BIOPHYSICAL AND BIOCHEMICAL INVESTIGATIONS OF PROTEINS INVOLVED IN GENOMIC MAITENANCE AND STABILITY

Elizabeth J. Sacho

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry**.**

Chapel Hill
2007

Approved by:

Dr. Dorothy Erie

Dr. Linda Spremulli

Dr. Tom Kunkel

Dr. Matthew Redinbo

Dr. Gary Pielak

ABSTRACT

Elizabeth J. Sacho: Biophysical and Biochemical Investigations of Proteins Involved in Genomic Stability
(Under the direction of Dr. Dorothy A. Erie)

The processes by which cells either make or avoid mutations in their DNA are varied and multifaceted. Often proteins involved in the repair of DNA damage are also found inducing mutations in the immune response. Despite the fact that these proteins and processes are found in all eukaryotes, the scientific community is just beginning to elucidate the mechanisms of these diverse pathways and the functions of the proteins involved in them. This dissertation looks at proteins in two seemingly different cellular processes: DNA mismatch repair and the immune response, as modulated by cytidine deaminases.

The DNA mismatch repair protein MutLα has been long thought to function as a molecular matchmaker; facilitating interactions between proteins in the mismatch repair pathway. Since little was known about this protein, biophysical and biochemical assays were designed and carried out to elucidate what role binding of ATP and DNA may have on MutLα. Surprisingly, MutLα can exist in four different conformational states, and these conformational states are governed by adenine nucleotide binding (but not hydrolysis).

It had been shown previously that MutLα binds to DNA in long, cooperative protein tracts. MutLα protein tracts do not appear to be effected by increasing salt concentration, but they are affected by the presence of ATP. Additionally, the ATP hydrolysis rate of full-length MutLα was measured, and found to be 2-6x larger than values previously reported for isolated N-terminal domains of MutLα, suggesting that conformational changes seen upon ATP binding, and the interactions that are likely made as a result of those conformational changes, may enhance the rate of ATP hydrolysis by MutLα.

Two of the cytidine deaminases (AID and APOBEC3G) have been shown to be involved in the immune response. The third cytidine deaminase (APOBEC2) is only a cytidine deaminase by sequence homology; and as yet has unknown function. Since it is widely believed that all cytidine deaminases are dimeric or tetrameric, a study of the oligomeric states of these proteins (with and without substrate present) was undertaken. Despite all three proteins belonging to the same superfamily, they each appear to exist in a different oligomeric state.

To Nana & Grampa, for everything.

ACKNOWLEDGEMENTS

I first off want to thank my advisor and mentor, Dorothy, for letting me work in her lab. Without her enthusiasm, encouragement and sheer scientific knowledge, this work would have never been possible. Dorothy allowed me to "test my wings" in her lab. While not everything may have worked, this experience and working with her has taught me many, many things, including perseverance, critical thinking and problem solving. I am very grateful for the education I have received under her direction.

I would like to thank my collaborator (and committee member) Tom Kunkel. He was kind enough to let me come and work in his lab in order to learn the finer points of expressing yeast proteins (only two, but it's a heck of a prep). This experience, and his insight into my project, has been invaluable to my growth as a scientist, and for that, I thank him.

I would like to thank my coworkers, past and present, in the Erie lab. Junghoon, Candi, Susan, Scott, Na, Cherie, Shannon, Brian and Bob have all contributed to my scientific education, and I am lucky to have met and grown along side every one of them. Special thanks must go out to Lauryn, Erika and Vanessa, for their submission to torture (reading thesis chapter drafts) as well as their willingness to always lend an ear. Special thanks also need to go out to some past members of the Erie lab: Ingrid, for her AFM expertise and friendship; Hong, for getting me started on this project and, along with Yong, training me on the AFM.

I would like to thank my collaborators and the other members of the biological division, for providing time, instrumentation, supplies and expertise. This includes Mercedes Arana, Shannon Holmes and other members of the Kunkel lab, Dr. Paul Modrich and Farid Kadyrov (Duke), Dr. Marilyn Diaz and Dave Brar (NIEHS), Dr. Myron Goodman and Linda Chelico (USC), Dr. Linda Spremulli and members of the Spremulli lab, Dr. Gary Pielak and members of the Pielak lab, Dr. Matthew Redinbo and members of the Redinbo lab, Dr. Nancy Thompson and the Thompson lab. Their support, along with the Department of Chemistry, has been invaluable.

I would also like to thank the many friends (too numerous to mention) who have become my extended family while I have been in NC. I am lucky to have you, and without your support and encouragement, this never would have been possible.

I would like to extend special thanks to my boyfriend, Robert Immormino. His love, support and encouragement of me during these five years, not to mention the many scientific conversations, kept me going long after I thought it was possible. This man is wonderful, and I don't know how this would have turned out if he were not in my life.

Finally, I need to thank my mom and grandparents. Their ceaseless love and encouragement and support got me to where I am today. Thank you.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

ABBREVIATIONS

1-ds binding	yMutLα binding on one strand of dsDNA
1-ds tract	yMutLα protein tract along one strand of dsDNA
2-ds binding	yMutLα binding on two strands of dsDNA
2-ds tract	yMutLα protein tract along two strands of dsDNA
3-D	Three-dimensional
A	Alanine
ADP	adenosine diphosphate
AFM	atomic force microscopy
AID	activation induced cytidine deaminase
Ala	Alanine
AMPPNP	5'-adenylyl-β-γ-imidodiphosphate
APOBEC	apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide like
APOBEC2	apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide like- 2
APOBEC3G	apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide like– 3G
ATP	adenosine triphosphate
ATPγS	adenosine 5'-(gamma-thio) triphosphate
bp	Base pair
C	Cysteine

C-	carboxy
C•C	Structred, dimerized C-terminal domains of yMutLα
C•C•N	Structure consisting of the dimerized C-terminal domains of yMutLα plus the linker arm and N-terminal domain of either Mlh1 or Pms1
C•C•N•N	Structure consisting of the dimerized C-terminal domains of yMutLα plus the linker arms and N-terminal domains of both Mlh1 and Pms1
CD	Circular dichroism
CD4+ T-cells	"helper" T-cells
Ci	Curie
CSR	Class Switch Recombination
D	Aspartic Acid
ddH_2O	doubly deionized water
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic acid
dsDNA	double stranded DNA
DTT	dithiothreitol
E	Glutamate
E. coli	*Escherichia coli*
EDTA	Ethylamine diamine tetraacetic acid
ExoI	Exonuclease I
G	Glycine

g	gram
Glu	Glutamate
GST	Glutathione-S-transferase
H	Histidine
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HIV	Human Immunodeficiency Virus
HMGB1	high mobility group box 1
hMutLα-EA	Hydrolysis-deficient hMutLα mutant
HNPCC	hereditary non-polyposis colorectal cancer
hsp90	Family of 90 kDa heat shock proteins found in the cytoplasm & endoplasmic reticulum
IDL	insertion-deletion loop
KCl	Potassium chloride
λ	wavelength
L	Liter
Li	Lithium
μm	micron
μM	micromolar
M	molar
Mg	Magnesium
MgCl	Magnesium Chloride
MgOAc	Magnesium Acetate

min	minute
mL	milliliter
Mlh	MutL homolog
mM	millimolar
mmol	millimole
MMR	DNA mismatch repair
Msh	MutS homolog
MutLα	Mlh1-Pms2 (yPms1)
MutSα	Msh2-Msh6
MutSβ	Msh2-Msh3
MW	Molecular weight
N	Asparagine
N	Structured N-terminal domain of either Mlh1 or Pms1
N-	amino
Na	Sodium
NaOAc	Sodium Acetate
NaOH	Sodium hydroxide
nM	nanomolar
nm	nanometer
nm^3	Cubic nanometers
nt	nucleotide
OD_{595}	Optical density at λ=595
OH^-	base

P_i	Inorganic phosphate
P	Proline
PCNA	proliferating cell nuclear antigen
PKA	Protein Kinase A
Pms	Post-meiotic segregation
PMSF	Phenylmethylsufonylfluoride
Polymin P	Poly(ethylenimine)
RFC	Replication Factor C
RNA	ribonucleic acid
RPA	Replication protein A
rpm	Revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
SHM	Somatic Hypermutation
ssDNA	single stranded DNA
T	Threonine
T	Thymine
TBE	Tris-borate-EDTA buffer
TLC	Thin layer chromatography
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
U	uracil
UDG	Uracil deglycosylase
V	Valine

Vif	Viral Infectivity factor
W	Adenine or Thymine
X	Any amino acid
Y	Pyrimidine (C or T)

CHAPTER 1

INTRODUCTION

The stability of an organism's genome is challenged every day by both internal and external forces. Consequently, a multitude of processes have evolved to deal with the inevitable genomic insults. Some of these genomic changes, for example the introduction of the wrong nucleotide into a growing DNA strand, can be deleterious to the organism. Other changes, such as the mutations of DNA in the process of adaptive immunity, are advantageous to the organism.

Mutations in genomic DNA, depending on the frame of reference, can be either a good or bad thing. A good thing, in the case of voluntary mutations caused by proteins in the process of passive and adaptive immunity, and a bad thing, in the case of involuntary mutations caused by internal and external sources. This dissertation focuses on characterizing the biophysical and biochemical properties of proteins involved in both the repair of mutations via the DNA mismatch repair system and the induction of mutations (for the purpose of adaptive immunity) by the cytidine deaminases that make up the AID/APOBEC family.

DNA Mismatch Repair

DNA mismatch repair (MMR) is one of several DNA repair processes that are vital for maintaining genome stability. In DNA synthesis, a mismatch (non-Watson-

Crick base pairing) occurs once in every 10^6-10^7 base pairs. The MMR system has been shown to post-replicatively improve the fidelity of DNA synthesis by 100-1000 fold (Bellacosa 2001). The MMR machinery is able to correct not only mismatches but also insertion-deletion loops (IDLs), which may occur during replication and recombination of DNA.

Mutations or deficiencies in the MMR system have been linked to an increased incidence of DNA mutations which can lead to cancer (Bellacosa 2001). Specifically, mutations in the DNA MMR proteins have been linked to hereditary non-polyposis colorectal cancer (HNPCC) (Liu, Yan et al. 2001; Hoffmann, Shcherbakova et al. 2003; Li 2003). HNPCC is characterized by an early onset of colorectal cancer as well as an increased frequency of other cancers, including cancers of the endometrium, stomach, ovary, pancreas, and urinary tract (Liu, Yan et al. 2001).

Mismatch Repair in E. coli

The process of MMR is best understood in *Escherichia coli* (*E. coli*), where it is carried out by the mismatch repair proteins MutS, MutL and MutH alongside the cellular replication machinery (Modrich and Lahue 1996). DNA repair via the MMR pathway (see Fig. 1.1) is initiated when MutS binds to the mismatch. MutL then binds to the MutS-mismatch complex in an ATP-dependant manner. Next, the MutS-MutL-mismatch complex activates a MutH bound at a hemi-methylated GATC site. The hemi-methylation of the DNA serves as the discrimination signal between the parent (methylated) and daughter (un-methylated) strand. MutH, a latent endonuclease, then cuts the daughter strand, making a nick that is accessible to a DNA helicases (UvrD) and an exonuclease. The excision of the mismatch starts at the nick and continues ~100 base

pairs past the mismatch. The replicative E. coli DNA polymerase (Pol III) then resynthesizes the DNA, and DNA ligase seals the nick (Au, Welsh et al. 1992; Modrich and Lahue 1996; Kunkel and Erie 2005; Iyer, Pluciennik et al. 2006).

The first two proteins involved in MMR in *E. coli*, MutS and MutL, are conserved in higher eukaryotes. In *E.coli* and other prokaryotes, MutS and MutL function as homodimers. However, in eukaryotes, the MutL and MutS homodimers are replaced by an array of heterodimers. A comparison between *E. coli*, *S. cerevisiae*, and human MutS and MutL proteins is found in Table 1.1.

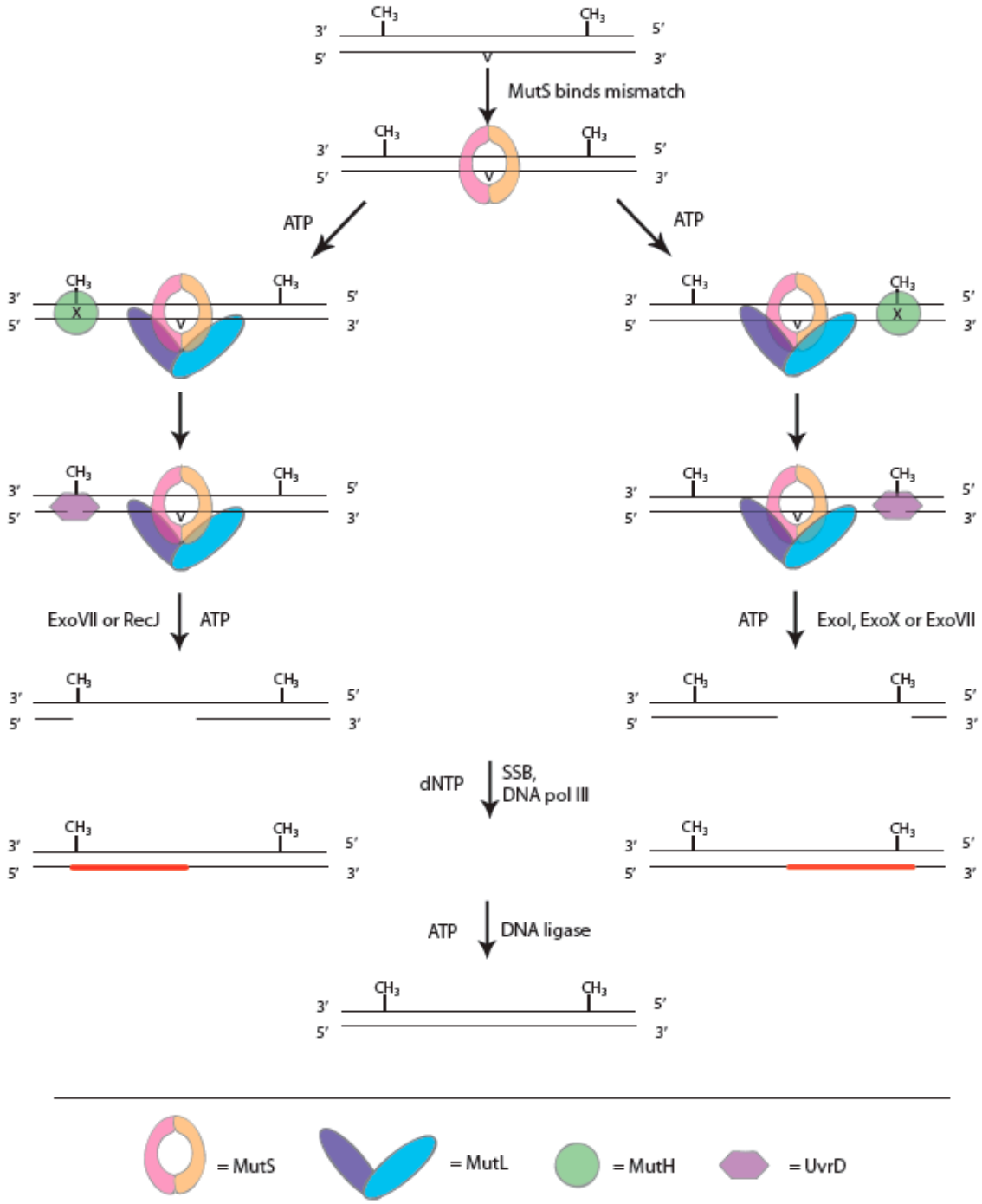

CH3
3′
5′
MutS binds mismatch
ATP
ATP
ExoVII or RecJ
ATP
ATP
ExoI, ExoX or ExoVII
dNTP
SSB,
DNA pol III
ATP
DNA ligase
= MutS
= MutL
= MutH
= UvrD

Figure 1.1

Figure 1.1: Schematic of E. coli MMR.

MMR can occur from both 5' to the mismatch (repair pathway on the left) or 3' to the mismatch (repair pathway on the right). DNA repair is initiated by the binding of MutS to the mismatch. MutL then binds to the MutS-mismatch complex in an ATP-dependant manner. Next, the MutS-MutL-mismatch complex activates a MutH bound at a hemi-methylated GATC site. MutH, a latent endonuclease, then cuts the daughter strand, making a nick that is accessible to a DNA helicase (UvrD) and an exonuclease. The excision of the mismatch starts at the nick, and continues ~100 base pairs past the mismatch. The replicative E. coli DNA polymerase (Pol III) then resynthesizes the DNA, and DNA ligase seals the nick

E. Coli	***S. cerevisiae***	***Human***
MutS	yMsh1	?
"	yMutSα (yMsh2-Msh6)	hMutSα (hMsh2-Msh6)
"	yMutSβ (yMsh2-Msh3)	hMutSβ (hMsh2-Msh3)
"	yMsh4	hMsh4
"	yMsh5	hMsh5
MutL	yMutLα (yMlh1-Pms1)	hMutlα (hMlh1-Pms2)
"	yMutLβ (yMlh1-Mlh2)	hMlhβ (hMlh1-Pms1)
"	yMlh3	hMlh3

Table 1.1: Comparison between E. coli, S. cerevisiae *and Human MMR proteins*

Eukaryotic Mismatch Repair

While the *E. coli* MMR system is well understood, it is becoming apparent that it may not be representative of MMR in eukaryotes. For example, there are more proteins involved in DNA replication in eukaryotes, there is no MutH homolog in eukaryotes, and the strand discrimination signal is unknown. Additionally, in eukaryotes the proteins MutS and MutL exist as the heterodimeric proteins MutSα and MutLα (Table 1.1). MutSα has been shown to be asymmetric in its binding to ATP as well as its interactions with DNA mismatches (Dufner, Marra et al. 2000; Drotschmann, Hall et al. 2002; Drotschmann, Yang et al. 2002; Warren, Pohlhaus et al. 2007), and MutLα has been shown to undergo asymmetric conformational changes in response to ATP binding (Hall, Shcherbakova et al. 2002; Tomer, Buermeyer et al. 2002). The MMR system in eukaryotes has been reconstituted from human proteins and has been found to require MutSα or MutSβ, MutLα (but only for repair in the 3' direction), ExoI, PCNA, HMGB1, RFC, RPA and DNA Polymerase δ (Constantin, Dzantiev et al. 2005; Zhang, Yuan et al. 2005). Because of its role in initiating MMR, many biochemical, biophysical and structural studies have been conducted on MutSα. Fewer of those types of studies have been conducted on MutLα.

MutLa

Until recently, most of what was known about MutLα was based on similarities of MutLα to *E. coli* MutL. It has been known that the ATPase activity of MutLα is necessary for MMR (Tran and Liskay 2000; Hall, Shcherbakova et al. 2002; Raschle, Dufner et al. 2002; Tomer, Buermeyer et al. 2002), and that MutLα is required for DNA

repair (Kramer, Kramer et al. 1989; Prolla, Christie et al. 1994), but until recently there wasn't any known function for MutLα. Kadyrov *et al* recently discovered an endonucleolytic activity in MutLα that is dependant on the presence of a mismatch, MutSα, PCNA, RFC and ATP (Kadyrov, Dzantiev et al. 2006). The authors also found that the endonucleolytic activity of MutLα, believed to be located in the C-terminal domain of Pms2 (yPms1), is dependant upon the enzyme being proficient for ATP hydrolysis (Kadyrov, Dzantiev et al. 2006).

Domain Organization of MutLα

Both proteins that make up MutLα have a similar organizational scheme (Fig 1.2). The C-termini of both proteins in MutLα are responsible for dimerization (Pang, Prolla et al. 1997), and the C-terminus of Pms1 (hPms2) contains the metal binding residues required for endonucleolytic activity (Kadyrov, Dzantiev et al. 2006). The N-terminus of each protein contains an ATPase functionality as well as residues important to DNA binding (Tran and Liskay 2000; Hall, Wang et al. 2001; Drotschmann, Hall et al. 2002; Hall, Shcherbakova et al. 2002; Raschle, Dufner et al. 2002; Tomer, Buermeyer et al. 2002). Dimerization via the C-termini is not enough to facilitate MMR (Pang, Prolla et al. 1997); the ATPase activity is required for the mismatch repair action of the protein (Hall, Shcherbakova et al. 2003) as well as the endonucleolytic function (Kadyrov, Dzantiev et al. 2006). The N- and C- terminal domains are connected by what is predicted to be a ~300 amino acid long flexible linker region (Jones 1999; McGuffin, Bryson et al. 2000; Bryson, McGuffin et al. 2005).

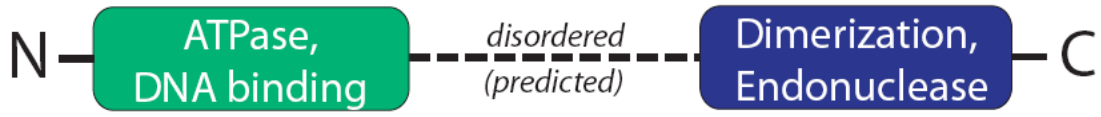


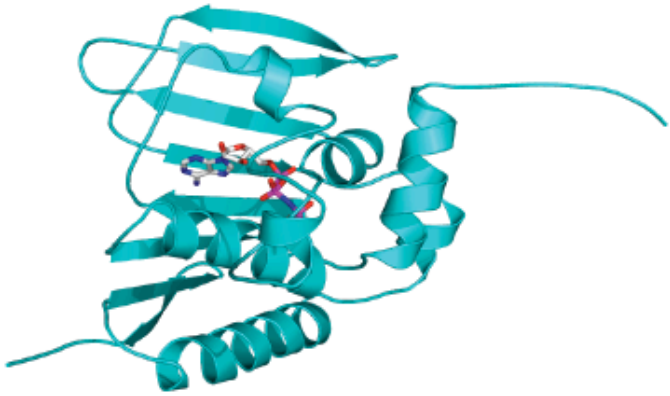
Figure 1.2: General organization of MutLα.

N- and C-terminal domains are predicted to have structure (represented by rounded rectangles), while the region connecting the two domains is predicted to be disordered. The N- terminal domain contains an ATPase activity, and is also able to bind DNA. The C-terminal domain contains the residues required for dimerization, as well as the metal binding residues (in hPms2 or yPms1 only) required for the endonucleolytic activity.

