implications for the mechanisms by which MMR mutations cause cancer and drug resistance. It has been suggested that treatment with drugs that cause DNA lesions recognized by the MutS homologues results in repeated futile cycles of MMR that ultimately result in strand breaks and signaling to induce
apoptosis (Karran and Bignami, 1994; Modrich, 1997). An alternative hypothesis suggests that MMR proteins may be involved in direct signaling of some types of damage (Duckett et al., 1996; Modrich, 1997). The findings presented in this work indicate that functional MMR is not necessary for the damage sensitivity conferred by the Msh2-Msh6 complex, since Msh6 Thr1217Asp does not induce MMR but does not behave like an Msh6 null mutant with respect to drug sensitivity. Biochemical analysis of the yeast mutant revealed that it was unable to interact with Mlh1-Pms1, suggesting that the MMR reaction terminates at an early step prior to strand excision (Hess et al., 2006). Taken together, these data therefore support a direct signaling model and appear to be inconsistent with a model in which futile cycles of repair are required for drug sensitivity.

It is significant that the mouse Msh6 Thr1217Asp mutation does not appear to cause defects in apoptosis in response to DNA damage, but still causes tumorigenesis. This indicates that increased rates of mutation alone are sufficient to drive tumorigenesis, suggesting that the primary defect that leads to cancer susceptibility in HNPCC patients is increased mutagenesis rather than defects in apoptosis or damage signaling.
5.2 STRUCTURE AND FUNCTION OF THE N-TERMINAL REGION OF MSH6

In the work described in Chapter 3, we characterized the N-terminal region (NTR) of Msh6 from *S. cerevisiae* using a combination of structural and functional techniques (Shell et al., 2007b). Partial proteolysis, size-exclusion chromatography and small-angle X-ray scattering (SAXS) were used to assess the structure of the NTR as an isolated domain and in the context of the full-length MutSα complex, both in the presence and absence of PCNA. Genetic analysis was then used to determine the requirements for the NTR for functional MMR, as well as the exchangeability of the NTRs of Msh3 and Msh6.

The first 304 residues of Msh6 appear to be unstructured when expressed and purified as an isolated construct. It is rapidly cleaved by trypsin to produce products too small to visualize on a gel, elutes very early from a size-exclusion column, is insensitive to heat, and has the X-ray scattering properties of an unfolded polymer. It binds stably to PCNA but this interaction does not induce detectable structuring. In the context of the full-length Msh2-Msh6 complex, the NTR is similarly unstructured. Partial proteolysis produces a truncated form of Msh6 that we analyzed by mass spectrometry and found to be missing approximately 251 residues from the N-terminus. The NTR remains unstructured upon binding PCNA, and SAXS analysis indicates that there is no substantial or sustained contact between PCNA and the globular portions of Msh2 or Msh6.
Genetic analysis revealed that a deletion of the first 123 residues of the NTR caused only a mild MMR defect in vivo that was indistinguishable from that of a point mutation in the PCNA-interaction motif, consistent with previous reports (Clark et al., 2000; Flores-Rozas et al., 2000). Larger deletions had more severe phenotypes, and a deletion of 251 residues was non-functional in one of the three in vivo mutation reporter assays used, despite being expressed normally, localized to the nucleus, and retaining normal DNA binding, mispair discrimination and ATP-responsiveness in vitro when overexpressed and purified. However, an internal deletion of residues 51-251 was mostly functional in vivo, indicating that the functions of the PCNA-interaction region and the internal portion of the NTR are partially redundant.