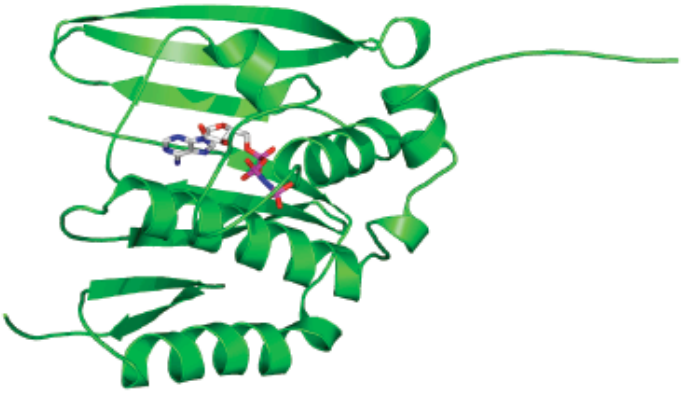
For *E. coli* MutL, there are multiple crystal structures of N-terminal domain, and one of the C-terminal domain. However, there is only one published crystal structure of a eukaryotic MutL homolog. The N-terminal domain of hPms2 (homologous to yPms1) has been crystallized and solved with ADP, ATPγS and no nucleotide bound (Guarne, Junop et al. 2001). The sequence and structural homology between MutL homologs (and even between eukaryotic and prokaryotic homologs) is primarily in their N-termini, which in all cases binds adenosine nucleotide using a set of four elements known as the Bergerat fold (Bergerat, de Massy et al. 1997; Guarne, Junop et al. 2001). The Bergerat fold is a non-traditional ATP binding domain, unique to the GHL protein family (of which MutL and homologs are members) (Ban and Yang 1998; Ban, Junop et al. 1999; Dutta and Inouye 2000; Guarne, Junop et al. 2001; Hu, Machius et al. 2003).

MutLα and the GHL ATPase family

In addition to MutL and MutL homologs, the GHL protein family includes two other namesakes, DNA Gyrase (and other type II topoisomerases, such as TopoVI) and Hsp90 (Figure 1.3) (Dutta and Inouye 2000). All members of the GHL ATPase family dimerize via their C-termini and bind ATP via the Bergerat fold, which is found in the N-termini (Figure 1.3).



MutL
(PDB id: 1B63)



type II topoisomerase
(Topo VI)
(PDB id: 1MX0)

Hsp90 homolog
(yHsp82)
(PDB id: 2CG9)

Figure 1.3

Figure 1.3: Comparison of ATP binding pockets for 3 GHL ATPase family members.

Crystal structures of the N-terminal domains of three GHL family members: MutL, Topo VI and yHsp82. Each protein has ADP or non-hydrolyzable ATP analog bound. Notice how the overall structures are similar: β-sheets form the "back wall" of the ATP binding pocket; loops and α-helices form the sides and lid

The Bergerat fold consists of four motifs, I, II, III and IV (Figure 1.4). Each motif contains a characteristic sequence: EXXXNXXD (Motif I), DXGXG (Motif II), GXXGXG/A (Motif III), and TX_nGT (Motif IV) (Dutta and Inouye 2000). The third motif, known as the "ATP lid", changes conformation upon ATP binding, generally folding over the nucleotide once it is in the binding pocket made by the other three motifs (Bergerat, de Massy et al. 1997; Ban and Yang 1998; Ban, Junop et al. 1999; Dutta and Inouye 2000; Prodromou, Panaretou et al. 2000; Guarne, Junop et al. 2001; Pearl and Prodromou 2001; Corbett and Berger 2003; Immormino, Dollins et al. 2004; Corbett and Berger 2005; Dollins, Immormino et al. 2005; Shiau, Harris et al. 2006). The ATP lid can move up to 20Å upon ATP binding from an open to a closed conformation (Figure 1.4, compare light purple or light green (- adenine nucleotide) to dark purple or dark green (+ adenine nucleotide)). In addition to Mg^{2+}, a monovalent cation is required for ATP lid stabilization and ATP hydrolysis; the exception being Li^{+}, which is believed to be too small to form the contacts necessary and ultimately inhibits ATP hydrolysis (Hu, Machius et al. 2003).

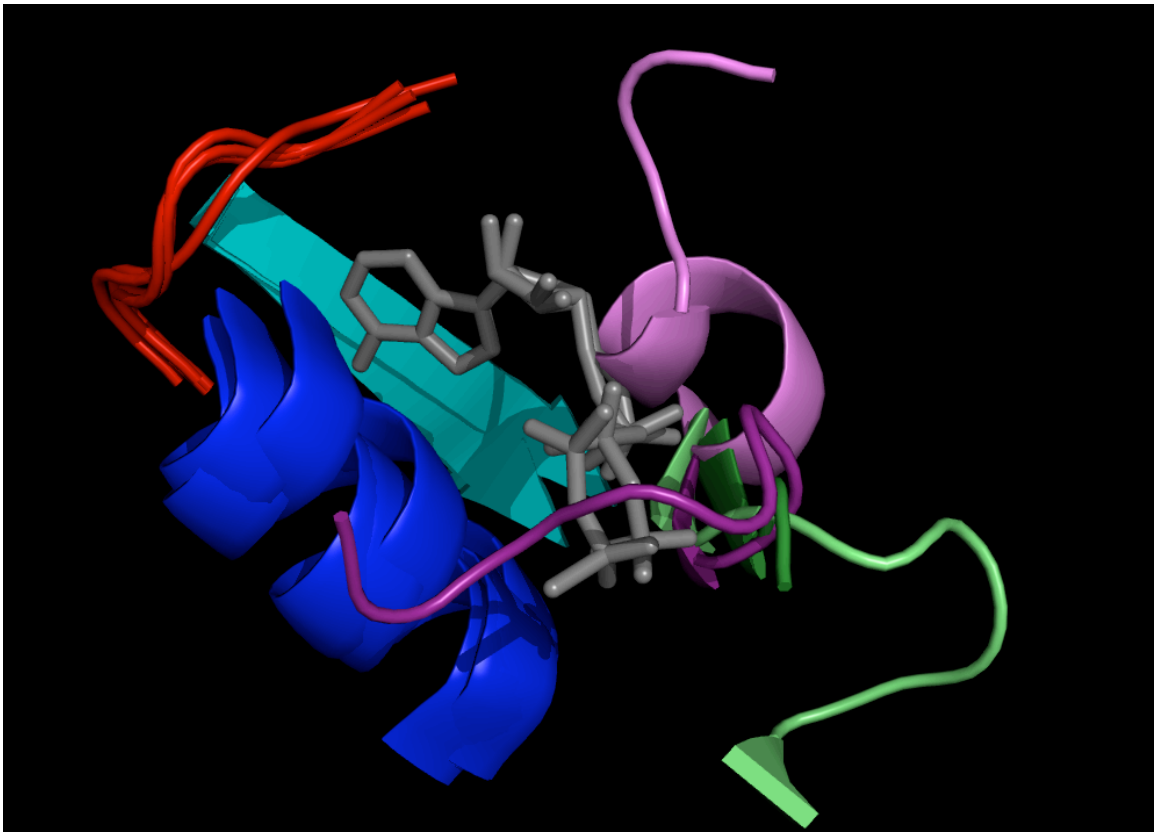


Figure 1.4: Close up of the four motifs that make up the ATP binding pocket (Bergerat fold)

The ATP binding pocket is shown for *E. coli* MutL & hPms2, each in the absence and presence of non-hydrolyzable ATP analog. Motifs I, II & IV (blue, red and cyan, respectively) overlay well. Motif III, the "ATP lid" changes conformation upon ATP binding in *E. coli* MutL (compare light purple to dark purple), and becomes disordered upon ATP binding in hPms2 (compare light green to dark green). PDB id's: 1B63, 1BKN, 1H7S, 1H7U

Members of the GHL ATPase family have slow rates of ATP hydrolysis in the absence of other cofactors (0.4-0.9 min^{-1}) (Ban, Junop et al. 1999; Dutta and Inouye 2000; Spampinato and Modrich 2000; Guarne, Junop et al. 2001; Hall, Shcherbakova et al. 2002), suggesting that ATP hydrolysis in GHL family members, unlike other ATPases, is used more for signaling or conformational changes.

Indeed, conformational changes in *E. coli* and yeast hsp90 proteins have been shown to occur as a result of ATP binding; these conformational changes are thought to be linked to the chaperoning of client proteins by the hsp90 family (Prodromou, Panaretou et al. 2000; Pearl and Prodromou 2001; Soti, Racz et al. 2002; Siligardi, Hu et al. 2004; Ali, Roe et al. 2006; Shiau, Harris et al. 2006). Similarly, the timing of the passage of one strand of DNA through another by the type II topoisomerases (Topo VI, for example) is regulated by ATP hydrolysis (Corbett and Berger 2005).

Cytidine Deaminases

Cytidine deaminases, as their name overtly suggests, deaminate cytidines to uracils. One group of cytosine deaminases is the AID/APOBEC (activation induced cytidine deaminase / apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide like) family. The majority of the members of this family are catalytically active on RNA, but a few members (APOBEC3G and the activation induced cytosine deaminase AID) have been shown to be active on ssDNA (Petersen-Mahrt, Harris et al. 2002; Dickerson, Market et al. 2003; Chelico, Pham et al. 2006).

The AID/APOBEC family of proteins, which is active on RNA and ssDNA, belongs to the cytidine deaminase superfamily. This superfamily contains a conserved zinc-binding signature motif, (H/C)(A/V)E(X_{24-30})(PCXXC), where X is any amino acid

(Navaratnam and Sarwar 2006); the cysteines and histidine coordinate a zinc ion, while the glutamate is involved in proton shuttling (Navaratnam and Sarwar 2006). The ancestral members of this superfamily deaminate free bases (as opposed to those found in RNA or DNA sequences) and are either dimeric or tetrameric (Navaratnam and Sarwar 2006). Dimerization is often used in the cytidine deaminase superfamily to form the active site (Figure 1.5A) (Lau, Zhu et al. 1994; Xie, Sowden et al. 2004). In contrast to other members of the superfamily, there is little known about the oligomeric state of AID, APOBEC2 and APOBEC3G, or how oligomerization affects function and cellular interactions.

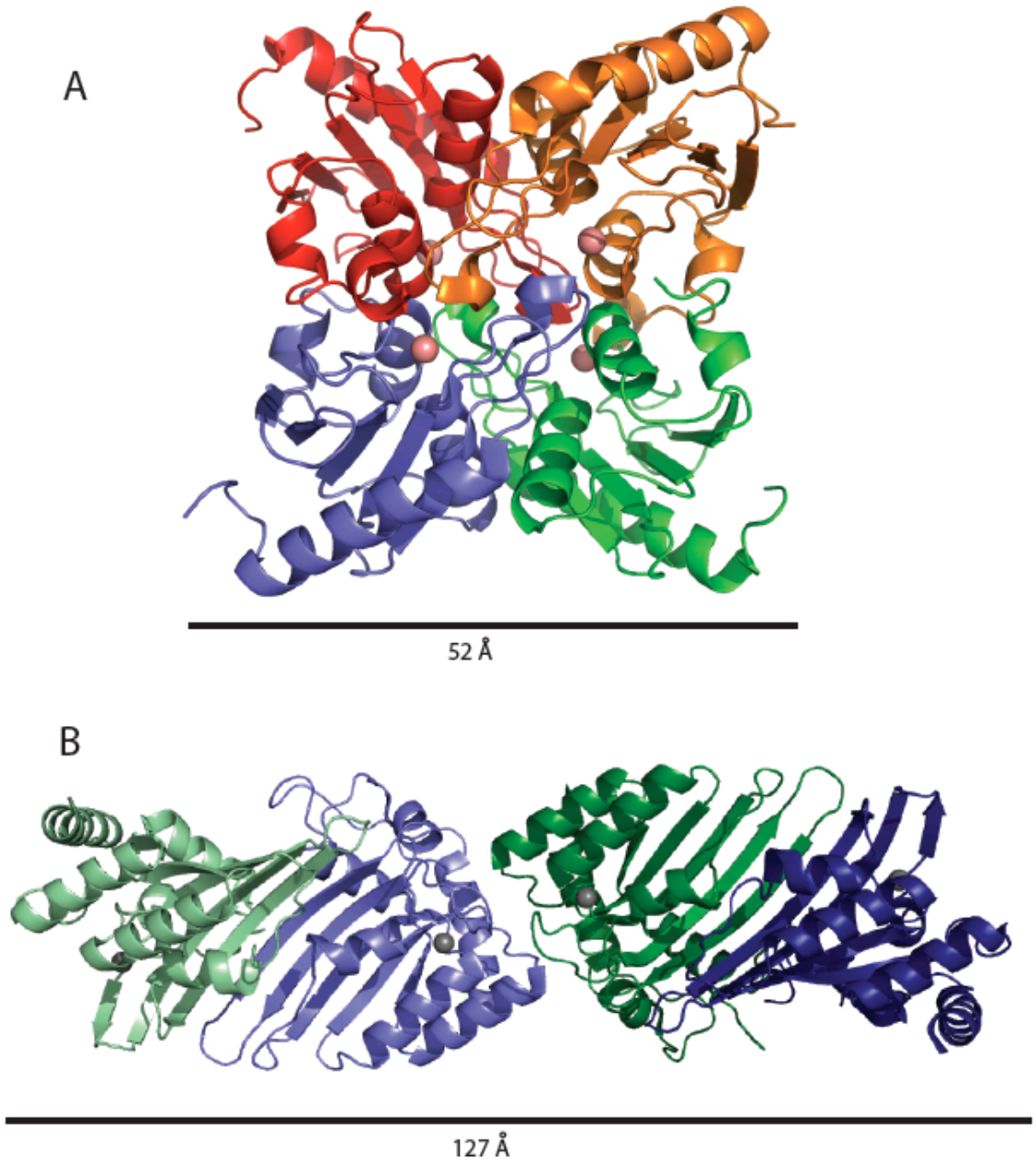


Figure 1.5

Figure 1.5: Two possible structures for proteins in the AID/APOBEC family.

(A) Structure of the yeast protein CDD1 (PDB id: 1R5T), which deaminates free cytosines in solution. Chains are shown in red, orange, green and blue. Zinc ions (found in the active site) are represented by pink spheres. (B) Structure of APOBEC2 (PDB id: 2NYT), which has no known function. This dimer of dimers is shown in blues and greens, with zinc ions represented by grey spheres. (Figure adapted from (Prochnow, Bransteitter et al. 2007))

AID

The activation induced cytidine deaminase AID is required for class switch recombination (CSR) and somatic hypermutation (SHM) (Muramatsu, Kinoshita et al. 2000; Revy, Muto et al. 2000; Arakawa, Hauschild et al. 2002; Harris, Sale et al. 2002; Yang, Obiakor et al. 2005). CSR and SHM are mechanisms that lead to the formation of high affinity antibodies characteristic of the secondary immune response. AID deaminates cytidines, generating uracils in DNA that encodes the variable and the switch regions of the immunoglobulin locus (Petersen-Mahrt, Harris et al. 2002; Rada, Williams et al. 2002).

In vitro, AID deaminates single-stranded but not double-stranded DNA (Bransteitter, Pham et al. 2003; Dickerson, Market et al. 2003; Pham, Bransteitter et al. 2003; Sohail, Klapacz et al. 2003). It has been shown that AID will function on dsDNA, but only if it is actively undergoing transcription, and especially if AID has been phosphorylated by PKA and is in complex with RPA (Chaudhuri, Khuong et al. 2004; Basu, Chaudhuri et al. 2005; Besmer, Market et al. 2006).

AID preferentially deaminates microsequences associated with SHM hotspots (WGYW, where W = A or T, Y = Pyrimidine) (Bransteitter, Pham et al. 2003; Beale, Petersen-Mahrt et al. 2004; Bransteitter, Pham et al. 2004; Rogozin and Diaz 2004; Yu, Huang et al. 2004). Data from pull-down assays using tagged-AID proteins and from structural modeling based on the free cytidine deaminases suggest that AID may function as a dimer to deaminate DNA (Ta, Nagaoka et al. 2003; Xie, Sowden et al. 2004; Prochnow, Bransteitter et al. 2007). However, it remains unclear whether the active site of AID is at the dimer interface, or if AID monomers are catalytically active.

APOBEC2

APOBEC2 was found by sequence homology to other APOBEC proteins, and is found primarily in cardiac and skeletal muscle (Liao, Hong et al. 1999). Unlike the other members of the AID/APOBEC family, APOBEC2 contains no deamination activity and its function remains unknown. The crystal structure of APOBEC2 has been solved, and because of the 44.6% sequence homology between the crystallized APOBEC2 fragment and AID, this crystal structure is serving as a model for AID and other AID/APOBEC family members (Prochnow, Bransteitter et al. 2007; Zhang, Mangeat et al. 2007).

The crystal structure of APOBEC2 (Prochnow, Bransteitter et al. 2007) (PDB id: 2NYT) revealed that APOBEC2 exists as a dimer of dimers (Figure 1.5B). One dimer interface is made by the packing of two beta-sheets (one from each monomer); whereas the second dimerization interface, that of the tetramer, is weaker, and made by hydrophobic contacts, salt bridges and hydrogen bonding between loop regions

APOBEC3G

APOBEC3G, unlike APOBEC2, does have a known function and deamination activity. This deamination activity is involved in viral immunity (Sheehy, Gaddis et al. 2002; Nguyen, Gummuluru et al. 2007; Noguchi, Hiraga et al. 2007), with APOBEC3G deaminating cytidines in the viral transcript to eventually cause mutations. These mutations eventually lead to the degradation of viral DNA (Sheehy, Gaddis et al. 2002).

APOBEC3G interacts with the viral infectivity factor (Vif), an HIV protein, which targets APOBEC3G for proteosomal degradation. In cells lacking Vif, APOBEC3G is able to survive, be packaged into the virion and deaminate cytidines in

the viral DNA. APOBEC3G has been shown to act processively along ssDNA, with a preference for deaminating the third cytosine of -CCC- sequences (Chelico, Pham et al. 2006).

An enzymatically active, low molecular weight form of APOBEC3G exists in resting CD4+ T cells, which are resistant to HIV infection, while an enzymatically inactive, higher molecular weight form exists in activated CD4+ T cells, which happen to be susceptible to HIV infection (Chiu, Soros et al. 2005). Treatment of the high molecular weight form of APOBEC3G with RNase converts it to the low molecular weight form (Chiu, Soros et al. 2005). These results suggest that the oligomeric state of APOBEC3G, and possibly other AID/APOBEC family members, may be related not only to deamination activity but also to interactions with other proteins.

Atomic Force Microscopy

Atomic force microscopy (AFM), also known as scanning force microscopy, was developed in 1986 from scanning tunneling microscopy. The advent of AFM enabled microscopic imaging of nonconductive, soft and live biological samples (Bustamante 1995; Fotiadis, Scheuring et al. 2002). With AFM, it is possible to view the topographical details on surfaces from the sub-molecular to the cellular level. AFM images can be collected in air and in solution, and for this reason, enables the study of dynamic processes of single molecules at physiologically relevant conditions. Because of this wide range of applications, AFM is well suited to the imaging of biomolecules, biomolecular complexes, organelles and cells.

AFM is a useful technique for the imaging of proteins, protein-protein and protein-DNA complexes. Because of the deposition process (Figure 1.6) AFM images

collected in air provide a “snapshot” of the dynamic processes that are going on in the solution the sample was taken from. A small volume of a sample containing protein and/or DNA is deposited directly onto mica. At this point, the proteins and/or DNA that “stick” to the mica, by virtue of electrostatic interaction, will remain on the surface as the rest of the sample is washed away and the surface is dried. In this process, the conformational state the proteins and/or DNA happened to be in upon deposition is the conformational state that they will remain in as the sample is imaged.

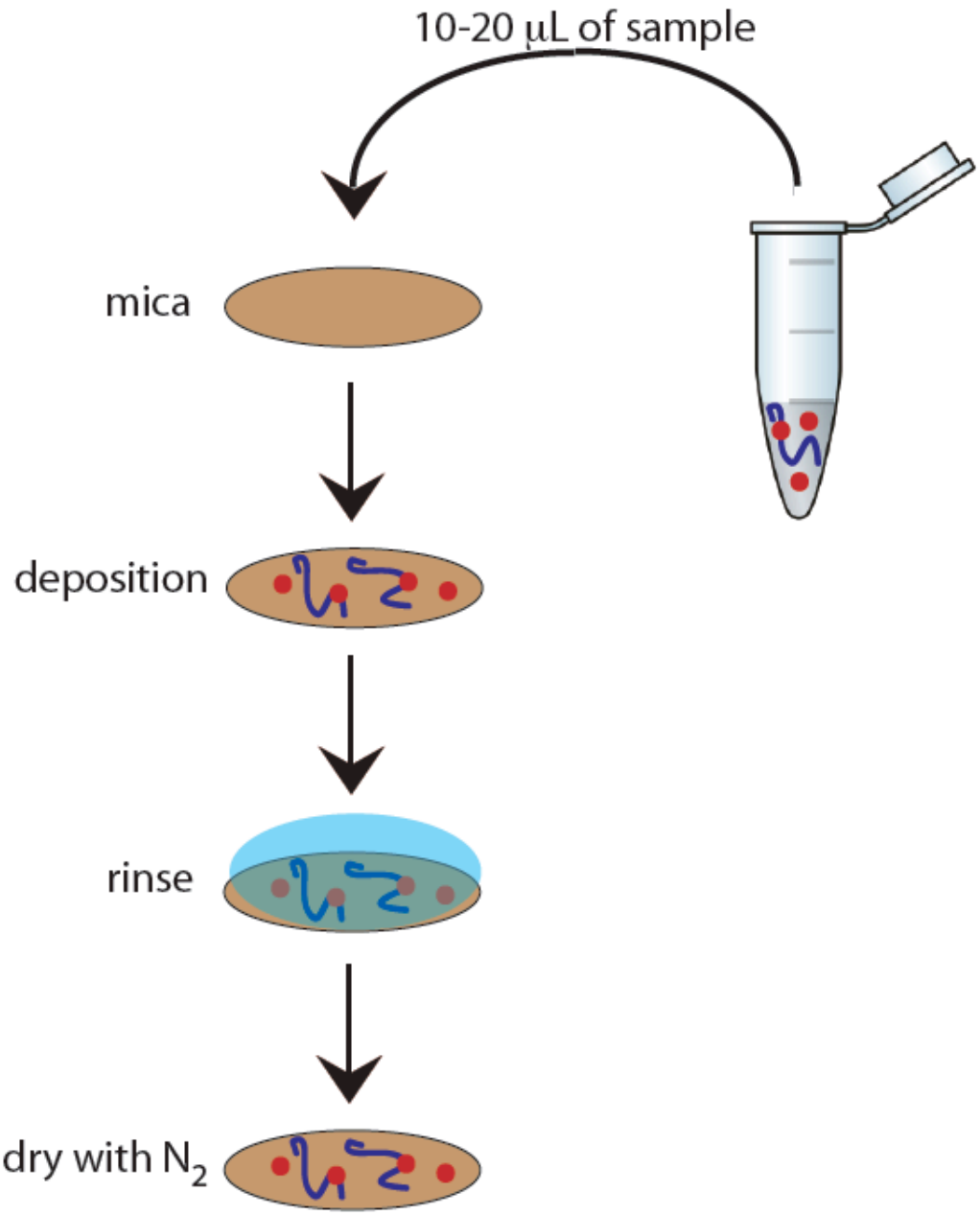


Figure 1.6

Figure 1.6: Schematic of the deposition process for imaging a sample with AFM.

A small volume of the sample (10-20 μL) is deposited directly onto freshly cleaved mica. The mica is then rinsed with ddH_2O and dried with a gentle stream of nitrogen before being placed in the AFM for imaging.

Much information is contained in an AFM image. For example, it is possible to estimate the oligomerization state of a protein based on its volume as measured by AFM, since a linear relationship between a protein's molecular weight and its AFM volume has been shown (Ratcliff and Erie 2001; Yang, Wang et al. 2003). Furthermore, the conformational effects of a protein binding to DNA (bending, wrapping, kinking) (Yang, Wang et al. 2003), as well as conformational changes within the protein itself can be seen directly.

AFM images are generated when a sample of interest is scanned beneath a tip mounted on a cantilever (Figure 1.7A). Forces of interaction between the tip and the sample, generated as the sample is scanned, cause deflections in the cantilever (Figure 1.7B). These deflections are monitored by a laser beam, which is reflected off the back of the cantilever into a 4-segment photodiode. The photodiode is linked to the sample station (piezo) through a feedback loop, and the height of the piezo is adjusted to keep the clipping amplitude of the tip, monitored by the photodiode, constant (Figure 1.7C). The change in height of the piezo is then plotted as a function of the X-Y position of the tip to yield a topographical image of the sample (Hansma, Sinsheimer et al. 1993; Bustamante 1995). (Figure 1.7D)

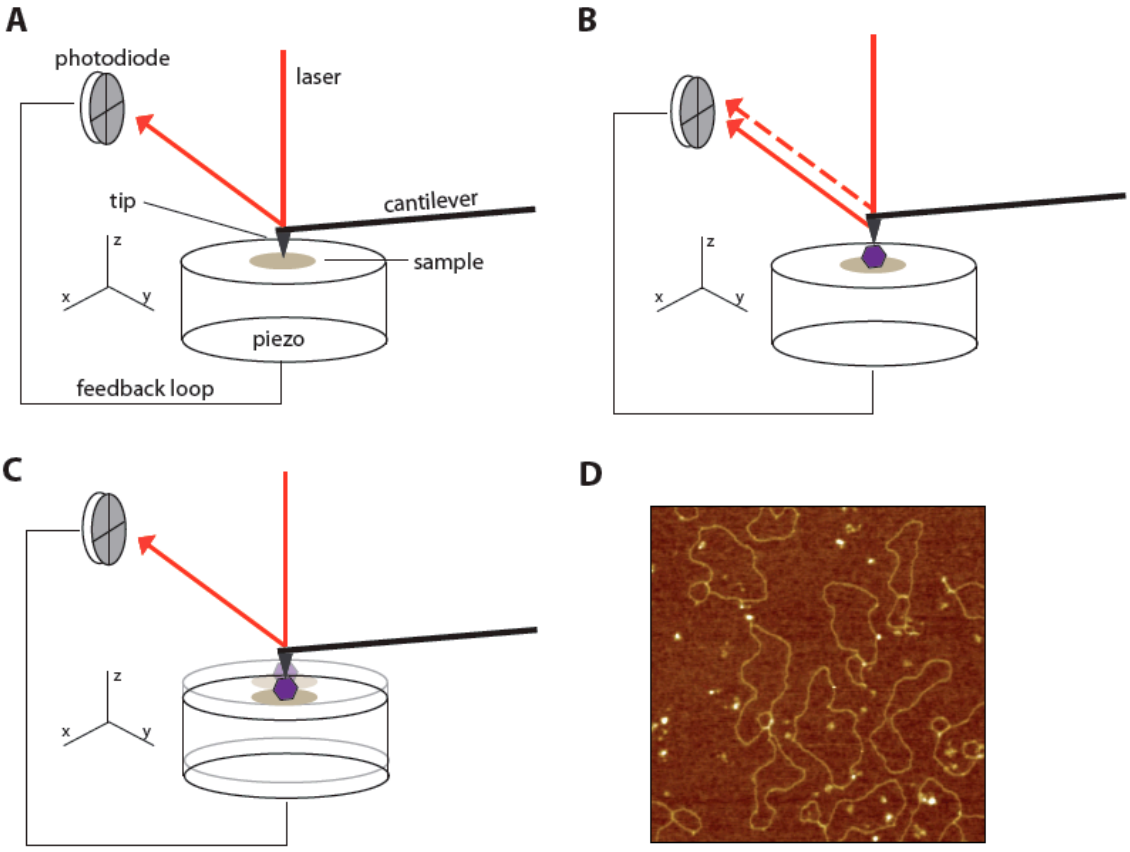

A
photodiode
laser
tip
cantilever
sample
piezo
feedback loop
z
x
y
B
C
D

Figure 1.7

Figure 1.7: Overview of how AFM works.

(A) The sample is placed on a piezo, which is scanned in the x-y direction by a tip attached to a cantilever. Laser light is reflected off of the back of the cantilever into a photodiode. Throughout the scanning process the position of the reflected laser light on the photodiode is kept constant by a feedback loop between the photodiode and the piezo. (B) As the tip encounters something on the surface (protein, DNA, etc), it moves over it, causing the reflection of the laser light to change (dashed line). (C) To compensate for this change in the reflection, the piezo is moved down (compare grey lines to black lines), and this movement in the z-direction, when plotted against the x-y position yields a tree dimensional image (D), represented here in 2 dimensions with the third dimension (z) being represented by color. This false color image (1 x 1 μm) has lower areas represented by darker colors, and higher areas represented by lighter colors

Dissertation Overview

In the body of this work, AFM was the primary technique used to study protein conformational changes, oligomeric states, and, protein-DNA interactions. In Chapter 2, the effect of adenine nucleotide on the conformation of MutLα is investigated. In addition to using AFM, these studies also required the use of partial trypsin proteolysis and circular dichroism (CD). The combination of these techniques reveals that MutLα can exist in one of four different conformational states, and ATP binding, but not hydrolysis mediates these conformational states. Chapter 3 is an investigation into the oligomeric state of three cytidine deaminases, AID, APOBEC2 and APOBEC3G. Despite all three proteins being cytidine deaminases, they all have different oligomeric states, which may reflect on the different nature of their cellular function. In Chapter 4, the effect of salt, ATP and the presence of a nick in DNA on the DNA binding and ATP hydrolysis of MutLα are investigated. The preliminary data presented in this chapter suggest that ATP has a significant effect on the binding of yMutLα to covalently closed circular DNA and that the full-length yMutLα hydrolyzes ATP at a rate faster than the N-terminal fragments of the protein. Appendix A contains specific details on AFM imaging and the purification of yMutLα.

References

Ali, M. M., S. M. Roe, et al. (2006). "Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex." Nature **440**(7087): 1013-7.

Arakawa, H., J. Hauschild, et al. (2002). "Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion." Science **295**(5558): 1301-6.

Au, K. G., K. Welsh, et al. (1992). "Initiation of methyl-directed mismatch repair." J Biol Chem **267**(17): 12142-8.

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

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

Basu, U., J. Chaudhuri, et al. (2005). "The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation." Nature **438**(7067): 508-11.

Beale, R. C., S. K. Petersen-Mahrt, et al. (2004). "Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo." J Mol Biol **337**(3): 585-96.

Bellacosa, A. (2001). "Functional interactions and signaling properties of mammalian DNA mismatch repair proteins." Cell Death Differ **8**(11): 1076-92.

Bergerat, A., B. de Massy, et al. (1997). "An atypical topoisomerase II from Archaea with implications for meiotic recombination." Nature **386**(6623): 414-7.

Besmer, E., E. Market, et al. (2006). "The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination." Mol Cell Biol **26**(11): 4378-85.

Bransteitter, R., P. Pham, et al. (2004). "Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase." J Biol Chem **279**(49): 51612-21.

Bransteitter, R., P. Pham, et al. (2003). "Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase." Proc Natl Acad Sci U S A **100**(7): 4102-7.

Bryson, K., L. J. McGuffin, et al. (2005). "Protein structure prediction servers at University College London." Nucleic Acids Res **33**(Web Server issue): W36-8.

Bustamante, C. a. K., D (1995). "Scanning Force Microscopy in Biology." Physics Today: 32-38.

Chaudhuri, J., C. Khuong, et al. (2004). "Replication protein A interacts with AID to promote deamination of somatic hypermutation targets." Nature **430**(7003): 992-8.

Chelico, L., P. Pham, et al. (2006). "APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA." Nat Struct Mol Biol **13**(5): 392-9.

Chiu, Y. L., V. B. Soros, et al. (2005). "Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells." Nature **435**(7038): 108-14.

Constantin, N., L. Dzantiev, et al. (2005). "Human mismatch repair: Reconstitution of a nick-directed bidirectional reaction." J Biol Chem.

Corbett, K. D. and J. M. Berger (2003). "Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution." Embo J **22**(1): 151-63.

Corbett, K. D. and J. M. Berger (2005). "Structural dissection of ATP turnover in the prototypical GHL ATPase TopoVI." Structure (Camb) **13**(6): 873-82.

Dickerson, S. K., E. Market, et al. (2003). "AID mediates hypermutation by deaminating single stranded DNA." J Exp Med **197**(10): 1291-6.

Dollins, D. E., R. M. Immormino, et al. (2005). "Structure of unliganded GRP94, the endoplasmic reticulum Hsp90. Basis for nucleotide-induced conformational change." J Biol Chem **280**(34): 30438-47.

Drotschmann, K., M. C. Hall, et al. (2002). "DNA binding properties of the yeast Msh2-Msh6 and Mlh1-Pms1 heterodimers." Biol Chem **383**(6): 969-75.

Drotschmann, K., W. Yang, et al. (2002). "Evidence for sequential action of two ATPase active sites in yeast Msh2-Msh6." DNA Repair (Amst) **1**(9): 743-53.

Dufner, P., G. Marra, et al. (2000). "Mismatch recognition and DNA-dependent stimulation of the ATPase activity of hMutSalpha is abolished by a single mutation in the hMSH6 subunit." J Biol Chem **275**(47): 36550-5.

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

Fotiadis, D., S. Scheuring, et al. (2002). "Imaging and manipulation of biological structures with the AFM." Micron **33**(4): 385-97.

Guarne, A., M. S. Junop, et al. (2001). "Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHL ATPase." Embo J **20**(19): 5521-31.

Hall, M. C., P. V. Shcherbakova, et al. (2003). "DNA binding by yeast Mlh1 and Pms1: implications for DNA mismatch repair." Nucleic Acids Res **31**(8): 2025-34.

Hall, M. C., P. V. Shcherbakova, et al. (2002). "Differential ATP binding and intrinsic ATP hydrolysis by amino-terminal domains of the yeast Mlh1 and Pms1 proteins." J Biol Chem **277**(5): 3673-9.

Hall, M. C., H. Wang, et al. (2001). "High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer." J Mol Biol **312**(4): 637-47.

Hansma, H. G., R. L. Sinsheimer, et al. (1993). "Recent advances in atomic force microscopy of DNA." Scanning **15**(5): 296-9.

Harris, R. S., J. E. Sale, et al. (2002). "AID is essential for immunoglobulin V gene conversion in a cultured B cell line." Curr Biol **12**(5): 435-8.

Hoffmann, E. R., P. V. Shcherbakova, et al. (2003). "MLH1 Mutations Differentially Affect Meiotic Functions in Saccharomyces cerevisiae." Genetics **163**(2): 515-26.

Hu, X., M. Machius, et al. (2003). "Monovalent cation dependence and preference of GHKL ATPases and kinases." FEBS Lett **544**(1-3): 268-73.

Immormino, R. M., D. E. Dollins, et al. (2004). "Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone." J Biol Chem **279**(44): 46162-71.

Iyer, R. R., A. Pluciennik, et al. (2006). "DNA mismatch repair: functions and mechanisms." Chem Rev **106**(2): 302-23.

Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.

Kadyrov, F. A., L. Dzantiev, et al. (2006). "Endonucleolytic function of MutLalpha in human mismatch repair." Cell **126**(2): 297-308.

Kramer, B., W. Kramer, et al. (1989). "Heteroduplex DNA correction in Saccharomyces cerevisiae is mismatch specific and requires functional PMS genes." Mol Cell Biol **9**(10): 4432-40.

Kunkel, T. A. and D. A. Erie (2005). "DNA mismatch repair." Annu Rev Biochem **74**: 681-710.

Lau, P. P., H. J. Zhu, et al. (1994). "Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene." Proc Natl Acad Sci U S A **91**(18): 8522-6.

Li, G. M. (2003). "DNA mismatch repair and cancer." Front Biosci **8**: d997-1017.

Liao, W., S. H. Hong, et al. (1999). "APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family." Biochem Biophys Res Commun **260**(2): 398-404.

Liu, T., H. Yan, et al. (2001). "The role of hPMS1 and hPMS2 in predisposing to colorectal cancer." Cancer Res **61**(21): 7798-802.

McGuffin, L. J., K. Bryson, et al. (2000). "The PSIPRED protein structure prediction server." Bioinformatics **16**(4): 404-5.

Modrich, P. and R. Lahue (1996). "Mismatch repair in replication fidelity, genetic recombination, and cancer biology." Annu Rev Biochem **65**: 101-33.

Muramatsu, M., K. Kinoshita, et al. (2000). "Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme." Cell **102**(5): 553-63.

Navaratnam, N. and R. Sarwar (2006). "An overview of cytidine deaminases." Int J Hematol **83**(3): 195-200.

Nguyen, D. H., S. Gummuluru, et al. (2007). "Deamination-independent inhibition of hepatitis B virus reverse transcription by APOBEC3G." J Virol **81**(9): 4465-72.

Noguchi, C., N. Hiraga, et al. (2007). "Dual effect of APOBEC3G on Hepatitis B virus." J Gen Virol **88**(Pt 2): 432-40.

Pang, Q., T. A. Prolla, et al. (1997). "Functional domains of the Saccharomyces cerevisiae Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations." Mol Cell Biol **17**(8): 4465-73.

Pearl, L. H. and C. Prodromou (2001). "Structure, function, and mechanism of the Hsp90 molecular chaperone." Adv Protein Chem **59**: 157-86.

Petersen-Mahrt, S. K., R. S. Harris, et al. (2002). "AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification." Nature **418**(6893): 99-103.

Pham, P., R. Bransteitter, et al. (2003). "Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation." Nature **424**(6944): 103-7.

Prochnow, C., R. Bransteitter, et al. (2007). "The APOBEC-2 crystal structure and functional implications for the deaminase AID." Nature **445**(7126): 447-51.

Prodromou, C., B. Panaretou, et al. (2000). "The ATPase cycle of Hsp90 drives a molecular 'clamp' via transient dimerization of the N-terminal domains." Embo J **19**(16): 4383-92.

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

Rada, C., G. T. Williams, et al. (2002). "Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice." Curr Biol **12**(20): 1748-55.

Raschle, M., P. Dufner, et al. (2002). "Mutations within the hMLH1 and hPMS2 subunits of the human MutLalpha mismatch repair factor affect its ATPase activity, but not its ability to interact with hMutSalpha." J Biol Chem **277**(24): 21810-20.

Ratcliff, G. C. and D. A. Erie (2001). "A novel single-molecule study to determine protein--protein association constants." J Am Chem Soc **123**(24): 5632-5.

Revy, P., T. Muto, et al. (2000). "Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2)." Cell **102**(5): 565-75.

Rogozin, I. B. and M. Diaz (2004). "Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process." J Immunol **172**(6): 3382-4.

Sheehy, A. M., N. C. Gaddis, et al. (2002). "Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein." Nature **418**(6898): 646-50.

Shiau, A. K., S. F. Harris, et al. (2006). "Structural Analysis of E. coli hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements." Cell **127**(2): 329-340.

Siligardi, G., B. Hu, et al. (2004). "Co-chaperone regulation of conformational switching in the Hsp90 ATPase cycle." J Biol Chem **279**(50): 51989-98.

Sohail, A., J. Klapacz, et al. (2003). "Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations." Nucleic Acids Res **31**(12): 2990-4.

Soti, C., A. Racz, et al. (2002). "A Nucleotide-dependent molecular switch controls ATP binding at the C-terminal domain of Hsp90. N-terminal nucleotide binding unmasks a C-terminal binding pocket." J Biol Chem **277**(9): 7066-75.

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

Ta, V. T., H. Nagaoka, et al. (2003). "AID mutant analyses indicate requirement for class-switch-specific cofactors." Nat Immunol **4**(9): 843-8.

Tomer, G., A. B. Buermeyer, et al. (2002). "Contribution of human mlh1 and pms2 ATPase activities to DNA mismatch repair." J Biol Chem **277**(24): 21801-9.

Tran, P. T. and R. M. Liskay (2000). "Functional studies on the candidate ATPase domains of Saccharomyces cerevisiae MutLalpha." Mol Cell Biol **20**(17): 6390-8.

Warren, J. J., T. J. Pohlhaus, et al. (2007). "Structure of the Human MutSalpha DNA Lesion Recognition Complex." Mol Cell **26**(4): 579-92.

Xie, K., M. P. Sowden, et al. (2004). "The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1." Proc Natl Acad Sci U S A **101**(21): 8114-9.

Yang, G., H. Obiakor, et al. (2005). "Activation-induced deaminase cloning, localization, and protein extraction from young VH-mutant rabbit appendix." Proc Natl Acad Sci U S A **102**(47): 17083-8.

Yang, Y., H. Wang, et al. (2003). "Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy." Methods **29**(2): 175-87.

Yu, K., F. T. Huang, et al. (2004). "DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine." J Biol Chem **279**(8): 6496-500.

Zhang, K. L., B. Mangeat, et al. (2007). "Model Structure of Human APOBEC3G." PLoS ONE **2**: e378.

Zhang, Y., F. Yuan, et al. (2005). "Reconstitution of 5'-directed human mismatch repair in a purified system." Cell **122**(5): 693-705.

CHAPTER 2

ASYMMETRIC ADENINE NUCLEOTIDE INDUCED CONFORMATIONAL CHANGES IN MUTLα

Introduction

Eukaryotes, as well as some prokaryotes, have no known MutH homolog, and the sequence of events from mismatch recognition to repair for these organisms is not completely understood. As such, the role of the eukaryotic MutL homolog, MutLα, is not fully known. MutLα is a heterodimer comprised of Mlh1 and Pms2 (Pms1 in yeast) and contains a latent endonuclease activity that is activated in the presence of a mismatch, MutSα, PCNA, RFC and ATP (Kadyrov, Dzantiev et al. 2006). In addition to this endonucleolytic activity, MutLα is able to suppress the activity of ExoI on homoduplex DNA (Jager, Rasmussen et al. 2001; Genschel and Modrich 2003; Nielsen, Jager et al. 2004), and it interacts with several other proteins in the MMR pathway (Jager, Rasmussen et al. 2001; Mendillo, Mazur et al. 2005; Kadyrov, Dzantiev et al. 2006; Lee and Alani 2006). It has been proposed that conformational changes in MutLα caused by ATP binding and/or hydrolysis modulate these varying protein interactions during the repair process (Guarne, Junop et al. 2001; Raschle, Dufner et al. 2002).

Like the homodimeric *E. coli* MutL (Guarne, Ramon-Maiques et al. 2004; Kosinski, Steindorf et al. 2005), each subunit in the heterodimeric MutLα contains a C-terminal dimerization domain (Pang, Prolla et al. 1997) connected to an N-terminal ATPase domain (Pang, Prolla et al. 1997; Tran and Liskay 2000) via a linker region that is predicted to be disordered. The ATPase activity of MutL and MutLα is required for MMR (Pang, Prolla et al. 1997; Spampinato and Modrich 2000; Tran and Liskay 2000; Hall, Shcherbakova et al. 2002; Raschle, Dufner et al. 2002) and is also required for the endonuclease activity in hMutLα (Kadyrov, Dzantiev et al. 2006).

MutL and MutL homologs are members of the GHL ATPase family (Ban and Yang 1998; Ban, Junop et al. 1999; Dutta and Inouye 2000; Guarne, Junop et al. 2001; Hu, Machius et al. 2003), that is characterized by a non-traditional ATP binding fold (Bergerat, de Massy et al. 1997). Other members of this family include the namesakes; DNA Gyrase and Hsp90 as well as Grp94 and the type II topoisomerases. MutL and other GHL family members have been shown to have slow ($0.4\ min^{-1} – 0.9\ min^{-1}$) rates of ATP hydrolysis in the absence of other cofactors (Ban, Junop et al. 1999; Dutta and Inouye 2000; Spampinato and Modrich 2000). In all proteins in the GHL family, ATP binding and/or hydrolysis appears to induce large conformational changes which are purported to be involved in the signaling of cellular processes (Ban, Junop et al. 1999; Dutta and Inouye 2000; Corbett and Berger 2003; Immormino, Dollins et al. 2004; Corbett and Berger 2005; Dollins, Immormino et al. 2005; Ali, Roe et al. 2006; Chu, Maynard et al. 2006; Shiau, Harris et al. 2006).

The conformational changes observed in *E. coli* MutL in response to adenine nucleotide have been previously explored indirectly by size-exclusion chromatography

and directly by crystallographic structures of the isolated N-terminal domain. Size-exclusion chromatography has shown that the full length *E. coli* MutL adopts a more compact size in the presence of a non-hydrolyzable ATP analog, AMPPNP, which the authors attribute to N-terminal dimerization (Ban, Junop et al. 1999). Similarly, crystal structures of the N-terminal domain of *E. coli* MutL show that AMPPNP binding results in the dimerization of the N-terminal domains, which is distinct from the monomeric crystal structure of the apo enzyme (Ban, Junop et al. 1999). The crystal structure of the AMPPNP bound form of *E. coli* MutL also shows that upon AMPPNP binding, the ATP lid of the Bergerat fold folds over and makes contacts with the AMPPNP.

In stark contrast to *E. coli* MutL, the N-terminal fragment of hPms2, one of two subunits in eukaryotic MutLα, is a monomer in the crystal structure, even in the presence of an ATP analog. Unlike *E.coli* MutL, the N-terminal fragment is hydrolytically proficient, and dimerization appears to not be a requirement for ATP hydrolysis (Guarne, Junop et al. 2001). Additionally, the ATP lid of the Bergeral fold actually becomes more disordered, and fewer residues of this lid are seen in the ATPγS structure than in the apo structure. To date, there are no data on the full-length structures of either MutL or MutLα.