Chimeric constructs encoding the NTR of Msh6 fused to Msh3 and vice versa produced functional proteins, indicating that the NTRs of Msh3 and Msh6 perform similar functions and behave as independent modules. The Msh3 NTR shares little homology to the Msh6 NTR aside from possessing an N-terminal PIP-box, but like the Msh6 NTR it contains a relatively high proportion of hydrophilic and charged residues, and computational predictions suggest that it too is unstructured. By adjusting the position at which the Msh3 NTR was fused to Msh6 we mapped the boundary between the NTR of Msh6 and the mispair-binding domain to between residues 273 and 282, and confirmed the results by making small deletions within Msh6. A surprising result was that when we fused the NTR of Msh6 to the N-terminus of Msh2, which normally lacks an NTR, the resulting chimera rescued the severe defect of an Msh6 NTR deletion mutant.
Our finding that Msh6, and probably Msh3 as well, interact with PCNA through long unstructured tethers is consistent with the genetic results indicating that the two NTRs are interchangeable and that the Msh6 NTR can tolerate a variety of internal deletions. The ability of the Msh6 NTR to function when located on the other subunit of the heterodimer strongly suggests that interaction with PCNA does not by itself contribute to strand discrimination or correct orientation of MutS\(\alpha\) on the DNA. Taken together, these results suggest that the function of the MutS homologue-PCNA interaction may be to achieve appropriate sub-nuclear localization of the MutS homologues by bringing them to regions of newly replicated DNA so that they can efficiently locate and initiate repair of mispairs. This is consistent with a previous report that a human Msh6 mutant in which the PIP-box was deleted failed to localize normally to replication foci (Kleczkowska et al., 2001).

The unstructured nature of the Msh6 NTR led us to wonder if other proteins that interact with PCNA do so via unstructured tethers. Disorder tendency predictions of regions containing putative PIP-boxes in all \textit{S. cerevisiae} ORFs showed that PIP-boxes that are known to mediate interaction with PCNA are frequently located in regions with high predicted disorder, while PIP-box-like sequences in proteins that have been found not to interact with PCNA or whose interaction status is unknown tend to lie in regions predicted to be ordered. These findings suggest that many interactions with PCNA likely occur via unstructured tethers, although most are not as long as the tethers formed by the N-termini of Msh3 and Msh6. Interaction by unstructured tethers may be required to bring proteins into close proximity to DNA.
without tightly restricting their positions. Unstructured tethers could also be required to allow multiple interactors to bind to PCNA simultaneously.

Because of the conservation of PIP-boxes on eukaryotic Msh3 and Msh6 proteins, the previously reported finding that this interaction is mostly dispensable for MMR was somewhat surprising (Clark et al., 2000; Flores-Rozas et al., 2000). Our finding that the internal region of the Msh6 NTR is functionally partially redundant with the PIP-box-containing region suggests that Msh6 may interact with another protein or proteins that can serve to facilitate localization to appropriate regions of the genome.
Chapter 4 describes a series of experiments that probed the functional modularity of mispair binding (Shell et al., 2007a). We created a chimeric *S. cerevisiae* MutS homologue that comprised Msh6 sequence except for the mispair-binding domain (MBD), which was exchanged for the MBD of Msh3. Breakpoints for construction of the chimera were based on sequence alignments with prokaryotic MutS sequences for which crystal structures have been solved (Lamers et al., 2000; Obmolova et al., 2000). The chimera, termed msh6(3-MBD), was functional for MMR *in vivo*, indicating that mechanisms of communication between the MBD and other domains of the protein are conserved between Msh3 and Msh6.

Analysis of the *in vivo* mutation spectrum of a strain expressing Msh6(3-MBD) revealed that like Msh3, it initiates repair of small insertions and deletions but does not initiate significant repair of base:base mispairs. The purified Msh2-Msh6(3-MBD) complex showed much lower affinity for a DNA substrate containing a base:base mispair than wild-type Msh2-Msh6, but had higher affinity for DNA containing loops of 3 or 4 unpaired nucleotides. This indicates that the MBD behaves as a functional module that confers mispair specificity to Msh3 and Msh6.

Reciprocal constructs in which the MBD of Msh3 was replaced by that of Msh6 failed to demonstrate MMR activity *in vivo*, despite testing a number of different junction positions. The most likely explanations are that either the
breakpoints we tested all resulted in disruptions of the structure of the protein, or that the Msh6 MBD fails to communicate appropriately with the other domains of Msh3.

Because recent studies had indicated differential requirements for the MBD-like domain of Msh2 by Msh3 and Msh6 (Goldfarb and Alani, 2005; Lee et al., 2007), we assessed the ability of Msh6(3-MBD) to function in combination with an Msh2 deletion mutant missing the first 133 amino acids. Surprisingly, unlike Msh3, Msh6(3-MBD) was fully functional in combination with the Msh2 deletion mutant. This result indicates that the mechanism of mispair binding by Msh3 does not require the MBD-like domain of Msh2 \textit{per se}, but rather that the Msh2 MBD-like domain may be involved in conformational changes that accompany mispair binding and transmit signals to other domains of the Msh2-Msh3 complex.