In this chapter an analysis of the full-length structure of MutLα at the level of individual molecules is presented. Using AFM coupled with biochemical analyses, the nucleotide-induced changes in the quaternary structure of human and yeast MutLα are investigated. AFM images reveal that MutLα can adopt four distinct conformations. Based on the observed conformational changes, a model for how ATP binding and hydrolysis drives conformation changes in MutLα is proposed. These observed

conformational changes are likely coupled to the signaling of downstream events in mismatch repair.

Results

MutLα Exists in Four Different Conformational States

Inspection of AFM images such as those in Figure 2.1A reveals that yMutLα can adopt four different conformational states. In the absence of additional nucleotide cofactors, MutLα exists primarily in an open and "v-shaped" extended conformation, in which a large compact central domain is connected to two smaller domains by flexible arms (Figure 2.1B and Figure 2.1F). Because it is known that yMlh1 and yPms1 form the heterodimeric complex via their C-termini (Pang, Prolla et al. 1997), the larger structure found at the vertex of the V is undoubtedly the dimerized C-termini, while the two domains located at the ends of the elongated linker arms are the N-termini.

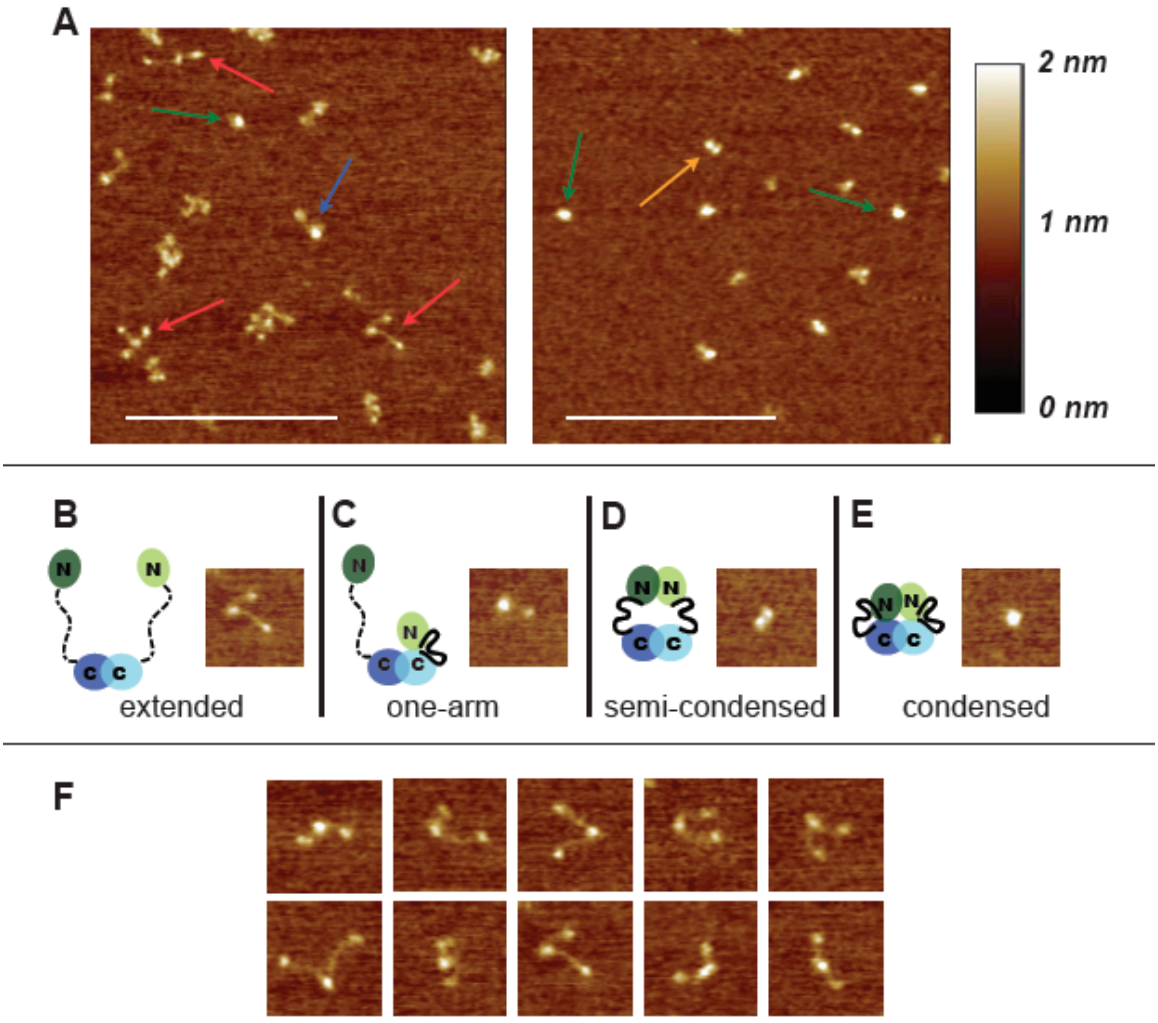

A
2 nm
1 nm
0 nm
B
C
D
E
N
C
extended
one-arm
semi-condensed
condensed
F

Figure 2.1

Figure 2.1: AFM images of yMutLα.

(A) 500 nm x 500 nm images of MutLα deposited in the absence of nucleotide (left) and in the presence of 5 mM ADP (right). Arrows point to examples of the four different conformational states seen in the images: red = extended; blue = one armed; orange = semi-condensed; green = condensed. Scale bar (white) is 250 nm. (B-E) 100 nm x 100 nm images of the four states accompanied by cartoons of MutLα in the different conformational states. In the cartoons, domains are indicated by ovals, connected by a flexible linker. Disordered linker is shown as a dashed line; ordered linker is shown as a heavy line. Mlh1 is shown in light green and blue, and Pms2 (yPms1) is shown in dark green and blue. (F) a selection of 100 nm x 100 nm images of the extended state to show the flexibility of the linker arms

A statistical analysis of these extended conformations revealed that one arm is approximately 40% longer than the other and a broad distribution of angles between the two arms (Figure 2.2 and Table 2.1). These results are consistent with secondary structure predictions made using PsiPred and PONDR (Jones 1999; McGuffin, Bryson et al. 2000; Bryson, McGuffin et al. 2005), which suggest that the N- and C- terminal domains of yMlh1 and yPms1 have well defined secondary structures that are connected by disordered regions of ~160 and ~300 amino acids, respectively. The flexible linker arms have a smaller cross-section than the coiled-coil domains of Rad50 as visualized by AFM (van Noort, van Der Heijden et al. 2003), further reinforcing the idea that the arms that connect the N- and C-termini lack secondary structure. Finally, the volume of proteins in AFM images has been shown to depend linearly on molecular weight (Ratcliff and Erie 2001; Yang, Wang et al. 2003; Yang, Sass et al. 2005), and the measured volumes of the three observed domains are consistent with the predicted molecular weights of the structured N-termini and dimerized C-terminal domains (Tables 2.2 and 2.3).

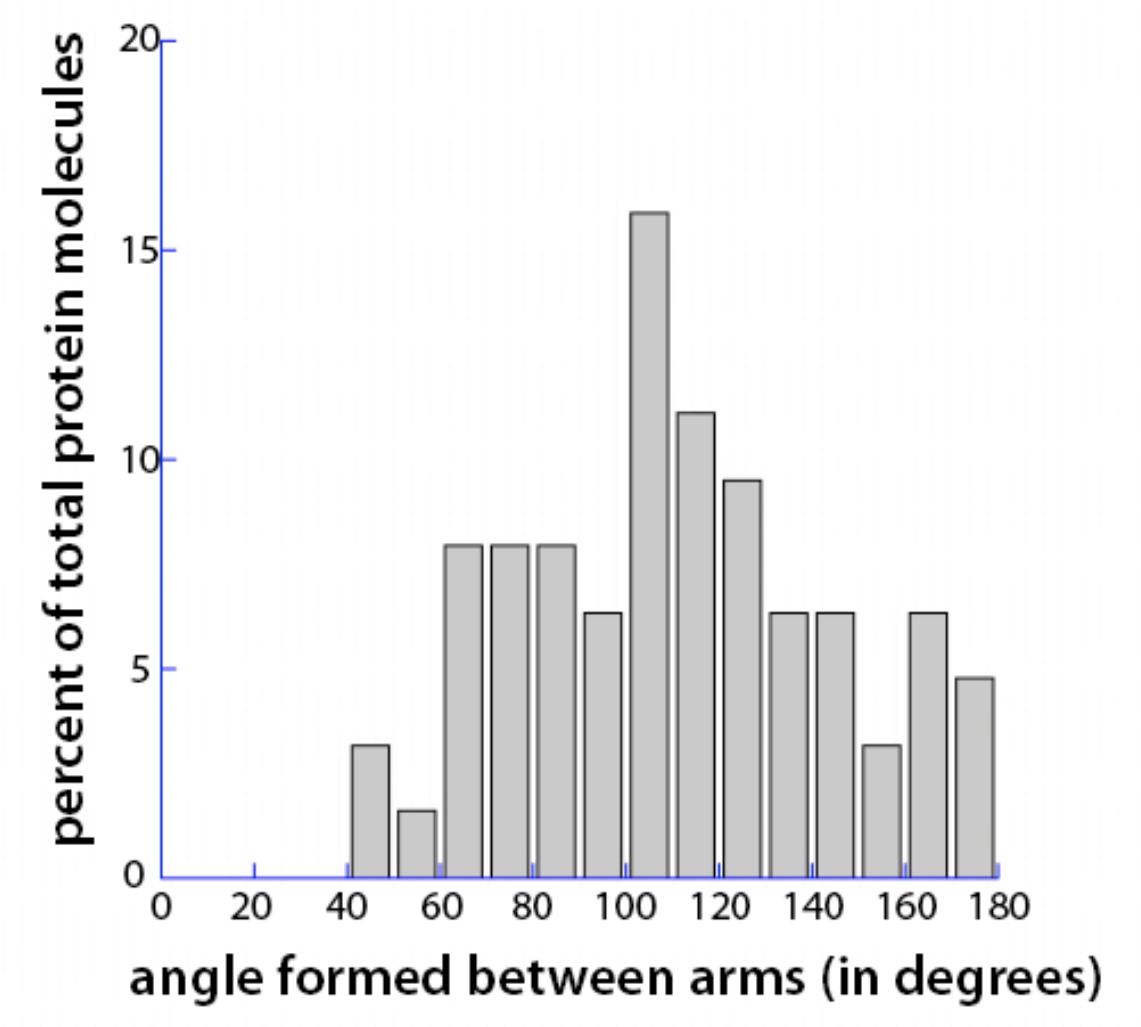


Figure 2.2: Distribution of angles formed between the two arms of MutLα in the extended conformation

Short Arm Length	Long Arm Length	Angle between the Two
21.68 ± 6.44 nm	30.00 ± 8.05 nm	110.75 ± 33.93 °

Table 2.1: Average arm length and angles of MutLα in the extended conformation. Lengths and angles were measured by hand (n=63). Lengths were measured from the edge of one domain to the edge of the other domain. Angles were measured by estimating a vertex in the center of the dimerized C-termini and then drawing lines from the vertex parallel with where the arm exited the dimerized C-termini.

	Predicted
C•C	56 kDa 52.5 nm^3
N_{Mlh1}	37 kDa 29.7 nm^3
N_{Pms1}	40 kDa 33.3 nm^3
C•C•N_{Mlh1}	112 kDa 119.7 nm^3
C•C•N_{Pms1}	130 kDa 141.3 nm^3
C•C•N•N (full length protein)	186 kDa 208.5 nm^3

Table 2.2: Predicted molecular weights & volumes for domains in yMutLα.

Molecular weights are based on PSIPRED & PONDR predictions for domain size. For the C•C•N structures, the linker belonging to the condensed N terminal domain is included in the predicted MW & volume calculations. For the C•C•N•N structure, both linker arms are included in the calculation.

	v-shape (n=10)	One-arm (n=11)	Semi-condensed & condensed (n=24)
C•C	90.9 ± 15.4 nm^3		
N	42.9 ± 11.6 nm^3	43.2 ± 12.3 nm^3	
C•C•N		137.7 ± 25.5 nm^3	
C•C•N•N			213.7 ± 42.4 nm^3

Table 2.3: Actual measured volumes (in nm^3) of C•C, N, C•C•N and C•C•N•N structures.

In addition to the extended conformation, three other minor populations of conformations are seen in the AFM images (Figures 2.1 & 2.3). Approximately 20% of the molecules exhibit a "one armed" conformation, with a large domain connected to a smaller one by a flexible linker arm (Figure 2.1C). Measurement of the volume of the domains (Ratcliff and Erie 2001; Yang, Wang et al. 2003; Yang, Sass et al. 2005) reveals that the large structure in this state is larger than the central structure in the extended conformation (Table 2.3). These results suggest that in these one-armed conformations, one of the N-terminal domains is either interacting directly with the dimerized C-terminal region and/or that one linker arm has folded ("condensed") to create a globular domain consisting of the two C-termini, one linker arm, and one N-terminus. Roughly 10% of the molecules are in a semi-condensed state, with two domains similar in size (Figure 2.1D). In this state, the linker arms are not visible and the centers of the two domains are separated by 15 ± 4 nm. These results suggest that in this state, the two N-terminal domains have associated with one another but not with the C-terminal domains. The semi-condensed state is distinguished from the one-armed state by the size of the protein domains and their relative orientations to each other. Finally, ~ 15% of the molecules exhibit a compact structure with no protrusions (condensed state, Figure 2.1E). In this condensed state, the N- and C-terminal domains of both proteins in the heterodimer appear to be interacting to form a single compact molecule with no visible separate domains or linker arms visible.

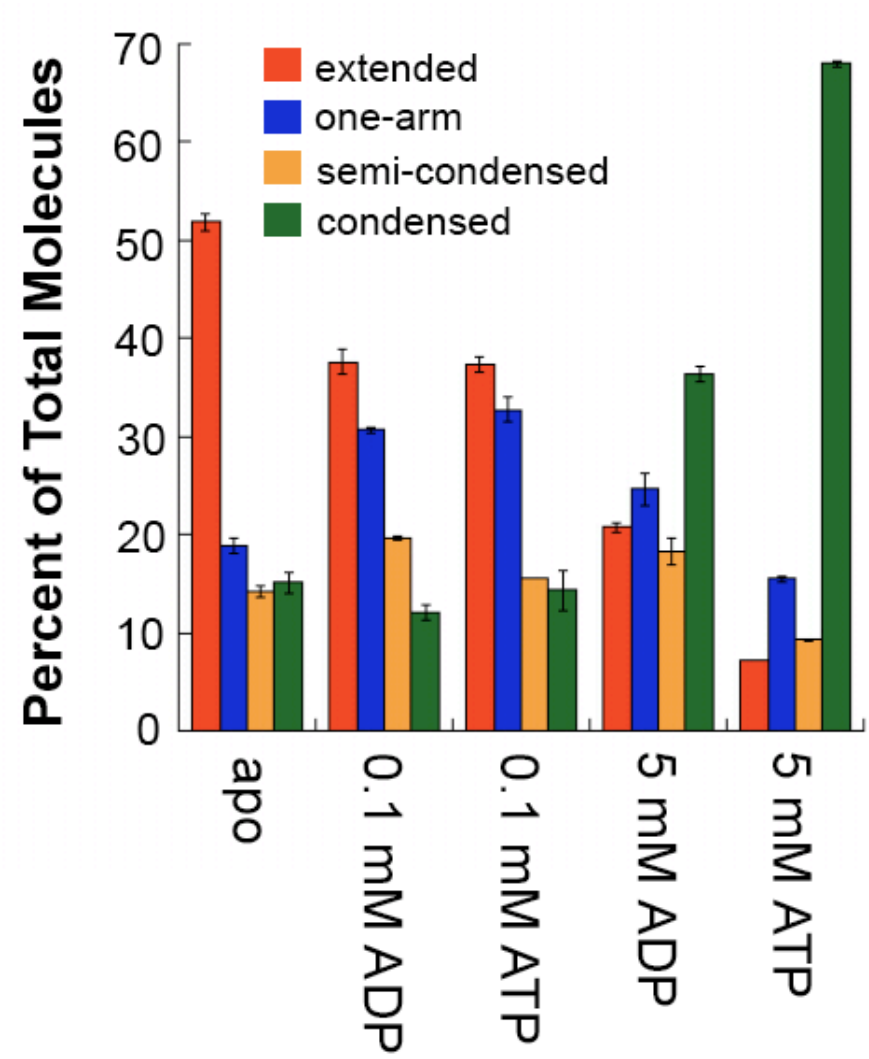


Figure 2.3: Distribution of yMutLα molecules among each of the four states under varying conditions.

apo (n=727), 0.1 mM ADP (n=912), 0.1 mM ATP (n=391), 5 mM ADP (n=559), and 5 mM ATP (n=1021). Red = extended; Blue = one-arm; Orange = semi-condensed; Green = condensed. Data were tallied by hand from 1 x 1 μm AFM images taken from multiple experiments.

Adenine nucleotides induce large conformational changes in MutLα

Genetic studies have shown that mutations that impair the binding or hydrolysis of ATP in either yPms1 or yMlh1 result in defective MMR *in vivo* (Tran and Liskay 2000; Hall, Shcherbakova et al. 2002). In addition, biochemical studies suggest that ATP binding and hydrolysis induce conformational changes in MutL homologs (Ban and Yang 1998; Tran and Liskay 2000; Hall, Shcherbakova et al. 2002; Raschle, Dufner et al. 2002). To assess the role of adenine nucleotides in controlling the conformations of MutLα, we imaged MutLα under a variety of nucleotide conditions (Figure 2.3). Based on previous studies of the ATPase activities of the N-terminal domains of yMlh1 and yPms1 (Hall, Shcherbakova et al. 2002), we chose conditions in which we expected either only the ATPase site of Mlh1 to be occupied (0.1 mM) or both sites to be occupied (5 mM).

Addition of either 0.1 mM ATP or ADP results in a decrease in the population of molecules in the extended conformation and increase in the population of the one-armed state (Figure 2.3). These results suggest that binding of ADP or ATP promotes a conformation of MutLα in which the N-terminal domain and linker arm of Mlh1 or Pms1 is interacting with or folded onto the C-terminal dimerization domains of Mlh1 and/or Pms1. Because Mlh1 appears to have a higher affinity for ATP (Hall, Shcherbakova et al. 2002), it is likely that these conformational changes are occurring in Mlh1 rather than Pms1.

At high concentrations of ATP (5 mM) where both yMlh1 and yPms1 should have nucleotide bound, MutLα exists predominately in the condensed conformation (70%) (Figure 2.3). Similar results were obtained with 1 mM ATP (data not shown).

These results indicate that binding and/or hydrolysis of ATP drives the formation of this highly compact state in which the N-termini are interacting with one another and with the linker arms and/or the C-terminal dimerization domains. Interestingly, high concentrations of ADP (5 mM) also cause a significant increase in the condensed state (40% versus 15% without nucleotide). These results suggest that occupancy of both of the ATPase sites with either ATP or ADP facilitates the formation of the condensed state, although ATP is more efficient than ADP.

Proteolysis Protection reveals ATP-induced structural changes in solution

To further assess these conformational changes, we used partial proteolysis to probe MutLα conformations in solution (Figure 2.4). Consistent with our AFM results and previous experiments using hMutLα (Raschle, Dufner et al. 2002), yMutLα is readily digested by trypsin in the absence of nucleotide cofactor. The addition of 5 mM ATP, and to a lesser extent 0.1 mM ATP, results in the protection of MutLα (especially Mlh1) from protease digestion (Figure 2.4). These results indicate that ATP induces a conformational change that masks the trypsin protease sites and supports the idea that the conformational changes observed by AFM also occur in solution.

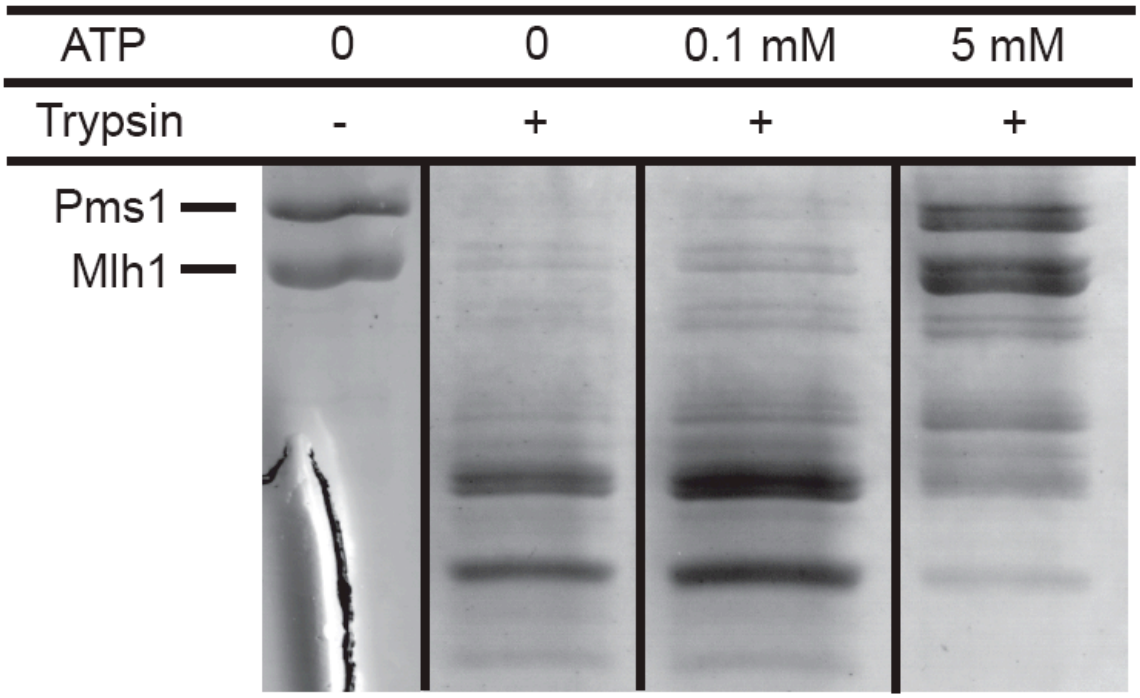


Figure 2.4: Partial trypsin proteolysis of yMutLα.

From left to right: undigested yMutLα, digested yMutLα, digested yMutLα + 0.1 mM ATP, digested yMutLα + 5 mM ATP

ATP binding causes an increase in secondary structure

Although the AFM data reveal that ATP can induce dramatic changes in the conformation of MutLα and the partial proteolysis data suggest conformational changes, neither technique provides direct information about secondary structure formation. To determine if ATP binding promotes the formation of structure in the flexible arms, we examined secondary structure formation using circular dichroism spectroscopy (CD). Due to the high extinction coefficient of ATP, we were able to only obtain CD spectra at 0.1 mM ATP (but not at 5 mM ATP). At this concentration of ATP, there is a significant increase in the one-armed state in the AFM images and a slight increase in the amount of protein protected from trypsin digestion. Comparison of the CD spectra in the absence and presence of 0.1 mM ATP (Figure 2.5) shows that the addition of 0.1 mM ATP causes a significant increase in the CD signal in the region centered at 225 nm. This negative peak, centered at 225 nm, is characteristic of α-helices and/or β-sheet secondary structures in proteins (Kelly, Jess et al. 2005). There is also a positive peak, centered at 260 nm, that is consistent with a rigid environment for aromatic amino acid residues (Kelly, Jess et al. 2005). These data, taken together with our AFM data, and partial proteolysis results, suggest that binding of ATP to one subunit of MutLα (probably Mlh1) induces the formation of secondary structure, likely in the flexible linker arm, and promoting the interaction between a linker arm and the N- and C- termini.

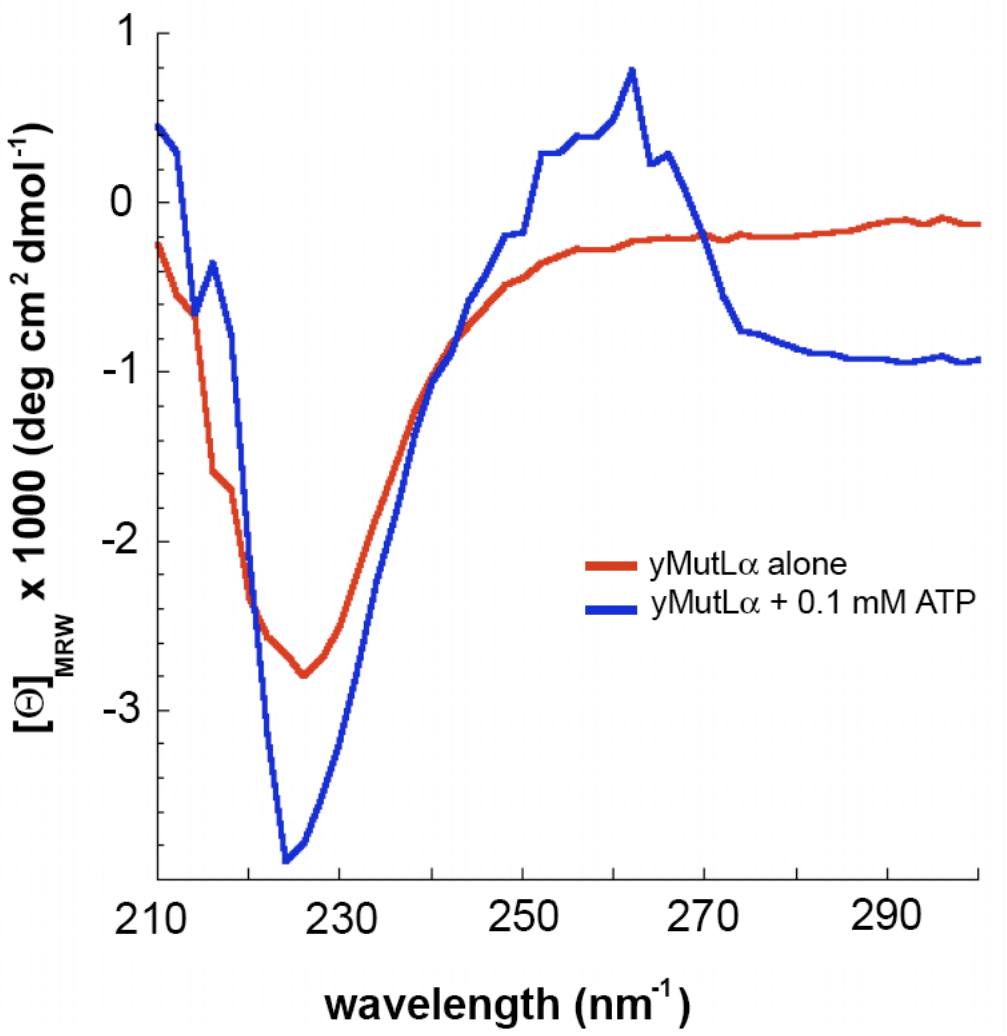


Figure 2.5: CD spectra of 0.4 μM yMutLα in the absence (red) and presence (blue) of 0.1 mM ATP.

The signal differences seen at 220-230 nm demonstrate that in the presence of ATP, the protein adopts more secondary structure (either α-helix or β-sheet). Plotted curves are the average of 3 experiments, with the background spectra (either imaging buffer alone or imaging buffer + 0.1 mM ATP) subtracted.

Human MutLα displays similar conformational changes

To determine if the results with yeast MutLα are general, we examined the conformations of human MutLα (hMlh1-hPms2) alone, with 0.1 mM ATP and with 5 mM ATP. As expected, we observe the same four states for hMutLα that we see for yMutLα. In addition, the distributions of the states under the same conditions are similar for hMutLα and yMutLα (Figure 2.6). Specifically, in the absence of ATP, both yeast and human MutLα exist predominantly in the extended state with defined N- and C-termini and disordered linker arms; whereas, in the presence of 5 mM ATP, both human and yeast MutLα are primarily condensed. There are some differences in the distributions in the presence of 0.1 mM ATP, which may result from differences in the ATP binding affinity of the yeast and human proteins (Figure 2.6), since very little is known about the binding affinity of any full-length eukaryotic MutLα.

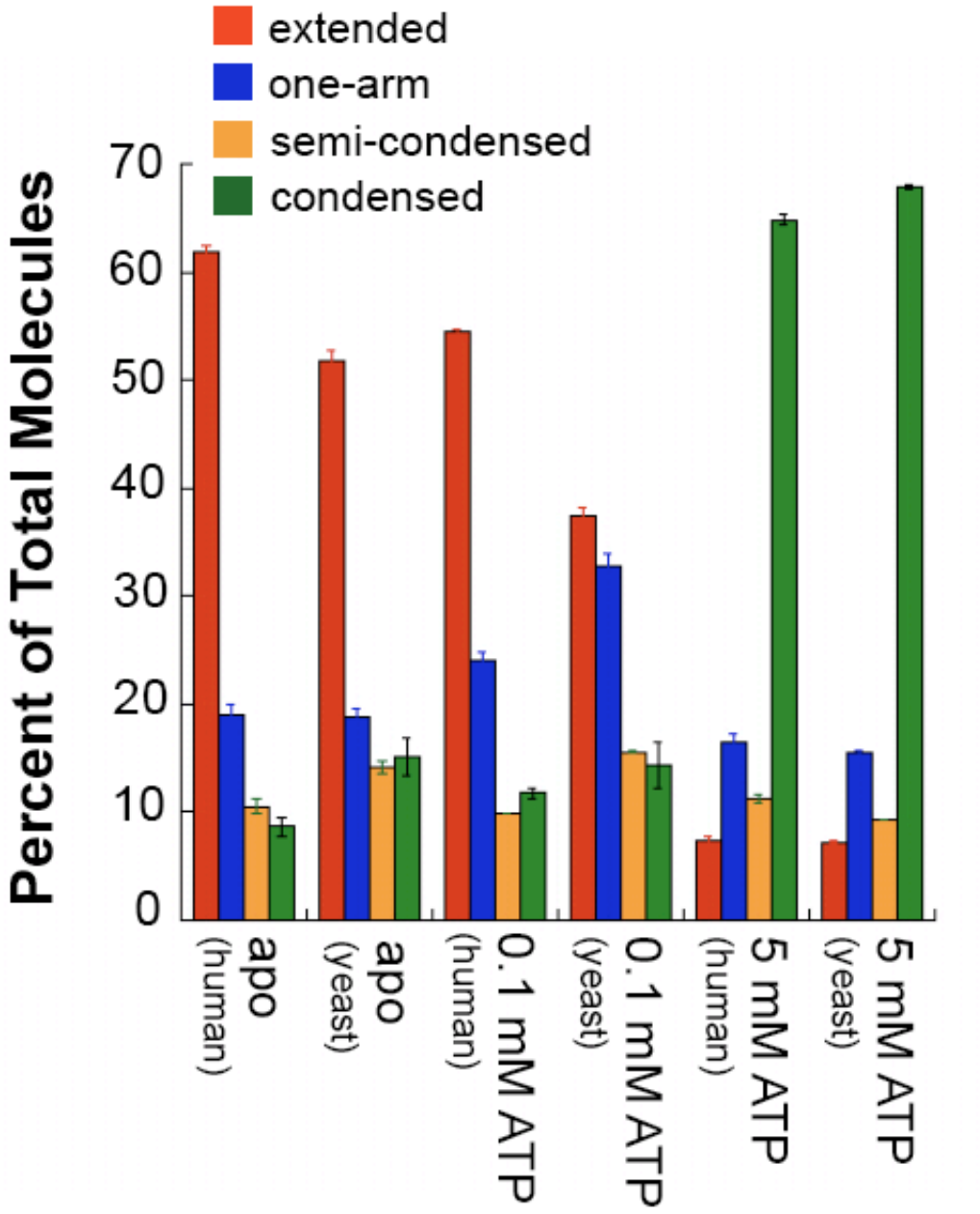

extended
one-arm
semi-condensed
condensed
Percent of Total Molecules
70
60
50
40
30
20
10
0
apo (human)
apo (yeast)
0.1 mM ATP (human)
0.1 mM ATP (yeast)
5 mM ATP (human)
5 mM ATP (yeast)

Figure 2.6

Figure 2.6: Comparison of the distributions of conformational states of hMutLα and yMutLα molecules.

hMutLα apo (n=1253), hMutLα + 0.1 mM ATP (n=979), hMutLα + 5 mM ATP (n=302). Red = extended; blue = one-arm; orange = semi-condensed; green = condensed. Data were tallied by hand from 1 x 1 μm AFM images taken from multiple experiments.

Hydrolysis is not required for the formation of the condensed state

To determine whether or not hydrolysis of ATP is required for the formation of the condensed state, we examined the conformational changes of a mutant of hMutLα in which the hydrolysis activity of both subunits is severely impaired (hMlh1•E34A-Pms2•E41A (hMutLα-EA)). In this mutant, the conserved glutamate residues in the active sites of both subunits have been changed to alanines, inhibiting hydrolysis of ATP (Raschle, Dufner et al. 2002; Tomer, Buermeyer et al. 2002). Images of hMutLα-EA in the absence of adenine nucleotide appear similar to images of the wild-type human and yeast MutLα, with the extended conformation dominating the population significantly (Figure 2.7A). In addition, like wild-type MutLα, high concentrations of ATP (5mM) cause hMutLα-EA to condense (Figure 2.7B and 2.7C).

The hMutLα-EA mutant is still capable limited ATP hydrolysis, with a rate of ~0.2 min^{-1} (Raschle et al., 2002; F.K and P.M. unpublished results). Because MutLα and ATP are incubated together for less than a minute prior to being deposited on the surface, less than 20% of the hMutLα-EA molecules will have hydrolyzed ATP at the time of deposition. Given that ~70% of the molecules are in the condensed state for both the wild-type hMutLα and hMutLα-EA, it is highly unlikely that hydrolysis is required for formation of the condensed state (Figure 2.7C).

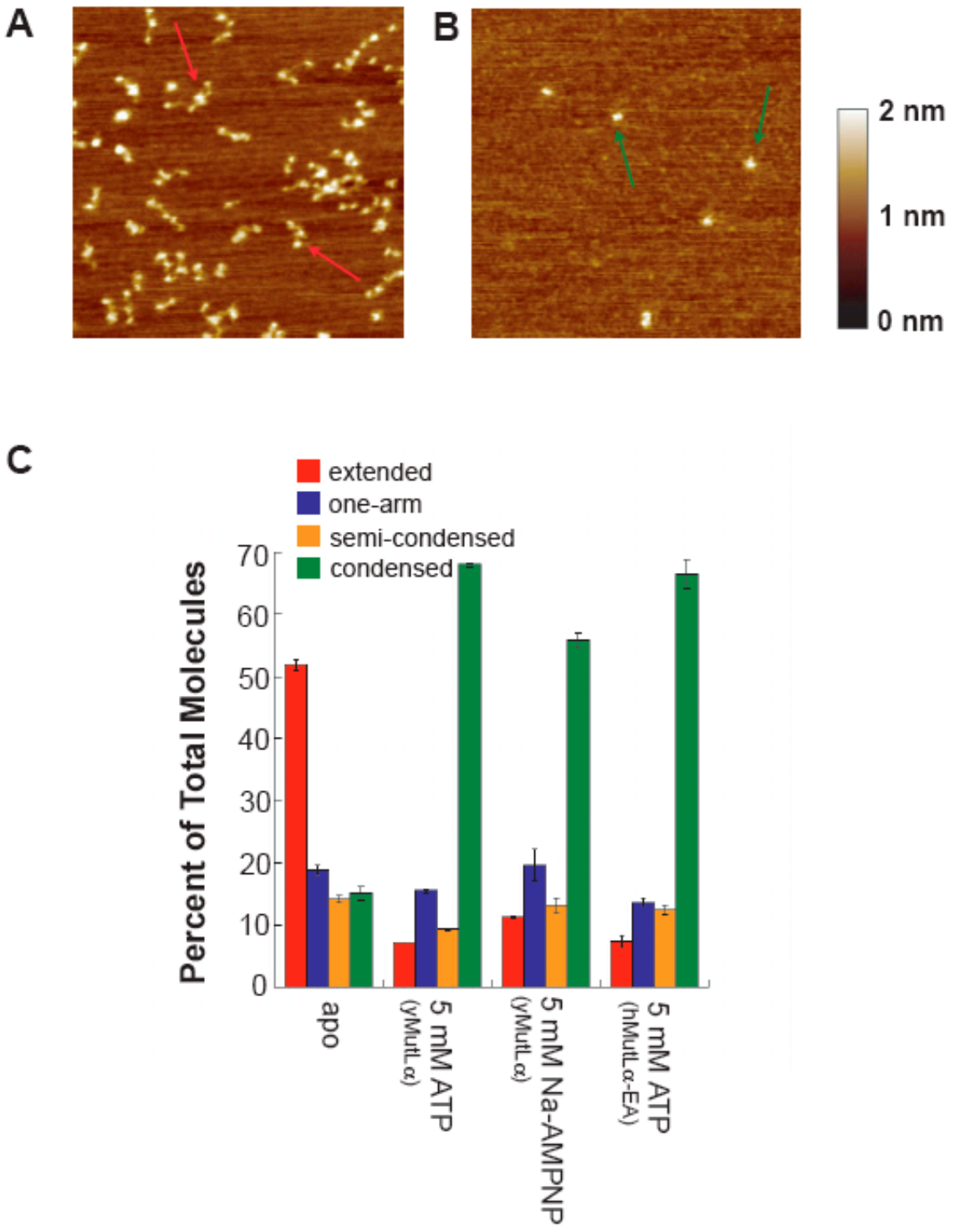

A
B
2 nm
1 nm
0 nm
C
extended
one-arm
semi-condensed
condensed
Percent of Total Molecules
70
60
50
40
30
20
10
0
apo
5 mM ATP
(yMutLα)
5 mM Na-AMPNP
(yMutLα)
5 mM ATP
(hMutLα-EA)

Figure 2.7

Figure 2.7: Comparison of hydrolysis-deficient experimental conditions.

Comparison of hMutLα-EA in the absence (A) and presence (B) of 5 mM ATP. Both are 500 x 500 nm scans; hMutLα concentration is 30 nM. Arrows point out individual molecules in the v-shaped (A) and condensed (B) state. Distribution of MutLa molecules among each of the four states. apo yMutLα (n=727), yMutLα + 5 mM ATP (n=1021), yMutLα + Na•AMPPNP (n=213), hMutLα•EA + 5mM ATP (n=161). Red = extended; blue = one-arm; orange = semi-condensed; green = condensed. Data were tallied by hand from 1 x 1 μm AFM images taken from multiple experiments.

To further investigate the role of ATP hydrolysis, we imaged yMutLα in the presence of the non-hydrolyzable ATP analog 5'-adenylyl-beta-gamma-imidodiphosphate (AMPPNP). In our initial experiments, we used AMPPNP as purchased from Sigma which comes as a tetralithium salt, and we found that it did not promote the condensation of MutLα, which is in direct contrast to the results obtained using the hMutLα-EA hydrolysis mutant. Because LiCl, in the absence of other salts, has been shown to inhibit the ATPase activity of GHL ATPases (Hu, Machius et al. 2003), we added 20mM LiCl to our 5 mM ATP reactions and found that it inhibited the formation of the condensed state (data not shown). Consequently, we purified the commercially available AMPPNP using an ion exchange column and eluted it using Na^+ (see materials & methods) to for the sodium salt (Na•AMPPNP). Depositions of yMutLα in the presence of Na•AMPPNP show that ~55% of the molecules are in the condensed state (Figure 2.7C). These results, taken together with the data on hMutLα-EA, strongly suggest that ATP binding, but not hydrolysis, is required to induce the formation of the fully condensed state.

Discussion

X-ray crystallography is a good technique for the indirect visualization of structured proteins and protein domains. However, there are no full-length structures of MutL or MutLα, likely because of the intrinsic disorder of the linker connecting the N- and C- termini as well as the large conformational changes that take place upon adenine nucleotide binding. AFM imaging has allowed us to not only visualize the protein in four

different conformational states but also to visualize the linker arm, which is unlikely to be detected with other techniques.

Nucleotide binding induces conformational changes and disorder-order transitions in MutLα

Our experiments show that in the absence of added adenine nucleotide, MutLα exists predominantly in an extended conformation (Figures 2.1 & 2.3). In this extended conformation, there are no visible interactions between the N- and C- terminal domains or between the two N-terminal domains of the heterodimer. There is likely no ATP bound to this dominating extended state, nor to the other, lesser-populated states.

The binding of one ATP to MutLα, likely to Mlh1 (K_M = 69 μM in an N-terminal construct of yMlh1 vs. K_M = 1500 μM in an N-terminal construct of yPms1) (Hall, Shcherbakova et al. 2002), drives the formation of the one-armed state. The one-armed state shows an increase in secondary structure compared to the extended state, as observed by CD in the absence (extended) and presence of 0.1 mM ATP (one-arm) (Figure 2.5). In addition, Mlh1 is slightly less sensitive (1.5 to 2-fold) to trypsin than Pms1 under all conditions (Figure 2.4 and data not shown). Taken together, these results suggest that binding of ATP to Mlh1 promotes secondary structure formation in its linker arm condensing the N-terminus of Mlh1 onto the dimerized C-terminal domains.

We also show that 5 mM ATP (or ADP), where both nucleotide-binding sites should be occupied, causes a condensation of both subunits in the heterodimer. In the condensed state, the interactions between linker regions and the N- and C-termini likely persist, with the addition of interactions between the two N-termini. The observation that

the hydrolysis mutant, hMutLα-EA, condenses in the presence of ATP strongly suggests that condensation is linked to binding of adenine nucleotide but not to ATP hydrolysis.

The semi-condensed state is likely a conformational rearrangement of the condensed state, with the contacts between the two N-termini remaining, but with the linker regions partially unfolded and no interaction between the N-termini and C-termini. This state suggests direct interaction between the two N-termini in MutLα. The presence of all four of these states in the absence of adenine nucleotide suggests that the four protein states are in dynamic equilibrium and that binding and hydrolysis of ATP followed by ADP release drive MutLα into the different conformational states.

The observed ordering of the previously disordered regions in MutLα upon ATP binding is analogous to the ordering that happens after ligand binding in intrinsically disordered proteins (Dunker, Lawson et al. 2001; Xie, Vucetic et al. 2007; Xie, Vucetic et al. 2007). Although ATP binding alone is unlikely to induce order in the ~160 and ~300 disordered amino acids in the linker arms, it probably promotes interactions of the linker regions with the other regions of MutLα, such as the C-terminal domains, and these interactions, together with the ATP binding may lead to the significant increase in secondary structure that is observed in the CD spectra. As such, this disorder-order transition adds another class of such transitions to natively disordered proteins (Dunker, Lawson et al. 2001; Xie, Vucetic et al. 2007; Xie, Vucetic et al. 2007) where binding of a small molecule to one region of a protein, which is predominately ordered, allosterically promotes the folding of an adjacent region of the protein.

Comparison of MutLα with other GHL ATPases

The members of the GHL ATPase family are diverse in their function. Despite this diversity of function, there remains a common theme: the binding and subsequent hydrolysis of ATP drives protein conformational changes that are linked to the timing of their mechanistic cycles (Ban, Junop et al. 1999; Dutta and Inouye 2000; Corbett and Berger 2003; Immormino, Dollins et al. 2004; Corbett and Berger 2005; Dollins, Immormino et al. 2005; Ali, Roe et al. 2006; Chu, Maynard et al. 2006; Shiau, Harris et al. 2006). Reflective of this common theme, the overall architectures of proteins in the GHL ATPase family are similar; the N-terminus contains a conserved ATPase binding domain and the C-terminus contains a varying dimerization domain (Dutta and Inouye 2000). One of the sources of variation between GHL proteins is in the connection between the N- and C- domains. Most GHL proteins have a structured middle domain; whereas, MutLα contains large disordered regions between the N- and C- terminal domains, which were previously predicted (Guarne, Ramon-Maiques et al. 2004) and confirmed by this study.

In spite of the functional differences between the MutLα and the other GHL family members, the four states observed here show similarity to conformational states of Grp94 and HtpG which have been studied using electron microscopy (EM) and/or x-ray crystallography (Wearsch and Nicchitta 1996; Shiau, Harris et al. 2006). Grp94 and HtpG have ordered middle domains (in comparison with the disordered linker of MutLα) and no one-armed state has been observed for Grp94 or HtpG.

In the absence of adenine nucleotide, both Grp94 and HtpG are seen by EM in an extended, v-shaped, conformation very similar to the extended conformation seen here

for MutLα by AFM. The difference between the extended MutLα and extended Grp94 and HtpG is the absence of the disordered linker arms connecting N- and C-termini in Grp94 and HtpG. In the presence of the non-hydrolyzable ATP analog AMPPNP, HtpG undergoes a large conformational change from the open, extended state to a more compact state (Shiau, Harris et al. 2006). This state appears similar to the semi-condensed state of MutLα seen by AFM (Figure 2.1D), with the N-termini dimerized, but not collapsed upon the C-termini. When visualized by EM in the presence of ADP, HtpG becomes even more condensed than it is in the presence of AMPPNP.