Previous work has indicated that MutS$\alpha$ functions exclusively through the Mlh1-Pms1 complex to initiate the downstream steps of MMR, while MutS$\beta$ can function with either Mlh1-Pms1 or Mlh1-Mlh3 (Flores-Rozas and Kolodner, 1998). It is not known where in the MutS homologue complexes interaction with the MutL homologues occurs, or how choice of MutL homologue complex is determined. We examined a strain bearing \textit{msh6(3-MBD)} in the presence and absence of functional \textit{MLH3} and found that \textit{MLH3} status had no effect on mutation rate, indicating that the Msh2-Msh6(3-MBD) complex is probably not able to function through the Mlh1-Mlh3 complex. This finding has two implications: first, that the MBDs of Msh3 and Msh6 likely do not contribute to the interaction interface with the MutL homologues; and second, that choice of MutL homologues seems not to be related to the type of mispair bound by the MutS homologue complexes.
5.4 FUTURE DIRECTIONS

The studies presented in Chapter 2 indicate that downstream steps of MMR are not required for Msh2-Msh6-dependent sensitivity to a number of DNA damaging agents. This suggests that the mechanism of MMR-dependent sensitivity does not involve futile cycles of repair, indicating the need for an alternative hypothesis that is consistent with the data. Previous work has indicated that loss of Mlh1 is sufficient to decrease sensitivity to MNNG (Koi et al., 1994), but the dominant Msh6 mutation described in Chapter 2 has subsequently been shown to abolish interaction of Mlh1-Pms1 with Msh2-Msh6, indicating that in the case of this mutant, Mlh1 is not involved in the mechanism that mediates sensitivity (Hess et al., 2006). Study of the roles of MMR proteins in DNA damage signaling has been hampered by the lack of effect on drug sensitivity of Msh2 and Mlh1 deletions in *S. cerevisiae*, which could otherwise provide a genetically tractable system for rigorous study (Xiao et al., 1995). However, it was recently shown that in the absence of both the methylguanine methyltransferase Mgt1 and homologous recombination, loss of either Msh2 or Mlh1 induces resistance to MNNG in *S. cerevisiae* (Cejka et al., 2005); it is hoped that further study in this system may shed light on the mechanism of action of MMR proteins in the processes of drug sensitivity and tolerance.

The work described in Chapter 3 raised several further questions. One is the identity of the putative factor(s) that interacts with the internal region of the Msh6 NTR. Subsequent experiments in our lab have been performed to address this question and will be continued. Another question is whether the NTR of Msh3 contains a
similar region that is functionally partially redundant with the PCNA-interacting region, as would be predicted based on results presented here and by others (Clark et al., 2000). Experiments are also underway to address this question. This work also indicates a need for further study to understand the role of PCNA in MMR. Our findings are consistent with a model in which PCNA serves to localize MutSα and MutSβ to regions of newly replicated DNA, but this hypothesis remains to be proven by assessing physical localization of these proteins within nuclei. The status of PCNA during and after mispair recognition is also unclear. A previous report suggested that PCNA dissociates from Msh6 upon mispair binding (Lau and Kolodner, 2003). However, in that study PCNA was not loaded onto the DNA strands and Mlh1-Pms1 was not present, so it is possible that under physiological conditions the interaction dynamics are different, particularly since a weak direct interaction between PCNA and Mlh1-Pms1 has recently been reported (Lee and Alani, 2006).

A broader and more challenging question raised by this work is how interactions between PCNA and its numerous binding partners are regulated, and how the various interactions are coordinated to allow efficient exchange of binding partners and ensure the presence of the appropriate interacting proteins during the various stages of DNA replication and repair. The cell cycle inhibitor p21 is known to function by interacting tightly with PCNA to prevent it from binding polymerase δ, thereby preventing DNA replication (Waga et al., 1994). However, little has been reported about the relative affinities of other PCNA-interacting proteins. Our data indicate that many of PCNA’s partners may interact via unstructured tethers, suggesting that multiple proteins could bind to each PCNA trimer simultaneously. More work is
clearly required to build an understanding of the dynamics of PCNA-interacting proteins during the processes of replication and repair.