Examination of the crystal structures of HtpG in the absence and the presence of ADP along with the structure of yHsp82, the yeast homolog of Hsp90, shows that the Hsp90 proteins also exist in an extended conformation in the absence of ATP (Ali, Roe et al. 2006; Shiau, Harris et al. 2006). The addition of ADP causes dimerization of the N-terminal domains (seen with HtpG), while the addition of ATP causes a twisted collapse around the dimerized N-terminal domains (seen with yHsp82) and a 20 Å movement of the N-terminal domains towards each other (Ali, Roe et al. 2006). Perhaps the semi-condensed structure we visualize by AFM is an "untwisted" version of the condensed state.

ATPase cycle for MutLα

Each of the four observed states of MutLα is a snapshot of the conformational changes involved in the process of ATP binding, hydrolysis and ADP release. Based on these four observed states, one possible model for an ATPase cycle suggests that adenine nucleotides bind sequentially to MutLα (Figure 2.8). The model is consistent with different K_Ms of the N-terminal domains of yMlh1 and yPms1 (Hall, Shcherbakova et al.

2002). The extended state corresponds to MutLα with no nucleotide bound (apo). The binding of the first ATP condenses one "arm", likely Mlh1, forming the one-armed state; the binding of a second ATP condenses the other arm, forming the condensed state.

The one-armed state is formed by the ordering of the linker region of one protein in MutLα, causing it to condense into the already dimerized C-termini. The occurrence of this state suggests that there may be contacts formed between the N- and C- terminal domains in MutLα after ATP binds to the protein. Partial proteolysis suggests that these interactions are dynamic, because the digestion patterns of yMutLα alone and yMutLα plus 0.1 mM ATP are similar (Figure 2.4).

The hydrolysis and subsequent release of both bound adenine nucleotides, either sequentially (following the one-armed pathway) or at the same time (following the semi-condensed pathway), returns the protein to the original extended conformation. Because high ADP concentration also drive the formation of the condensed state, it is likely that ADP release, not ATP hydrolysis, causes the cycling of the protein through the four states. As the cycle progresses, conformational changes expose different surfaces and new structured regions to solution, making them available for binding by other proteins.

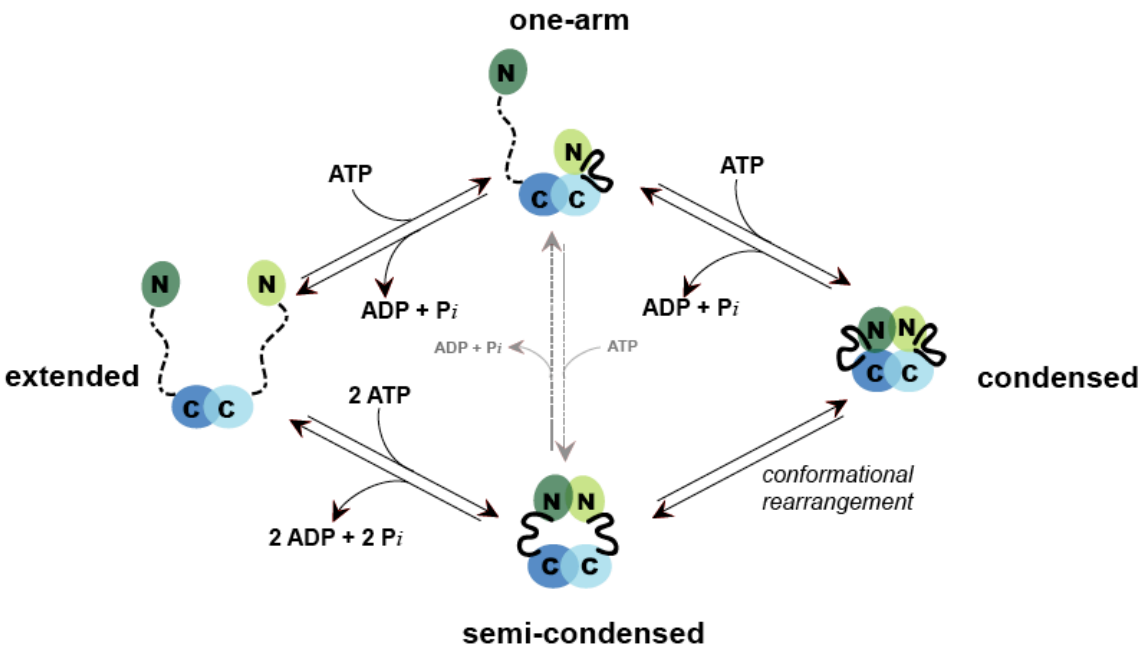


Figure 2.8: ATPase cycle for MutLα.

The extended state corresponds to MutLα with no nucleotide bound (apo). The binding of the first ATP condenses one "arm" (probably Mlh1), forming the one-armed state; the binding of a second ATP condenses the other arm (Pms2), forming the condensed state. The hydrolysis and release of both bound adenine nucleotides, either sequentially (following the one-armed pathway) or at the same time (following the semi-condensed pathway), returns the protein to the extended conformation.

Our model contrasts with a previously proposed ATPase cycle for MutL that was based upon mutational studies alongside biochemical and crystallographic data (Ban, Junop et al. 1999). In the *E. coli* MutL model, both MutL monomers bind ATP at the same time, followed by N-terminal dimerization. Each MutL monomer then simultaneously hydrolyzes ATP, the N-terminal domains dissociate and ADP is released. The extended state we observe for the eukaryotic MutLα correlates well with the N-terminal dissociated state seen for MutL, while the semi-condensed state correlates with the dimerized N-terminal state.

However, MutLα is a heterodimer, and each protein contained therein has differing ATP affinities and rates of ATP hydrolysis. These important differences make simultaneous ATP binding and hydrolysis unlikely for the subunits of MutLα. As we have observed in the one-armed state, the condensation of one protein subunit (after the binding of one ATP) occurs in the absence of the condensation of the other.

MutLα conformational changes in the context of MMR and cellular activities

It has long been suggested that MutL and MutL homologs function as a switch or molecular matchmaker, changing conformation to coordinate the functions of the mismatch repair machinery. In this paper, we report direct observations of conformational changes to MutLα in response to adenine nucleotide. ATP hydrolysis is required for repair (Pang, Prolla et al. 1997; Spampinato and Modrich 2000; Tran and Liskay 2000; Hall, Wang et al. 2001; Raschle, Dufner et al. 2002) as well as the endonucleolytic activity of MutLα (Kadyrov, Dzantiev et al. 2006). Since MutLα also interacts with PCNA and ExoI in the MMR pathway, perhaps our observed

conformational changes mediate the timing of these interactions as well as the enzymatic function of MutLα.

The *in vivo* concentrations of ADP and ATP (~1 mM and ~3 mM, respectively) (Williams, Fulton et al. 1993) suggest that MutLα would rarely exist in the apo form unless this form is stabilized by other proteins. It has been demonstrated that interactions between MutSα and MutLα are mediated by the N-terminal domain of Mlh1 alone and are not linked to the ATPase activity of MutLα (Raschle, Dufner et al. 2002; Plotz, Raedle et al. 2003). Since the ATPase activity of MutLα does not affect interactions with MutSα, MutSα may be able to interact with MutLα in any conformational state, but to do so, MutSα must be in the appropriate state.

Stable complexes containing MutSα, MutLα, PCNA, mismatched DNA and ATP have been observed *in vitro* (Dzantiev, Constantin et al. 2004; Hidaka, Takagi et al. 2005; Lee and Alani 2006); (Constantin, Dzantiev et al. 2005; Zhang, Yuan et al. 2005). Based on observed *in vitro* interactions and on *in vivo* ATP concentrations, MutLα is likely in the condensed state when interacting with MutSα and PCNA. With the subsequent hydrolysis of one or both bound ATP molecules, the protein likely changes from the condensed state to the semi-condensed and/or apo state, changing the surface of MutLα and possibly hiding the surface with which PCNA interacts, exposing the ExoI interaction surface. Once ADP and P_i are released into solution, MutLα will be "reset" and ready to participate in another round of mismatch repair. The time required to hydrolyze ATP and then release ADP and P_i, as well as the time required for conformational changes likely causes a "pause" in the process of mismatch repair,

allowing for the proteins currently interacting with MutLα to complete their intended processes.

As mentioned in the introduction, the proteins that make up MutLα also participate in other DNA transactions. The conformational changes reported here may mediate protein-protein interactions in those cellular activities as well.

Materials & Methods

Chemicals

ATP, ADP and AMPPNP were purchased from Sigma.

Protein Expression and Purification

yMutLα was purified as previously described (Hall and Kunkel 2001), with the exception of HEPES, pH 7.3 being used in the dialysis buffer instead of Tris, pH 7.3. hMutLα and hMutLα-EA were purified as described (Kadyrov, Dzantiev et al. 2006).

Atomic Force Microscopy

For AFM imaging, yMutLα, hMutLα or hMutLα-EA was diluted to a final concentration of 30 nM in imaging buffer (20 mM HEPES, pH 7.3, 5 mM MgOAc, 25mM NaOAc) at room temperature. A volume of 10 μl of the diluted protein was deposited onto freshly cleaved ruby mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water; excess water was blotted from the surface and the surface was then dried using a stream of nitrogen.

For experiments where ATP, ADP or AMPPNP (Na•AMPPNP) were present, the adenine nucleotide and MutLα were mixed and then diluted in imaging buffer to a final

adenine nucleotide concentration of 0.1 mM, 1 mM or 5 mM and protein concentration of 30 nM. The sample was then deposited as for MutLα protein alone. The sample preparation and deposition process (from addition of adenine nucleotide to drying of sample) took ~1 minute.

AFM images were captured in air using either a Nanoscope III or IIIa (Digital Instruments, Santa Barbara, CA) microscope in tapping mode. Pointprobe Plus tapping mode silicon probes (Agilent Technology, Tempe, AZ) with resonance frequencies of approximately 170 kHz were used for imaging. Images were collected at a speed of 2-3 Hz with an image size of 1 μm at 512 x 512 pixel resolution.

Images were analyzed using Nanoscope III v5.12r3 software (Veeco, Santa Barbara, CA), Image SxM, v 1.69 (Barrett 2006) and NIH ImageJ (Rasband 1997-2006) with OpenAFM and Cell Counter plug-ins. Volume analysis was performed as previously described (Ratcliff and Erie 2001; Yang, Wang et al. 2003). Statistical plots were generated with Kaleidagraph (Synergy Software, Reading, PA).

Purification of AMPPNP

A 1 mL HiTrap DEAE column (GE Biosciences) was prepared according to manufacturers instructions. AMPPNP was dissolved in ddH_2O (0.5 mL; final concentration 0.1 M) and loaded onto the column by syringe. The column was washed with ten column volumes of ddH_2O, and the AMPPNP was eluted with 0.25 M NaCl (~100 μL fraction size). The AMPPNP eluted in the first three fractions (after the 1 mL column volume came off), and was pooled.

Circular Dichroism Spectropolarimetry

CD spectra (λ = 210-300 nm) of 0.4 μM yMutLα in imaging buffer in the presence and absence of 0.1 mM ATP were acquired on an AVIV spectropolarimeter at 10^{o}C, with 2 nm resolution and an averaging time of 3 sec per point, in 1-cm quartz cuvettes. Plotted curves are the average of 3 experiments, with the background spectra (either imaging buffer alone or imaging buffer + 0.1 mM ATP) subtracted out.

Partial Proteolysis

In experiments testing for protease protection, 5 ng of trypsin was added to reactions containing 8 μg of yMutLα alone and with 0.1 mM ATP or 5 mM ATP in imaging buffer. The reaction was allowed to proceed for 5 minutes at room temperature before stopping the reaction by the addition of SDS load and incubation at 95°C for 10 minutes. All reactions were then run on a 10% SDS-PAGE gel. Protein bands were visualized by staining with Coomassie Brilliant Blue R-250.

References

Ali, M. M., S. M. Roe, et al. (2006). "Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex." Nature **440**(7087): 1013-7.

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

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

Barrett, S. D. (2006). Image SXM.

Bellacosa, A. (2001). "Functional interactions and signaling properties of mammalian DNA mismatch repair proteins." Cell Death Differ **8**(11): 1076-92.

Bergerat, A., B. de Massy, et al. (1997). "An atypical topoisomerase II from Archaea with implications for meiotic recombination." Nature **386**(6623): 414-7.

Bryson, K., L. J. McGuffin, et al. (2005). "Protein structure prediction servers at University College London." Nucleic Acids Res **33**(Web Server issue): W36-8.

Chu, F., J. C. Maynard, et al. (2006). "Identification of novel quaternary domain interactions in the Hsp90 chaperone, GRP94." Protein Sci **15**(6): 1260-9.

Constantin, N., L. Dzantiev, et al. (2005). "Human mismatch repair: Reconstitution of a nick-directed bidirectional reaction." J Biol Chem.

Corbett, K. D. and J. M. Berger (2003). "Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution." Embo J **22**(1): 151-63.

Corbett, K. D. and J. M. Berger (2005). "Structural dissection of ATP turnover in the prototypical GHL ATPase TopoVI." Structure (Camb) **13**(6): 873-82.

Dollins, D. E., R. M. Immormino, et al. (2005). "Structure of unliganded GRP94, the endoplasmic reticulum Hsp90. Basis for nucleotide-induced conformational change." J Biol Chem **280**(34): 30438-47.

Dunker, A. K., J. D. Lawson, et al. (2001). "Intrinsically disordered protein." J Mol Graph Model **19**(1): 26-59.

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

Dzantiev, L., N. Constantin, et al. (2004). "A defined human system that supports bidirectional mismatch-provoked excision." Mol Cell **15**(1): 31-41.

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

Guarne, A., M. S. Junop, et al. (2001). "Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHL ATPase." Embo J **20**(19): 5521-31.

Guarne, A., S. Ramon-Maiques, et al. (2004). "Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair." Embo J **23**(21): 4134-45.

Hall, M. C. and T. A. Kunkel (2001). "Purification of eukaryotic MutL homologs from Saccharomyces cerevisiae using self-affinity technology." Protein Expr Purif **21**(2): 333-42.

Hall, M. C., P. V. Shcherbakova, et al. (2002). "Differential ATP binding and intrinsic ATP hydrolysis by amino-terminal domains of the yeast Mlh1 and Pms1 proteins." J Biol Chem **277**(5): 3673-9.

Hall, M. C., H. Wang, et al. (2001). "High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer." J Mol Biol **312**(4): 637-47.

Hidaka, M., Y. Takagi, et al. (2005). "PCNA-MutSalpha-mediated binding of MutLalpha to replicative DNA with mismatched bases to induce apoptosis in human cells." Nucleic Acids Res **33**(17): 5703-12.

Hu, X., M. Machius, et al. (2003). "Monovalent cation dependence and preference of GHKL ATPases and kinases." FEBS Lett **544**(1-3): 268-73.

Immormino, R. M., D. E. Dollins, et al. (2004). "Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone." J Biol Chem **279**(44): 46162-71.

Iyer, R. R., A. Pluciennik, et al. (2006). "DNA mismatch repair: functions and mechanisms." Chem Rev **106**(2): 302-23.

Jager, A. C., M. Rasmussen, et al. (2001). "HNPCC mutations in the human DNA mismatch repair gene hMLH1 influence assembly of hMutLalpha and hMLH1-hEXO1 complexes." Oncogene **20**(27): 3590-5.

Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.

Kadyrov, F. A., L. Dzantiev, et al. (2006). "Endonucleolytic function of MutLalpha in human mismatch repair." Cell **126**(2): 297-308.

Kelly, S. M., T. J. Jess, et al. (2005). "How to study proteins by circular dichroism." Biochim Biophys Acta **1751**(2): 119-39.

Kosinski, J., I. Steindorf, et al. (2005). "Analysis of the quaternary structure of the MutL C-terminal domain." J Mol Biol **351**(4): 895-909.

Kunkel, T. A. and D. A. Erie (2005). "DNA mismatch repair." Annu Rev Biochem **74**: 681-710.

Lee, S. D. and E. Alani (2006). "Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA." J Mol Biol **355**(2): 175-84.

McGuffin, L. J., K. Bryson, et al. (2000). "The PSIPRED protein structure prediction server." Bioinformatics **16**(4): 404-5.

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

Modrich, P. (2006). "Mechanisms in eukaryotic mismatch repair." J Biol Chem **281**(41): 30305-9.

Modrich, P. and R. Lahue (1996). "Mismatch repair in replication fidelity, genetic recombination, and cancer biology." Annu Rev Biochem **65**: 101-33.

Nielsen, F. C., A. C. Jager, et al. (2004). "Characterization of human exonuclease 1 in complex with mismatch repair proteins, subcellular localization and association with PCNA." Oncogene **23**(7): 1457-68.

O'Brien, V. and R. Brown (2006). "Signalling cell cycle arrest and cell death through the MMR System." Carcinogenesis **27**(4): 682-92.

Pang, Q., T. A. Prolla, et al. (1997). "Functional domains of the Saccharomyces cerevisiae Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations." Mol Cell Biol **17**(8): 4465-73.

Plotz, G., J. Raedle, et al. (2003). "N-terminus of hMLH1 confers interaction of hMutLalpha and hMutLbeta with hMutSalpha." Nucleic Acids Res **31**(12): 3217-26.

Rasband, W. S. (1997-2006). ImageJ. US National Institutes of Health, Bethesda, Maryland, USA

Raschle, M., P. Dufner, et al. (2002). "Mutations within the hMLH1 and hPMS2 subunits of the human MutLalpha mismatch repair factor affect its ATPase activity, but not its ability to interact with hMutSalpha." J Biol Chem **277**(24): 21810-20.

Ratcliff, G. C. and D. A. Erie (2001). "A novel single-molecule study to determine protein--protein association constants." J Am Chem Soc **123**(24): 5632-5.

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

Shiau, A. K., S. F. Harris, et al. (2006). "Structural Analysis of E. coli hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements." Cell **127**(2): 329-340.

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

Stojic, L., R. Brun, et al. (2004). "Mismatch repair and DNA damage signalling." DNA Repair (Amst) **3**(8-9): 1091-101.

Tomer, G., A. B. Buermeyer, et al. (2002). "Contribution of human mlh1 and pms2 ATPase activities to DNA mismatch repair." J Biol Chem **277**(24): 21801-9.

Tran, P. T. and R. M. Liskay (2000). "Functional studies on the candidate ATPase domains of Saccharomyces cerevisiae MutLalpha." Mol Cell Biol **20**(17): 6390-8.

van Noort, J., T. van Der Heijden, et al. (2003). "The coiled-coil of the human Rad50 DNA repair protein contains specific segments of increased flexibility." Proc Natl Acad Sci U S A **100**(13): 7581-6.

Wearsch, P. A. and C. V. Nicchitta (1996). "Endoplasmic reticulum chaperone GRP94 subunit assembly is regulated through a defined oligomerization domain." Biochemistry **35**(51): 16760-9.

Williams, S. P., A. M. Fulton, et al. (1993). "Estimation of the intracellular free ADP concentration by 19F NMR studies of fluorine-labeled yeast phosphoglycerate kinase in vivo." Biochemistry **32**(18): 4895-902.

Xie, H., S. Vucetic, et al. (2007). "Functional Anthology of Intrinsic Disorder. 3. Ligands, Post-Translational Modifications, and Diseases Associated with Intrinsically Disordered Proteins." J Proteome Res.

Xie, H., S. Vucetic, et al. (2007). "Functional Anthology of Intrinsic Disorder. 1. Biological Processes and Functions of Proteins with Long Disordered Regions." J Proteome Res.

Yang, Y., L. E. Sass, et al. (2005). "Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions." Nucleic Acids Res **33**(13): 4322-34.

Yang, Y., H. Wang, et al. (2003). "Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy." Methods **29**(2): 175-87.

Zhang, Y., F. Yuan, et al. (2005). "Reconstitution of 5'-directed human mismatch repair in a purified system." Cell **122**(5): 693-705.

CHAPTER 3

CYTIDINE DEAMINASES

Introduction

Unlike the rest of this dissertation, the work contained within this chapter deals with proteins designed to make mutations in DNA, not avoid them. This chapter focuses on the oligomeric states of three members of the AID/ABOBEC family of proteins, AID, APOBEC2 and APOBEC3G. Members of the AID/APOBEC family of proteins (AID = activation induced cytosine deaminase; APOBEC = apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like) have been shown to deaminate cytidines found in RNA or ssDNA to uracils in the process of innate and adaptive immunity, including somatic hypermutation (SHM).

The AID/APOBEC family of proteins belongs to the cytidine deaminase superfamily. The ancestral members of this superfamily deaminate free bases (as opposed to those found in RNA or DNA) and are either dimeric or tetrameric. Dimerization is often used in the cytidine deaminase superfamily to form the active site. In contrast to their forefathers, there is little known about the oligomeric state of the members of the AID/APOBEC family.

AID

The cytidine deaminase, AID, is required for class switch recombination (CSR) and somatic hypermutation (SHM), two antigen-activated mechanisms that lead to the formation of isotype-switched, high affinity antibodies characteristic of the secondary immune response (Muramatsu, Kinoshita et al. 2000; Revy, Muto et al. 2000; Arakawa, Hauschild et al. 2002; Harris, Sale et al. 2002; Yang, Obiakor et al. 2005). AID deaminates cytidines, generating uracils in DNA encoding the variable and the switch regions of the immunoglobulin locus (Petersen-Mahrt, Harris et al. 2002; Rada, Williams et al. 2002).

In vitro, AID deaminates single-stranded but not double-stranded DNA (Bransteitter, Pham et al. 2003; Dickerson, Market et al. 2003; Pham, Bransteitter et al. 2003; Sohail, Klapacz et al. 2003). It has been shown that AID will function on dsDNA, but only if it is actively undergoing transcription, especially if AID has been phosphorylated by PKA and is in complex with RPA (Chaudhuri, Khuong et al. 2004; Basu, Chaudhuri et al. 2005; Besmer, Market et al. 2006).

AID preferentially deaminates microsequences associated with SHM hotspots (WGYW, where W = A or T, Y = Pyrimidine) (Bransteitter, Pham et al. 2003; Beale, Petersen-Mahrt et al. 2004; Bransteitter, Pham et al. 2004; Rogozin and Diaz 2004; Yu, Huang et al. 2004). Data from pull-down assays using tagged-AID proteins and from structural modeling based on the free cytidine deaminases suggest that AID may function as a dimer to deaminate DNA (Ta, Nagaoka et al. 2003; Xie, Sowden et al. 2004; Prochnow, Bransteitter et al. 2007). However, it remains unclear whether the active site of AID is at the dimer interface, or if AID monomers are catalytically active.