Chapter 4 addresses the mechanism of mispair recognition by Msh3. Our work indicates that the mispair-binding domains of Msh3 and Msh6 interact with other domains of these proteins by homologous mechanisms, while differences in the mispair-interaction sequences confer mispair specificity. An important remaining question is how Msh3 recognizes mispairs on a molecular basis; crystal structures of Msh2-Msh3 bound to various mispairs will be needed to unambiguously answer this.

Study of the mispair-binding domain chimera indicated that interaction with the MutL homologues does not involve the mispair-binding domains. This narrows down the possibilities but does not answer the important question of which domains of the MutS homologues form the interaction interface with the MutL homologues. Further studies are needed to fill this gap in our knowledge of the mechanism of MMR.
5.5 REFERENCES


APPENDIX A
Table A-1: Summary of all plasmid-born mutants created and tested.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Complementation of an msh3Δ msh6Δ strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1063</td>
<td>msh6G1067I</td>
<td>None (dominant mutator in WT strain)</td>
</tr>
<tr>
<td>pRDK1063</td>
<td>msh6G1067T</td>
<td>Close to wild-type</td>
</tr>
<tr>
<td>pRDK1067</td>
<td>msh6K336Y, F337K</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1068</td>
<td>msh6K336Y, F337K, E339K</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1156</td>
<td>msh F10A, F11A</td>
<td>Moderate</td>
</tr>
<tr>
<td>pRDK1159</td>
<td>msh3Y157K, K158F</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1162</td>
<td>msh3Y157K, K158F, K160E</td>
<td>Slight to modest</td>
</tr>
<tr>
<td>pRDK1172</td>
<td>msh3Y157K, K158F, K160E, F162Y</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1177</td>
<td>msh3Y157K, K158F, Y159F, K160E</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1169</td>
<td>msh3Y157K, K158F, K160E, C161L, F162Y, A163E, E164K</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1160</td>
<td>Encodes Msh6 with residues 305-421 replaced by Msh3 residues 126-256 that in addition contains Y157K and K158F mutations in the Msh3 sequence</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1161</td>
<td>Encodes Msh6 with residues 305-421 replaced by Msh3 residues 126-256 that in addition contains Y157K, K158F, and K160E mutations in the Msh3 sequence</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1087</td>
<td>Encodes Msh6 with residues 305-421 replaced with Msh3 residues 126-256 ( ^* )</td>
<td>Close to wild-type MSH3 (chapter 4)</td>
</tr>
<tr>
<td>pRDK1092</td>
<td>Encodes Msh6 with residues 305-371 replaced with Msh3 residues 126-204</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1093</td>
<td>Encodes Msh6 with residues 335-421 replaced with Msh3 residues 156-256</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1097</td>
<td>Encodes Msh6 with residues 335-371 replaced with Msh3 residues 156-204</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1112</td>
<td>Encodes Msh6 with residues 328-385 replaced with Msh3 residues 149-218</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1103</td>
<td>Encodes Msh3 with residues 126-256 replaced with Msh6 residues 305-421</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1119</td>
<td>Encodes residues 1-421 of Msh6 fused to residues 257-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1128</td>
<td>Encodes residues 1-455 of Msh6 fused to residues 282-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1151</td>
<td>Encodes residues 1-466 of Msh6 fused to residues 293-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1153</td>
<td>Encodes residues 1-483 of Msh6 fused to residues 310-1018 of Msh3</td>
<td>None</td>
</tr>
</tbody>
</table>
Table A-1 continued: Summary of all plasmid-born mutants created and tested.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Complementation of an (msh3\Delta msh6\Delta) strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1175</td>
<td>Encodes residues 1-370 of Msh6 fused to residues 204-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1182</td>
<td>Encodes residues 1-441 of Msh6 fused to residues 270-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1183</td>
<td>Encodes residues 1-441 of Msh6 fused to residues 271-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1176</td>