APOBEC2

APOBEC2 was found by sequence homology to other APOBEC proteins, and is found primarily in cardiac and skeletal muscle (Liao, Hong et al. 1999). Unlike the other members of the AID/APOBEC family, APOBEC2 contains no deamination activity and its function remains unknown. Despite this lack of enzymatic activity, the crystal structure of APOBEC2 has been solved, and because of the 44.6% sequence homology between the crystallized APOBEC2 fragment and AID, this crystal structure serves as a model for other AID/APOBEC family members (Prochnow, Bransteitter et al. 2007; Zhang, Mangeat et al. 2007).

The crystal structure of APOBEC2 (Prochnow, Bransteitter et al. 2007) (PDB id: 2NYT) revealed that APOBEC2 exists as a dimer of dimers. One dimer interface is made by the joining of two sets beta-sheets (one from each monomer) (Figure 5.1A); whereas the second dimerization interface is made by hydrophobic packing, salt bridges and hydrogen bonding (Figure 5.1B).

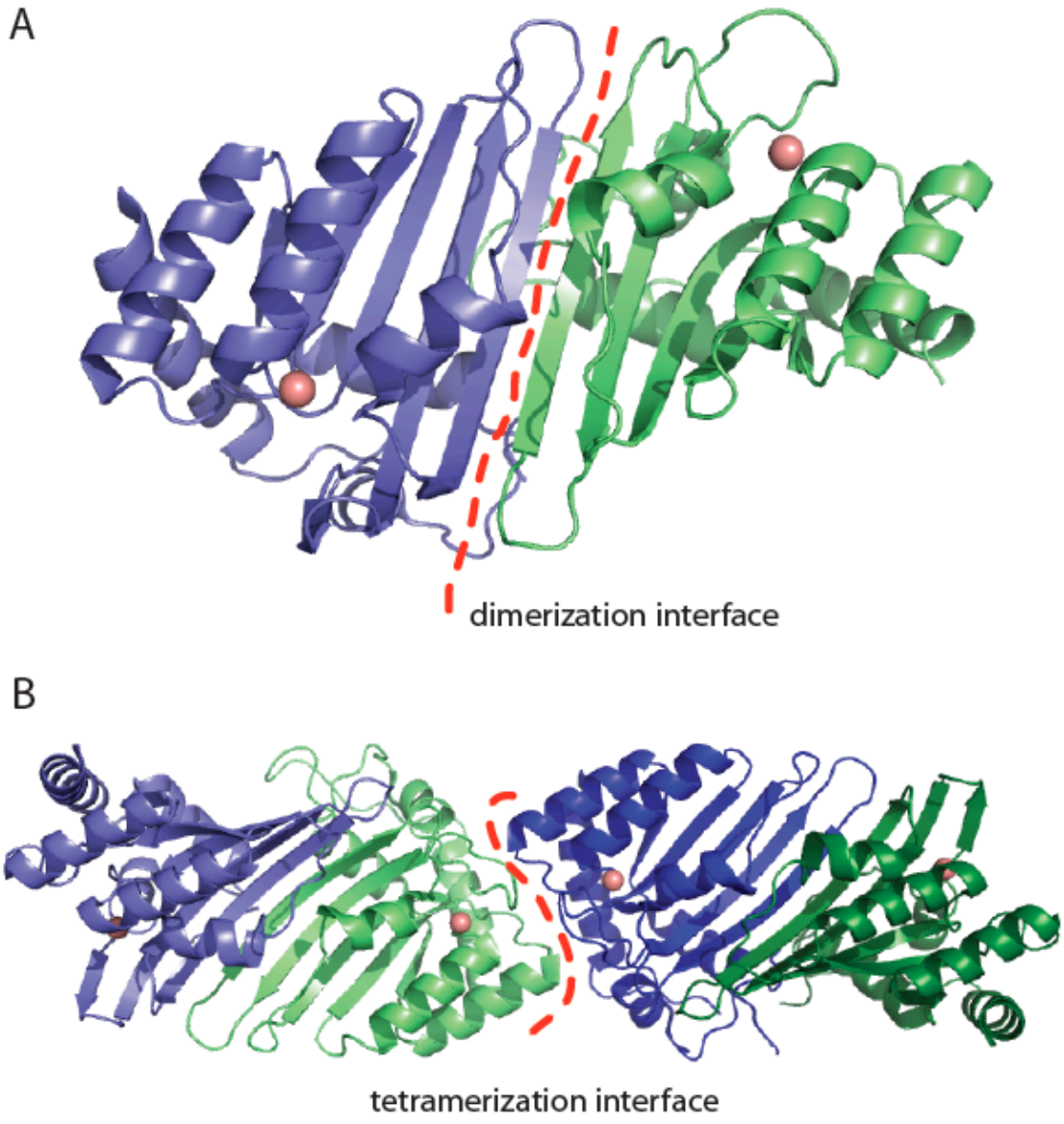


Figure 3.1: Illustration of dimer (A) and tetramer (B) interfaces for APOBEC2 (PDB ID 2NYT).

Monomers of APOBEC2 are shown in blues and greens; zinc ions are shown as pink spheres; red dashed lines show dimerization and tetramerization interfaces.

APOBEC3G

APOBEC3G works to inhibit the infectivity of viruses like HIV and Hepatitis B, presumably by deaminating cytosines to generate uracils, eventually targeting viral DNA for degradation (Sheehy, Gaddis et al. 2002; Nguyen, Gummuluru et al. 2007; Noguchi, Hiraga et al. 2007). Interaction of the HIV Viral infectivity factor protein (Vif) with APOBEC3G targets APOBEC3G for degradation via the proteosomal pathway. In Vif-deficient cells, however, APOBEC3G is able to survive and deaminate cytosines in the viral DNA, causing mutations in and eventual degradation of the viral DNA (Sheehy, Gaddis et al. 2002). APOBEC3G has been shown to act processively along ssDNA, with a preference for deaminating the third cytosine of -CCC- sequences (Chelico, Pham et al. 2006).

Results

AFM, which yields topographic images of molecules deposited on a surface, has proven to be a powerful tool for determining the oligomerization state of proteins as well as protein-protein association constants (Ratcliff and Erie 2001; Xue, Ratcliff et al. 2002; Bao, Wang et al. 2003; Yang, Wang et al. 2003; Tessmer, Moore et al. 2005; Yang, Sass et al. 2005). It has been demonstrated that the volume of a protein or protein complex as measured by AFM exhibits a linear dependence on the molecular weight of proteins (Equation 3.1), where V is the volume as measured by AFM and MW is the molecular weight of the protein (Ratcliff and Erie 2001; Yang, Wang et al. 2003).

Equation 3.1 $$V = 1.2(MW) - 14.7$$

Using this relationship, the oligomerization states for many different proteins and protein-protein complexes have been determined in the presence and absence of DNA (Ratcliff and Erie 2001; Yang, Wang et al. 2003; Tessmer, Moore et al. 2005; Yang, Sass et al. 2005).

AID

AID deaminates DNA loops and bubbles embedded in double-stranded DNA

To examine the deamination activity of AID on various single-stranded DNA substrates including substrates containing loops or bubbles, and to determine substrates for AFM analysis, we designed oligonucleotides containing the structures depicted in Figure 3.2. Deamination reactions were done for the various substrates following the scheme depicted in Figure 3.2A. AID deaminated all substrates tested, including the loop and bubble substrates, but not double-stranded DNA (Figures 3.2B and C). The best substrate for AID is the L-oligo, modeled after the oligonucleotide designed by Yu and colleagues (Yu, Huang et al.), with over 50% deamination after 15 minutes, followed by the single-stranded oligonucleotide modeled after V regions of human IgV heavy chains (over 40% deaminated product at 15 minutes) (Figure 3.2C). Even at high protein concentrations, in the single-stranded oligonucleotide that contains multiple dC's throughout its length, the majority of product is deaminated at or near the site resembling

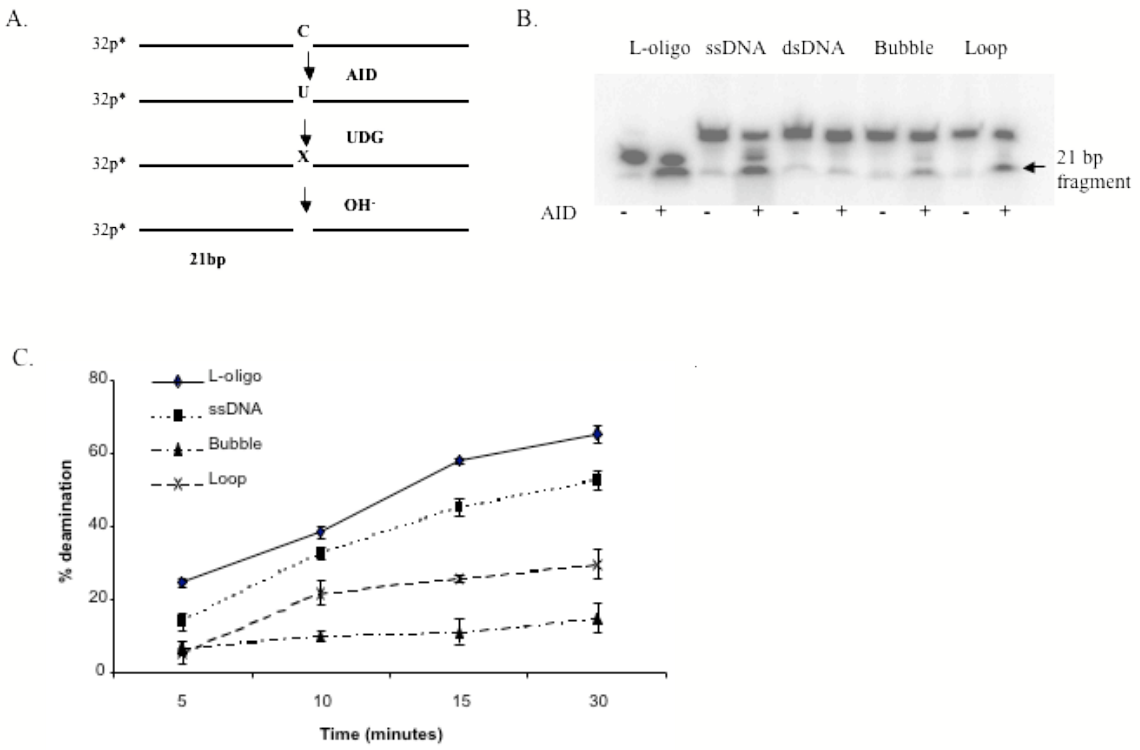


Figure 3.2: AID deaminates single-stranded DNA in various configurations.

(A) Schematic representation of deamination reaction and expected product following incubation of end-labeled oligonucleotide with AID and UDG. (B) AID deaminates most oligonucleotides where single-stranded DNA (ssDNA) is exposed but fails to deaminate double-stranded DNA (dsDNA), (C) Quantitation of autoradiographs after deamination reactions. The substrate devoid of secondary structure (L-oligo) deaminated best but there was substantial deamination of loop structures over time. All reactions contained 2 nM oligonucleotide and 500 nM AID protein and were repeated at least three times. (experiments done by S. Brar)

AID-mediated immunoglobulin somatic hypermutation hotspots (AGCA) (Figures 3.2b and 3.3), as seen previously by us and other groups (Bransteitter, Pham et al. 2003; Beale, Petersen-Mahrt et al. 2004; Bransteitter, Pham et al. 2004; Rogozin and Diaz 2004; Yu, Huang et al. 2004). Although the substrate containing a single-stranded 8-nucleotide loop is not deaminated as rapidly as the single-stranded oligonucleotides, its deamination is significantly faster than that of the substrate containing a 10-nucleotide bubble. For example, after a 15 minute incubation with AID, greater than 20 % of the loop substrate was deaminated while only 11% of the bubble substrate was deaminated (Figure 3.2C). Deamination of single-stranded hairpin loops was consistently better than that of double-stranded loops, regardless of the size of the loops, at all concentrations of protein and DNA tested (data not shown). This lower rate of deamination of bubble substrates probably results from formation of non-canonical base pairs in the loop, thereby disrupting the accessibility of the cytidine to AID (Chou, Chin et al.). Taken together, these results indicate that AID preferentially deaminates cytosines that are maximally exposed. This preference for single-stranded vs. double-stranded loops may reflect AID catalyzed-deamination of exposed single-stranded DNA in hairpin loops that might be formed due to the palindromic sequences often associated with CSR regions (Tashiro, Kinoshita et al.) or single-stranded DNA regions formed during transcription.

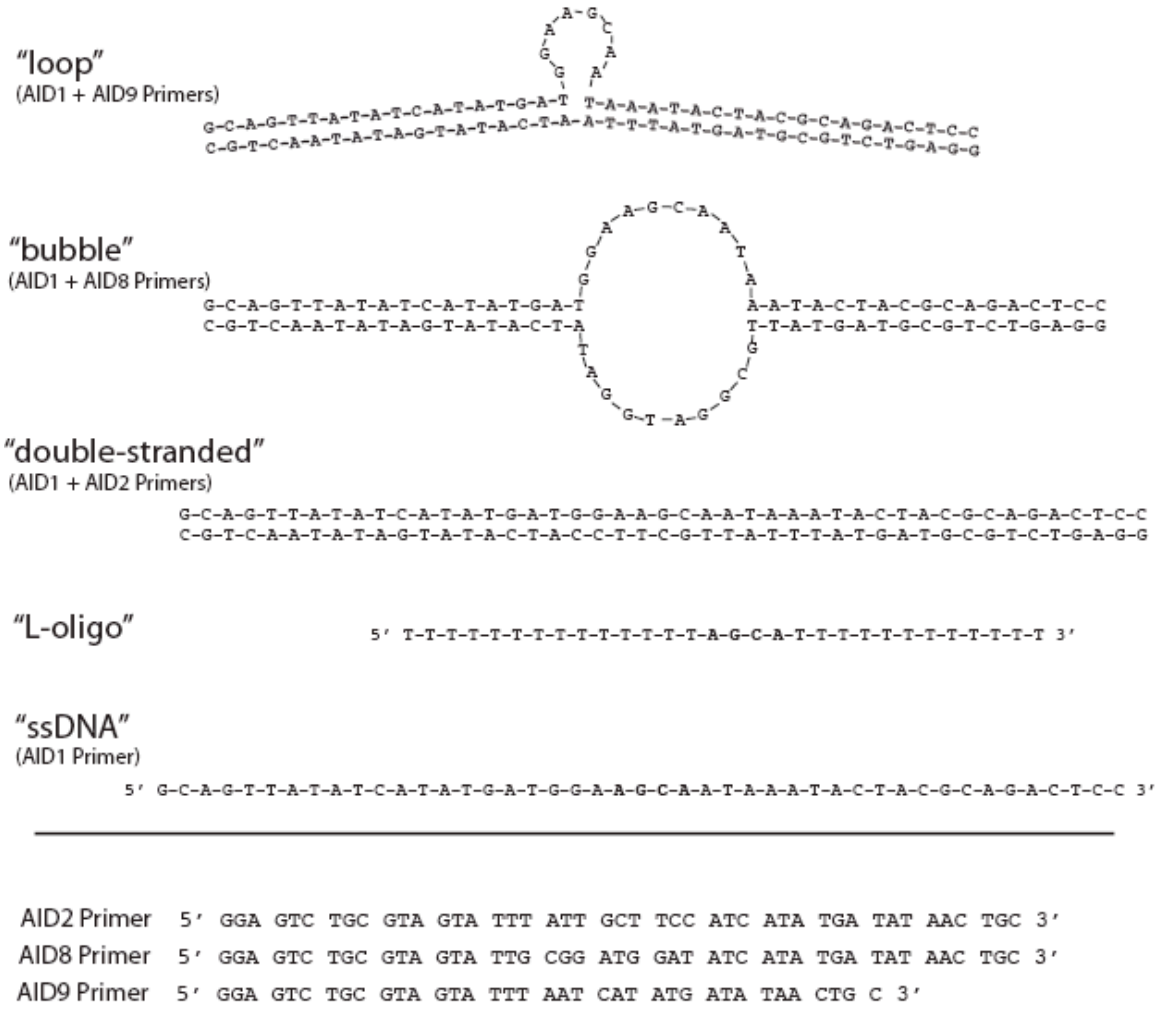


Figure 3.3: Predicted structures and sequences of substrates used for deamination assays and AFM imaging

AID exists as a monomer in the absence of DNA

We have used atomic force microscopy to determine the oligomerization state of AID. Figure 3.4A shows representative AFM images of AID in the absence of DNA substrate. To determine the oligomerization state of AID, the volumes of each protein as seen by AFM were determined from a large number of images. In the absence of DNA, a distribution of volumes shows a single peak at ~15 nm^3 (Figure 3.4C, black bars). This volume is consistent with the predicted volume of 15 nm^3 for a monomer of His-tagged AID (AID monomer MW = 24.8kDa).

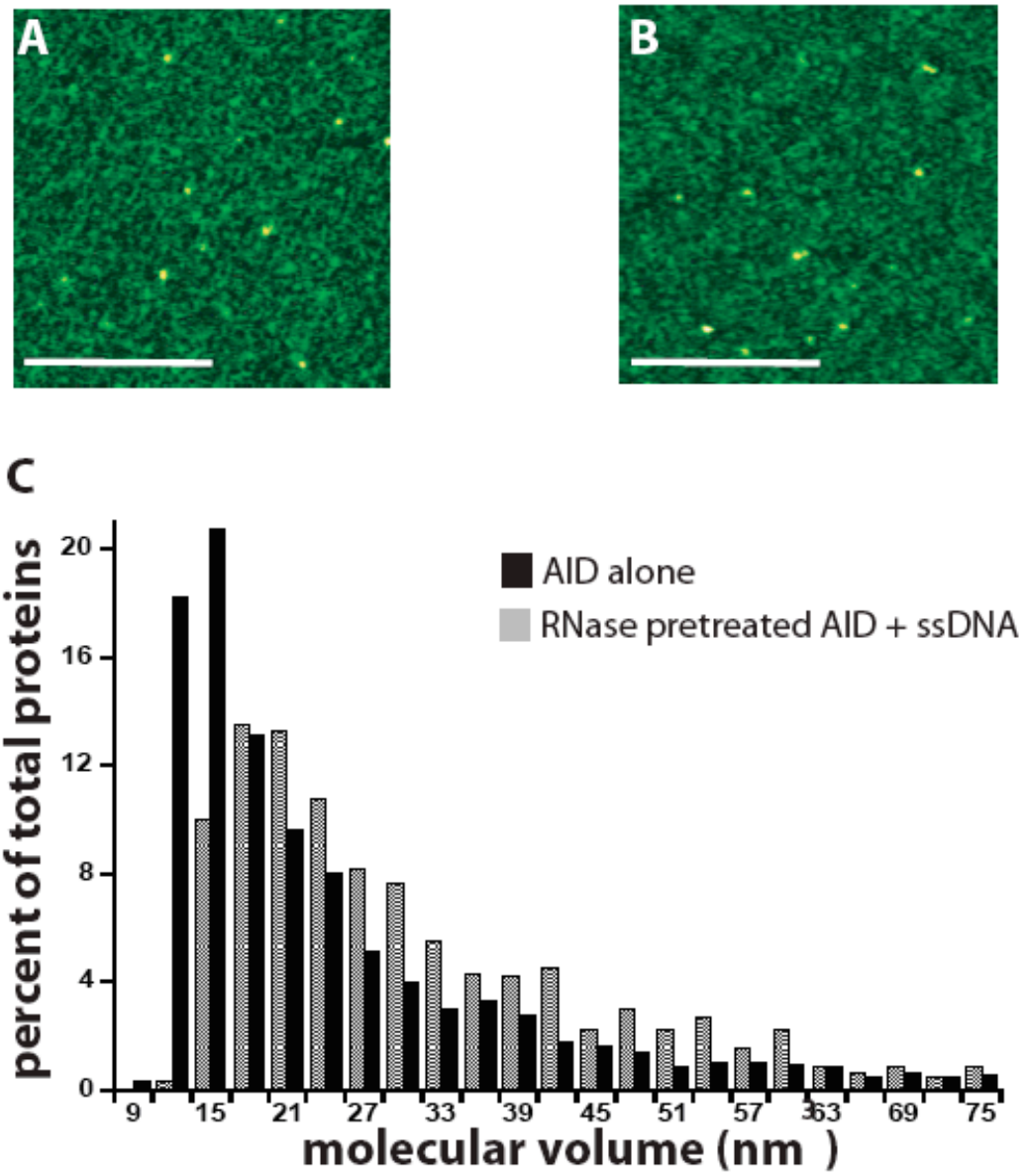

A
B
C
percent of total proteins
20
16
12
8
4
0
AID alone
RNase pretreated AID + ssDNA
9
15
21
27
33
39
45
51
57
63
69
75
molecular volume (nm)

Figure 3.4

Figure 3.4: AFM data on AID in the absence and presence of ssDNA substrate.

All AFM images are 500 x 500 nm; scale bars (white) represent 250 nm. (A) representative AFM scan of AID in the absence of ssDNA. (B) Representative AFM scan of RNaseA-pretreated AID in the presence of L-oligo. (C) Volume analysis of AID. Black bars represent AID in the absence of substrate. Grey bars represent RNase-pretreated AID in the presence of L-oligo

A very small fraction of molecules have volumes corresponding to the predicted volume of a dimer of AID (45 nm^3). Increasing the protein concentration from 5 nM to 20 nM or incubating at 500nM and diluting to 10 nM only immediately before deposition (see methods) did not increase the population of species with larger volumes (data not shown). These results indicate that AID exists predominantly as a monomer at these concentrations and suggests that any dimeric form would be transient, and have a relatively weak association constant (Ratcliff and Erie 2001; Yang, Wang et al. 2003).

AID exists primarily as a monomer in the presence of substrate DNA

To determine if ssDNA induces dimerization, AID was incubated with 10nM to 50nM of the L-oligo for 30 minutes before deposition. The concentrations of AID protein and oligonucleotide used in these AFM experiments are within the range of concentrations where significant deamination of substrate by AID after 30 minutes of incubation is observed.

Since RNA co-purifies with AID, and the removal of RNA (usually with RNaseA) is necessary for the activation and deamination activity of AID (Bransteitter, Pham et al. 2003), we considered the possibility that bound RNA might interfere with dimerization and/or DNA binding. To test this possibility, we pre-incubated AID with RNaseA for 30 minutes before adding the L-oligo, and then imaged as usual (Figure 3.4B). To assure that AID was active under these conditions, the deamination activity of AID when preincubated with RNaseA for 30 minutes versus adding RNaseA at the start of the deamination assay was tested. As seen previously, RNaseA was required for deamination to occur, however, pre-incubation with RNaseA did not change the rate of deamination.