<td>Encodes Msh3 with residues 177-202 replaced by residues 356-369 of Msh6 and in addition contains (msh3) point mutations (Y157K, K158F,) and (K160E)</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1120</td>
<td>Encodes residues 1-304 of Msh6 fused to residues 126-1018 of Msh3</td>
<td>Moderate</td>
</tr>
<tr>
<td>pRDK1279</td>
<td>Encodes residues 1-304 of Msh6 with F33A and F34A mutations, fused to residues 126-1018 of Msh3</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1285</td>
<td>Encodes residues 1-251 of Msh6 fused to residues 126-1018 of Msh3</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1286</td>
<td>Encodes residues 1-251 of Msh6 with F33A and F34A mutations, fused to residues 126-1018 of Msh3</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1284</td>
<td>Encodes residues 1 and 124-251 of Msh6 fused to residues 126-1018 of Msh3</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1289</td>
<td>Encodes residues 1 and 51-251 of Msh6 fused to residues 126-1018 of Msh3</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1181</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 252-1242 of Msh6</td>
<td>Good (chapter 3)</td>
</tr>
<tr>
<td>pRDK1291</td>
<td>Encodes residues 1-125 of Msh3 with F10A and F11A mutations, fused to residues 252-1242 of Msh6</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1209</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 262-1242 of Msh6</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1198</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 272-1242 of Msh6</td>
<td>Like wild-type (chapter 3)</td>
</tr>
<tr>
<td>pRDK1292</td>
<td>Encodes residues 1-125 of Msh3 with F10A and F11A mutations, fused to residues 272-1242 of Msh6</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1260</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 285-1242 of Msh6</td>
<td>Little (chapter 3)</td>
</tr>
<tr>
<td>pRDK1261</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 295-1242 of Msh6</td>
<td>Little (chapter 3)</td>
</tr>
<tr>
<td>pRDK1122</td>
<td>Encodes residues 1-125 of Msh3 fused to residues 305-1242 of Msh6</td>
<td>None (chapter 3)</td>
</tr>
<tr>
<td>pRDK1287</td>
<td>Encodes residues 1 and 22-125 of Msh3 fused to residues 252-1242 of Msh6</td>
<td>Poor</td>
</tr>
<tr>
<td>pRDK1288</td>
<td>Encodes residues 1 and 73-125 of Msh3 fused to residues 252-1242 of Msh6</td>
<td>Poor</td>
</tr>
</tbody>
</table>
Table A-1 continued: Summary of all plasmid-born mutants created and tested.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Complementation of an msh3Δ msh6Δ strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1126</td>
<td>Encodes residues 1-256 of Msh3 fused to residues 422-1242 of Msh6</td>
<td>Moderate</td>
</tr>
<tr>
<td>pRDK1127</td>
<td>Encodes residues 1-643 of Msh6 fused to residues 445-1018 of Msh3</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1130</td>
<td>Encodes residues 1-444 of Msh3 fused to residues 644-1242 of Msh6</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1118</td>
<td>Encodes Msh3 with residues 742-1018 replaced with residues 922-1242 of Msh6</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1147</td>
<td>Encodes Msh6Δ2-24</td>
<td>Close to wild-type</td>
</tr>
<tr>
<td>pRDK1129</td>
<td>Encodes Msh6Δ2-50</td>
<td>Good (chapter 3)</td>
</tr>
<tr>
<td>pRDK1141</td>
<td>Encodes Msh6Δ2-123</td>
<td>Good (chapter 3)</td>
</tr>
<tr>
<td>pRDK1203</td>
<td>Encodes Msh6Δ2-143</td>
<td>Moderate (chapter 3)</td>
</tr>
<tr>
<td>pRDK1204</td>
<td>Encodes Msh6Δ2-163</td>
<td>None (chapter 3)</td>
</tr>
<tr>
<td>pRDK1205</td>
<td>Encodes Msh6Δ2-183</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1143</td>
<td>Encodes Msh6Δ2-201</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1145</td>
<td>Encodes Msh6Δ2-251</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1247</td>
<td>Encodes Msh6Δ2-271</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1245</td>
<td>Encodes Msh6Δ2-284</td>
<td>Little or none</td>
</tr>
<tr>
<td>pRDK1095</td>
<td>Encodes Msh6Δ2-294</td>
<td>None (chapter 3)</td>
</tr>
<tr>
<td>pRDK1242</td>
<td>Encodes Msh6Δ2-304</td>
<td>None (chapter 3)</td>
</tr>
<tr>
<td>pRDK1206</td>
<td>Encodes Msh6Δ51-123</td>
<td>Close to wild-type</td>
</tr>
<tr>
<td>pRDK1207</td>
<td>Encodes Msh6Δ51-201</td>
<td>Close to wild-type</td>
</tr>
<tr>
<td>pRDK1208</td>
<td>Encodes Msh6Δ51-210</td>
<td>Close to wild-type</td>
</tr>
<tr>
<td>pRDK1197</td>
<td>Encodes Msh6Δ51-251</td>
<td>Good (chapter 3)</td>
</tr>
<tr>
<td>pRDK1210</td>
<td>Encodes Msh6Δ51-251 with F33A and F34A</td>
<td>Moderate (chapter 3)</td>
</tr>
<tr>
<td>pRDK1248</td>
<td>Encodes Msh6Δ51-271</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1246</td>
<td>Encodes Msh6Δ51-284</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1244</td>
<td>Encodes Msh6Δ51-294</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1243</td>
<td>Encodes Msh6Δ51-304</td>
<td>None (chapter 3)</td>
</tr>
<tr>
<td>pRDK1256</td>
<td>Encodes Msh6Δ252-271</td>
<td>Like wild-type (chapter 3)</td>
</tr>
<tr>
<td>pRDK1257</td>
<td>Encodes Msh6Δ252-284</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1258</td>
<td>Encodes Msh6Δ252-294</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1259</td>
<td>Encodes Msh6Δ252-304</td>
<td>Poor (chapter 3)</td>
</tr>
<tr>
<td>pRDK1281</td>
<td>Encodes Msh3Δ2-21</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1282</td>
<td>Encodes Msh3Δ2-72</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1315</td>
<td>Encodes Msh3Δ2-83</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1316</td>
<td>Encodes Msh3Δ2-93</td>
<td>Good</td>
</tr>
<tr>
<td>pRDK1283</td>
<td>Encodes Msh3Δ2-125</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1303</td>
<td>Encodes Msh3Δ73-125</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1317</td>
<td>Encodes Msh3Δ84-125</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1318</td>
<td>Encodes Msh3Δ84-125 with F10A and F11A mutations</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1319</td>
<td>Encodes Msh3Δ94-125</td>
<td>None</td>
</tr>
<tr>
<td>pRDK1320</td>
<td>Encodes Msh3Δ94-125 with F10A and F11A mutations</td>
<td>None</td>
</tr>
</tbody>
</table>
Table A-1 continued: Summary of all plasmid-born mutants created and tested.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Complementation of an $msh2\Delta msh3\Delta msh6\Delta$ strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1219</td>
<td>Encodes Msh2 with the first residue replaced by residues 1-304 of Msh6</td>
<td>With $MSH6$, like wild-type; with $msh6\Delta251$, null.</td>
</tr>
<tr>
<td>pRDK1221</td>
<td>Encodes Msh2 with the first residue replaced by residues 1-251 of Msh6</td>
<td>With $MSH6$, like wild-type; with $msh6\Delta251$, near wild-type. See Chapter 3</td>
</tr>
<tr>
<td>pRDK1314</td>
<td>Encodes Msh2 with the first residue replaced by residues 1-251 of Msh6 with F33A and F34A mutations</td>
<td>With $msh6\Delta251$, null.</td>
</tr>
<tr>
<td>pRDK1290</td>
<td>Encodes Msh2 with the first residue replaced by residues 1-125 of Msh3</td>
<td>With $MSH3$, moderate; with $msh3\Delta125$, none</td>
</tr>
</tbody>
</table>
Table A-1 continued: Summary of all plasmid-born mutants created and tested.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Complementation of either a pms1Δ strain, an mlh1Δ strain, or an mlh3Δ msh6Δ strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRDK1149</td>
<td>Encodes residues 1-373 of Mlh3 fused to residues 348-769 of Mlh1</td>
<td>None (mlh3Δ msh6Δ and mlh1Δ)</td>
</tr>
<tr>
<td>pRDK1184</td>
<td>Encodes residues 1-373 of Mlh3 fused to residues 254-873 of Pms1, with Pms1 promoter</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1185</td>
<td>Encodes residues 1-373 of Mlh3 fused to residues 354-873 of Pms1, with Mlh3 promoter</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1186</td>
<td>Encodes residues 1-353 of Pms1 fused to residues 374-715 of Mlh3, with Pms1 promoter</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1223</td>
<td>Encodes residues 1-460 of Mlh3 fused to residues 460-873 of Pms1, with Pms1 promoter</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1336</td>
<td>Encodes residues 1-671 of Pms1 fused to residues 492-715 of Mlh3</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1337</td>
<td>Encodes residues 1-491 of Mlh3 fused to residues 672-873 of Pms1</td>
<td>None (mlh3Δ msh6Δ and pms1Δ)</td>
</tr>
<tr>
<td>pRDK1199</td>
<td>Encodes Pms1 Δ361-460</td>
<td>Like wild-type (pms1Δ)</td>
</tr>
<tr>
<td>pRDK1200</td>
<td>Encodes Pms1 Δ461-560</td>
<td>Like wild-type (pms1Δ)</td>
</tr>
<tr>
<td>pRDK1202</td>
<td>Encodes Pms1 Δ561-660</td>
<td>Good (pms1Δ)</td>
</tr>
</tbody>
</table>

a All Msh3 codon numbering begins at the second ATG (Harrington and Kolodner, 2007).

Figure A-1. Summary of Msh3/Msh6 chimeras.
Msh6 sequence is displayed in gray and Msh3 sequence is displayed in black.
Numbering refers to amino acid positions in wild-type Msh3 or Msh6 as indicated by color. The letters “AA” superimposed over a construct indicate mutation of Phe10 and Phe11 (Msh3) or Phe33 and Phe34 (Msh6) to Ala. Ability of plasmids encoding the indicated chimeras to complement the high single-nucleotide deletion rate of an msh3Δ msh6Δ double mutant is indicated on a qualitative scale as follows: ++++, similar to wild-type MSH3 or MSH6; +++, modestly defective relative to wild-type MSH3 or MSH6; ++, substantially defective relative to wild-type MSH3 or MSH6; +, highly defective relative to wild-type MSH3 or MSH6 but still measurably different than a null allele; -, no apparent complementation.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Schematic of protein</th>
<th>Complementation of an msh3Δ msh6Δ strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT Msh6 pRDK439</td>
<td>1-1242</td>
<td>++++</td>
</tr>
<tr>
<td>WT Msh3 pRDK1088</td>
<td>1-1018</td>
<td>++++</td>
</tr>
<tr>
<td>pRDK1120</td>
<td>1-394</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1285</td>
<td>1-251</td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1286</td>
<td>1-251</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1289</td>
<td>124-251</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1284</td>
<td>1-126</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1122</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1261</td>
<td>1-125</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1260</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1198</td>
<td>1-125</td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1292</td>
<td>1-125</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1209</td>
<td>1-125</td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1181</td>
<td>1-125</td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1291</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1287</td>
<td>1-125</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1288</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1087</td>
<td>1-125</td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1092</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1093</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1097</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1112</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1119</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1120</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1151</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1153</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1175</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1182</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1183</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1184</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1127</td>
<td>1-125</td>
<td></td>
</tr>
<tr>
<td>pRDK1130</td>
<td>1-125</td>
<td></td>
</tr>
</tbody>
</table>
Figure A-2. Summary of Msh3 and Msh6 deletion mutants.
Msh6 sequences are displayed in gray and Msh3 sequences are displayed in black. Numbering refers to amino acid positions in wild-type Msh3 or Msh6 as indicated by color. The letters “AA” superimposed over a construct indicate mutation of Phe10 and Phe11 (Msh3) or Phe33 and Phe34 (Msh6) to Ala. Ability of plasmids encoding the indicated chimeras to complement the high single-nucleotide deletion rate of an msh3Δ msh6Δ double mutant is indicated on a qualitative scale as follows: ++++, similar to wild-type MSH3 or MSH6; ++, modestly defective relative to wild-type MSH3 or MSH6; +, substantially defective relative to wild-type MSH3 or MSH6; +, highly defective relative to wild-type MSH3 or MSH6 but still measurably different than a null allele; -, no apparent complementation.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Schematic of protein</th>
<th>Complementation of an \textit{msh3Δ msh6Δ} strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT MshΔ</td>
<td>1-1242</td>
<td>++++</td>
</tr>
<tr>
<td>pRDK439</td>
<td>1-1018</td>
<td>++++</td>
</tr>
<tr>
<td>pRDK1156</td>
<td>1-1018</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1281</td>
<td>22-1018</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1282</td>
<td>73-1018</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1315</td>
<td>84-1018</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1316</td>
<td>94-1018</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1203</td>
<td>126-1018</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1303</td>
<td>126-1018</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1317</td>
<td>126-1018</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1319</td>
<td>126-1018</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1147</td>
<td>29-1242</td>
<td>++++</td>
</tr>
<tr>
<td>pRDK1141</td>
<td>174-1242</td>
<td>++++</td>
</tr>
<tr>
<td>pRDK1203</td>
<td>144-1242</td>
<td>++</td>
</tr>
<tr>
<td>pRDK1204</td>
<td>164-1242</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1205</td>
<td>184-1242</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1143</td>
<td>202-1242</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1145</td>
<td>252-1242</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1247</td>
<td>272-1242</td>
<td>+</td>
</tr>
<tr>
<td>pRDK1245</td>
<td>285-1242</td>
<td>+/-</td>
</tr>
<tr>
<td>pRDK1095</td>
<td>295-1242</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1242</td>
<td>305-1242</td>
<td>-</td>
</tr>
<tr>
<td>pRDK1206</td>
<td>1-50</td>
<td>124-1242</td>
</tr>
<tr>
<td>pRDK1207</td>
<td>1-13</td>
<td>202-1242</td>
</tr>
<tr>
<td>pRDK1208</td>
<td>1-13</td>
<td>202-1242</td>
</tr>
<tr>
<td>pRDK1197</td>
<td>1-50</td>
<td>252-1242</td>
</tr>
<tr>
<td>pRDK1210</td>
<td>1-50</td>
<td>252-1242</td>
</tr>
<tr>
<td>pRDK1240</td>
<td>1-50</td>
<td>272-1242</td>
</tr>
<tr>
<td>pRDK1248</td>
<td>1-50</td>
<td>285-1242</td>
</tr>
<tr>
<td>pRDK1246</td>
<td>1-50</td>
<td>285-1242</td>
</tr>
<tr>
<td>pRDK1244</td>
<td>1-50</td>
<td>305-1242</td>
</tr>
<tr>
<td>pRDK1243</td>
<td>1-251</td>
<td>272-1242</td>
</tr>
<tr>
<td>pRDK1256</td>
<td>1-251</td>
<td>272-1242</td>
</tr>
<tr>
<td>pRDK1257</td>
<td>1-251</td>
<td>285-1242</td>
</tr>
<tr>
<td>pRDK1258</td>
<td>1-251</td>
<td>285-1242</td>
</tr>
<tr>
<td>pRDK1259</td>
<td>1-251</td>
<td>305-1242</td>
</tr>
</tbody>
</table>
Figure A-3. Summary of Mlh1, Mlh3, and Pms1 mutants.
Mlh1 sequences are displayed in gray, Mlh3 sequences in black, and Pms1 sequences in white. Numbering refers to amino acid positions in wild-type Mlh1, Mlh3, or Pms1 as indicated by color. Ability of plasmids encoding the indicated mutants to complement the high mutation rates of each of the three indicated strains is indicated on a qualitative scale as follows: ++++, defect fully complemented; ++, modestly defective relative to complementation by wild-type \textit{MLH1}, \textit{MLH3}, or \textit{PMS1}; +, substantially defective relative to wild-type \textit{MLH1}, \textit{MLH3}, or \textit{PMS1}; +, highly defective relative to wild-type \textit{MLH1}, \textit{MLH3}, or \textit{PMS1} but distinguishable from a null allele; -, no apparent complementation. N/D, not determined.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Schematic of protein</th>
<th>Complementation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>\textit{mhl1}\Delta</td>
</tr>
<tr>
<td>WT Mlh1</td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td>pRDK1338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT Mlh3</td>
<td>1-716</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT Pms1</td>
<td>1-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK435</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRDK1199</td>
<td>1-360-461-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1200</td>
<td>1-460-561-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1202</td>
<td>1-560-661-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1149</td>
<td>1-373-348-769</td>
<td></td>
</tr>
<tr>
<td>pRDK1185</td>
<td>1-373-354-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1186</td>
<td>1-553-374-715</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1337</td>
<td>1-491-672-873</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK1336</td>
<td>1-671-492-715</td>
<td>N/D</td>
</tr>
<tr>
<td>pRDK12ZZ3</td>
<td>1-480-480-873</td>
<td>N/D</td>
</tr>
</tbody>
</table>