The concentration of AID was 20-fold higher than that of RNaseA so that RNaseA would only be a minor population (5 %) of the protein in the AFM images. The distribution observed for AID that has been pre-treated with RNaseA and then incubated with L-oligo is indistinguishable from that of AID that was not pre-treated with RNaseA and incubated with the L-oligo (Figure 3.4C, grey bars and data not shown). The major peak in the distribution is found at ~18 nm^3. This peak, while still corresponding to monomeric AID, is shifted to a slightly higher volume compared to AID in the absence of oligonucleotide (Figure 3.4C, compare black bars to grey bars). This increase in volume in the presence of oligonucleotide is consistently seen for both AID treated and untreated with RNaseA, and it likely represents oligonucleotide being bound to monomeric AID. These results suggest that binding of DNA to AID does not induce the dimerization of AID, since there is no significant increase in the fraction of peaks with volumes corresponding to the dimer (~45 nm^3) in the presence of DNA substrate.

AID can deaminate single-stranded DNA as a monomer

To determine if AID can function as a monomer, we measured the rates of deamination of the L-oligo substrate (100 nM) at low concentrations of AID (5 nM and 20nM), where our AFM data indicate that AID is primarily monomeric (Fig. 3.4C). For 5 nM and 20 nM AID with 100 nM oligonucleotide, the initial rates (determined by linear fits to the data for timepoints ≤ 30 minutes) are 0.12 min^{-1} and 0.54 min^{-1}, respectively (Figure 3.5a). This increase in rate (4.5-fold) is within error of the expected increase for the 4-fold increase in enzyme concentration (Segel). These results strongly suggest that AID can function as a monomer to deaminate cytosine in single-stranded DNA.

To assure that this observed activity was due the monomer and not to a small population of dimers seen in the AFM images (less than 20% of the population has volumes that could be attributed to the dimer in the presence of oligo), we examined the rates of deamination of 10 nM L-oligo with 2-fold (20 nM) and 50-fold (500 nM) excess AID. If a small fraction of AID dimers were responsible for the observed activity, we would expect to see an increase in deamination rate with increasing AID concentration from 20 nM to 500 nM, because at 20 nM AID, the dimer concentration (20% of the population corresponds to ~ 3nM AID dimer) would be well below the DNA concentration. Consequently, if the dimer were responsible for the observed activity, we should see and increase in deamination activity upon increasing the AID concentration from 20 to 500 nM, as we observed for the experiments with the DNA in excess (Figure 3.5a). Inspection of Figure 3.5b reveals that the rate of deamination is the same for both of these concentrations. These results indicate that at 10 nM L-oligo, a two-fold excess of AID is sufficient to bind all of the L-oligo and catalyze deamination. This result is consistent with the measured binding affinities of AID for oligonucleotide substrates, which are in the sub to low nanomolar range (Larijani, Petrov et al.). Taking these kinetic data together with the AFM data strongly support the conclusion that AID is functional as a monomer on single-stranded DNA *in vitro*.

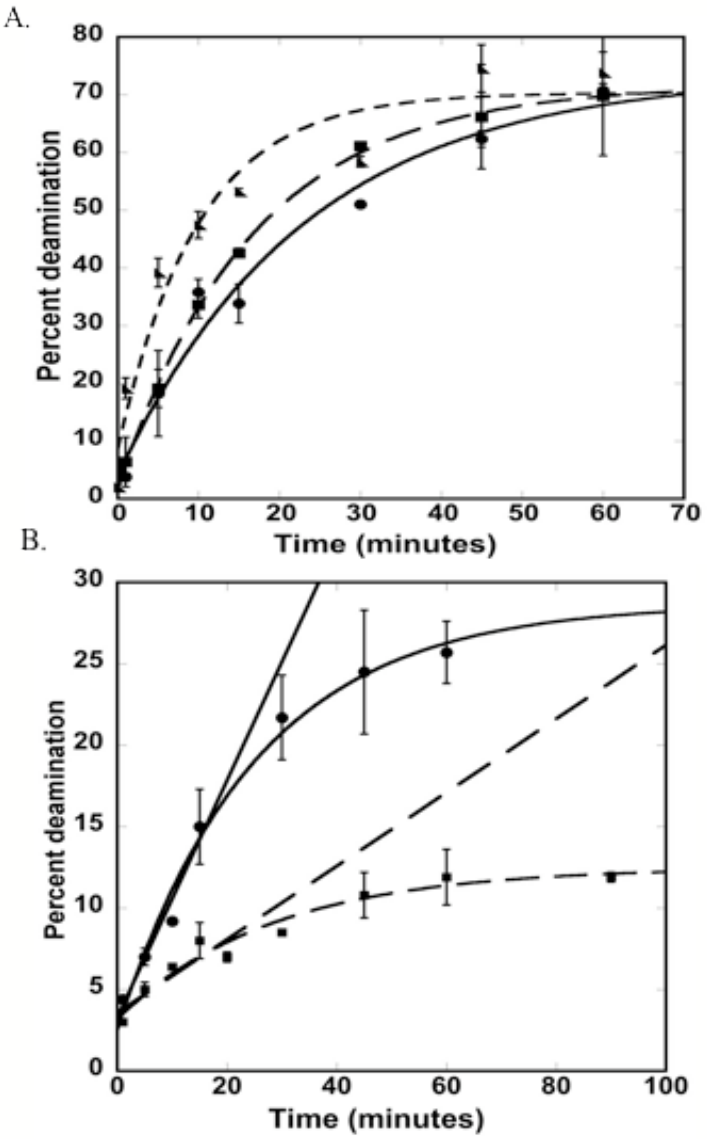

A.
Percent deamination
Time (minutes)
B.
Percent deamination
Time (minutes)

Figure 3.5

Figure 3.5: Deamination rate when AID is in limiting (A) or excess (B)

These data indicate that most of the catalysis on single-stranded DNA oligonucleotide is from the monomer. A) Rate of deamination of 100 nM L-oligo in the presence of 5 nM (squares) and 20 nM (circles) AID. The straight lines are linear fits to data points for times ≤ 30 minutes. The initial rates were determined from the linear fits. B) Rate of deamination of 10 nM L-oligo in the presence of 20 nM (squares) and 500 nM (circles) AID. The curves are first order exponential fits to the data. The error bars represent the standard error from 3 or more experiments. (experiments done by S. Brar)

APOBEC2

APOBEC2 exists primarily as a dimer

To investigate the oligomerization state of APOBEC2, we imaged APOBEC2 at concentrations of 40 nM (Figure 3.6A). Volume analysis done on the AFM scans at 40 nM show that APOBEC2 is primarily a dimer (Figure 3.6B). The major peak from these data is at ~30 nm^3, which is consistent with the molecular weight of a dimer (36.4 kDa) of the APOBEC2 fragment used in these studies. Our AFM results are in contrast to the crystal structure of APOBEC2, which showed APOBEC2 as a long, rod-shaped tetramer. Had APOBEC2 existed as a long, rod-shaped tetramer, it would have been visible in the AFM scans in one of two ways. The first is if the protein were "standing on end" on the mica. If this had been occurring, the protein peaks would have been both very tall and narrow. The other possibility is if the rod-shaped tetrameric protein were "laying down on its side", in which case the proteins observed would have been short, but wide in one dimension. Since neither of these theoretical forms were observed, nor were volumes consistent with that of a tetramer, we can conclude that the protein does not exist as a tetramer under the conditions of these AFM experiments.

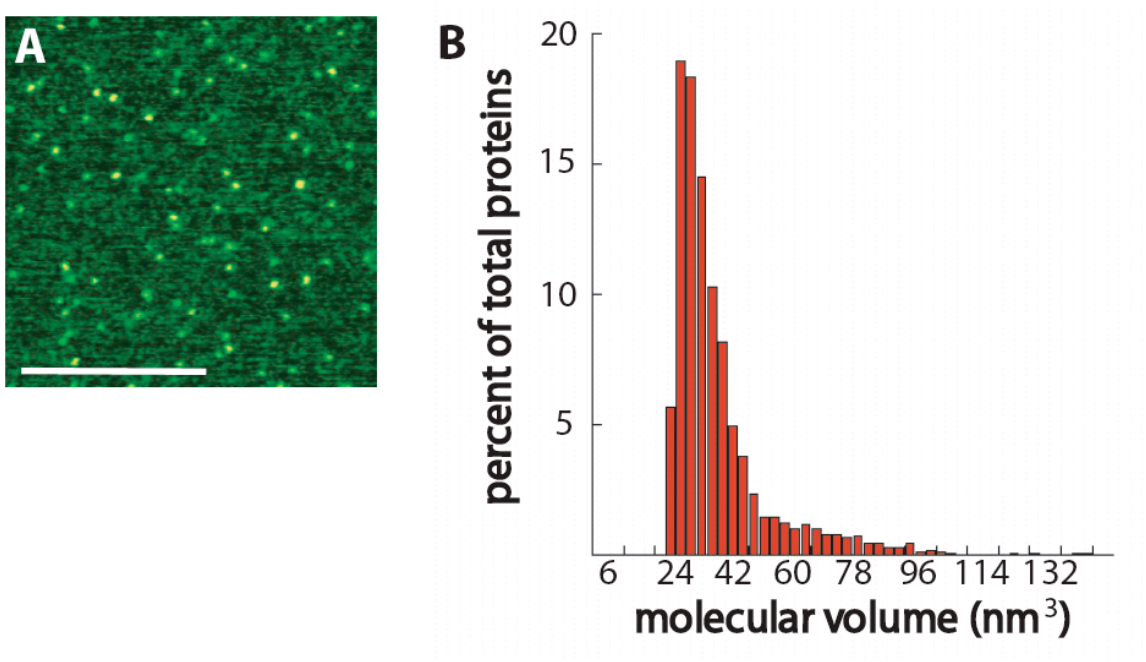


Figure 3.6: AFM data on APOBEC2.

(A) Representative AFM image of APOBEC2. AFM image is 500 x 500 nm, scale bar (white) represents 250 nm. (B) volume analysis of APOBEC2

APOBEC3G

APOBEC3G exists as a multimer of dimers

To investigate the oligomeric state of APOBEC3G, we imaged the protein at a concentration of 20nM in the presence and absence of a ssDNA substrate containing two sequences that APOBEC3G preferentially acts upon. APOBEC3G exists as multimers of dimers (Figure 3.7A & 3.7B). In the absence of DNA, the major volume peak is centered at ~75 nm^3, which correlates with the predicted volume of an APOBEC3G dimer (90.9 nm^3; predicted volume of monomer = 38.1 nm^3). The other peaks seen in the volume analysis (Figure 3.7B) correlate well with the predicted volumes of APOBEC3G tetramers (196.5 nm^3), hexamers (302.1 nm^3) and octamers (407.7 nm^3).

In the presence of ssDNA (Figure 3.7C), APOBEC3G seems to follow a similar trend as it has in the absence of ssDNA, except the peaks in volume analysis have been shifted to slightly higher volumes (Figure 3.7D), presumably because of the binding of ssDNA substrate.

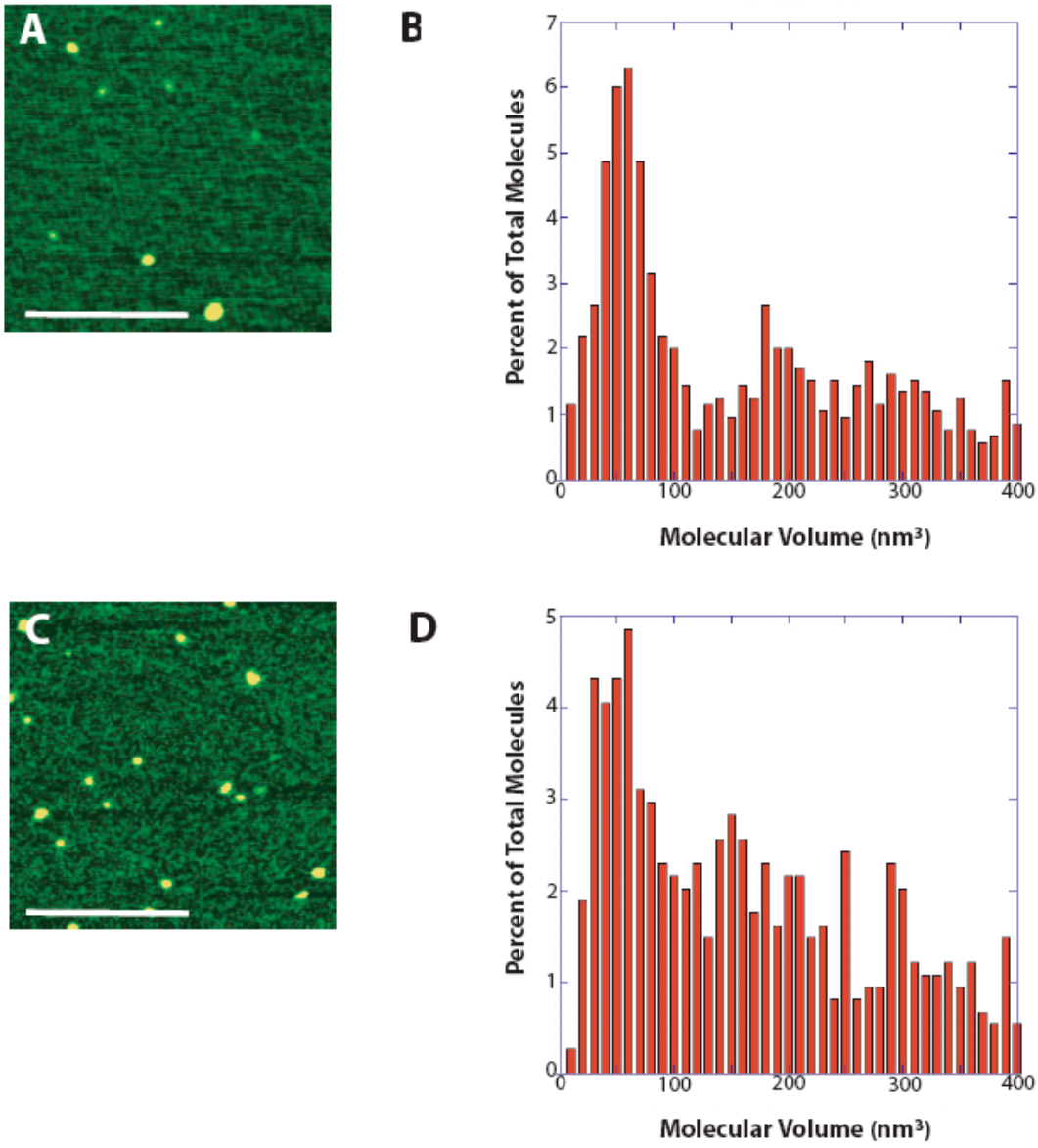

A
B
Percent of Total Molecules
Molecular Volume (nm³)
C
D
Percent of Total Molecules
Molecular Volume (nm³)

Figure 3.7

Figure 3.7: AFM data on APOBEC3G in the presence and absence of ssDNA substrate. All AFM images are 500 x 500 nm; scale bars (white) represent 250 nm. (A) Representative AFM scan of APOBEC3G in the absence of ssDNA substrate. (B) volume analysis of APOBEC3G in the absence of DNA substrate. (C) Representative AFM scan of APOBEC3G in the presence of equimolar ssDNA substrate (D) volume analysis of APOBEC3G in the presence of ssDNA substrate.

Discussion

AID

We found that AID is catalytically active as a monomer on single-stranded DNA *in vitro*, even when the single-stranded DNA region is embedded in double-stranded DNA in the form of loops. We did not observe a significant population of volumes consistent with AID dimers by AFM for any of the protein/oligo concentrations, including concentrations where well over 50% of the oligo is deaminated (Figure 3.4C, data not shown). Pre-incubation of AID at concentrations up to 500nM also failed to alter the percentage of volumes corresponding to the dimer. In addition, even when we examined deamination at protein concentrations ranging from exceedingly low (barely detectable deamination) to in large excess over oligonucleotide concentrations, the rate of deamination by AID on single-stranded DNA substrates was not substantially changed with increasing AID concentration (Figure 3.6). While these results demonstrate that AID is active as a monomer on single-stranded DNA *in vitro*, they do not preclude the possibility that AID may also function as a dimer *in vivo* where it may be post-translationally modified via phosphorylation (Basu, Chaudhuri et al.).

Comparison of the rates of deamination on single-stranded DNA loops within double-stranded DNA with deamination rates on substrates devoid of any secondary structure such as the L-oligo indicates that even short loops of single-stranded DNA can be deaminated by AID relatively well, albeit with less efficiency than DNA sequences in which the cytidine is expected to be completely exposed, such as the L-oligo (Figures 3.2 & 3.3). These results imply that even at low concentrations, as found in the nucleus of

cells expressing AID, monomeric AID could catalyze the deamination of cytidine to uracil as long as it can find a suitable substrate in the form of exposed single-stranded DNA patches or loops. In support of this idea, recent data suggests that AID can bind and deaminate single-stranded DNA in a variety of configurations such as displaced single-stranded DNA loops or patches (Duquette, Pham et al.; Yu, Roy et al.; Larijani, Petrov et al.; Ronai, Iglesias-Ussel et al.), and even in negatively supercoiled double-stranded DNA (Shen and Storb) formed during transcription. These results might explain why AID can access and deaminate DNA in non-natural substrates such as in DNA of *E. coli* cells, non-lymphoid cells, and non-Ig genes that are actively transcribed (Martin and Scharff 2002; Petersen-Mahrt, Harris et al. 2002; Yoshikawa, Okazaki et al. 2002; Okazaki, Hiai et al. 2003), and wherein B-cell specific post-translational modifications of AID do not occur. If for example, dimerization is optimally promoted when AID is specifically phosphorylated in activated B cells, these results imply that non-phosphorylated monomeric AID in other cells could potentiate inappropriate deamination of dC in non-immunoglobulin gene targets. It will be interesting to determine whether monomeric AID contributes to non-targeted AID-mediated deamination in B cell lymphomas or when expressed in non-B cells. If so, the oligomeric status of AID may contribute yet another layer of regulation similar to transcriptional regulation and intracellular localization that keep cells expressing AID from experiencing its mutagenic activity in non-intended targets (Brar, Watson et al.; Ito, Nagaoka et al.; McBride, Barreto et al.). In addition it is likely that, in B cells, targeting co-factors sharpen AID's substrate specificity uniquely to the Ig V and CSR regions (Shinkura, Ito et al.), although these putative co-factors have yet to be identified.

Our results do not imply that the dimer does not form or is not functional. In the AFM images, a few percent of the molecules have volumes that correspond to the dimer and other higher order oligomeric states; increasing the concentration of AID from 5 nM to 20 nM, or depositing the sample following incubation at 500nM, does not result in any detectable changes in the population of molecules with higher volumes. However, It is clear from our studies that *in vitro* that: a) the monomer is the predominant species in the absence of substrate, b) the monomer is catalytically active on single-stranded DNA, and c) binding to single-stranded DNA substrate does not induce any significant dimerization.

A recent study that examined the binding of GST-tagged and His-tagged AID supports our conclusion that AID can function as a monomer. Specifically, in band-shift assays of AID with a DNA substrate containing a single-stranded bubble, the GST-tagged-AID, which is forced to be a dimer by GST dimerization (Dirr, Reinemer et al.), causes dramatically greater shifts of the DNA than does the His-tagged-AID, consistent with His-tagged-AID binding to DNA as a monomer.

A recent report by Wang and colleagues, suggests that AID dimerization is required for optimal class switch recombination *in vivo*; however, these studies examined co-immunoprecipitation of AID with two different tags from cell extracts, and therefore, the apparent dimerization in these studies could be dimerization mediated by other proteins and not a direct interaction of two AID molecules (Wang, Shinkura et al.). In addition, whether or not this "dimerization" is also required for immunoglobulin hypermutation where deamination of dC's throughout the length of the V region is evident was not tested. Several other reasons may account for the differences between our results and those reported in the previous study (Wang, Shinkura et al.). Namely, we have

found that in its most proximal activity, as a deaminase, monomeric AID can deaminate dC's *in vitro*, whereas class switch recombination *in vivo*, may require a higher order oligomeric state of AID to help align specific dC's in switch regions or for other functionalities. Also, it is possible that AID acts both as a monomer and as a dimer *in vivo*, perhaps differentially for hypermutation and class switch recombination; a hypermutation phenotype was not reported in that study for the dimer mutants. While both of these processes require AID, they also likely require different co-factors, occur independently of each other, and target different regions of the heavy chain Ig locus. Finally, it is conceivable that phosphorylation or other post-translational modifications of AID in activated B cells may promote dimerization, while the non-modified form of AID acts as a monomer, with less efficiency. In fact, ectopically expressed AID is not phosphorylated and yet it can deaminate dC in non-lymphoid cells, but at a lower frequency (Martin and Scharff 2002; Petersen-Mahrt, Harris et al. 2002; Yoshikawa, Okazaki et al. 2002; Okazaki, Hiai et al. 2003; Chaudhuri, Khuong et al. 2004).

The observation that AID exhibits catalytic activity as a monomer *in vitro* is in contrast to other deaminases studied to date, including the RNA editing enzymes such as Apobec-1, and the adenosine deaminases (Lau, Zhu et al. 1994; Cho, Yang et al. 2003; Gallo, Keegan et al. 2003; Ko, Lin et al. 2003). Interestingly, these deaminases require dimerization not only for their enzymatic activity but also for specificity of the substrate. For example, homodimers comprising a wild type monomer and a mutant monomer of ADAR1 or ADAR2 lose only half the ability to deaminate a non-specific dsRNA substrate, in a sequence-independent manner but lose nearly 70% site-selective activity on natural substrates, such as the RNA encoding serotonin 2C subtype receptor (Cho,

Yang et al.). Consequently, it is possible that while monomeric AID is catalytically active, a homodimer may confer some kind of substrate specificity wherein interaction between the monomers helps align a specific cytosine residue for deamination. This idea is interesting because deamination of cytosine in an RNA target would require exquisite targeting not only to a particular RNA species, but a particular cytosine in its sequence. Such a level of specificity may require dimerization as seen with RNA editing enzymes such as Apobec-1 (Lau, Zhu et al.). However, an RNA target for AID has not been identified although indirect evidence supports the possibility (Honjo, Muramatsu et al.).

Alternatively, an AID homodimer, while not required for deamination, may help the simultaneous deamination of cytosine on both DNA strands, as previously suggested (Bransteitter, Pham et al.), an idea supported by our findings predicting that each monomer of AID in a putative dimer can catalyze deamination of single-stranded DNA independently. Additional support for this notion comes from structural analysis of AID suggesting that for the homodimer, both monomers are able to catalyze deamination independently on a double-stranded DNA or RNA substrate (Xie, Sowden et al.). In addition, a recent structure of the AID homolog APOBEC2 shows that while APOBEC2 is a dimer of dimers in the crystal structure, the active site is not near the dimer interface (Prochnow, Bransteitter et al.). Finally, it is important to note that although it is likely that AID can form a dimer, there is no direct evidence for dimerization *in vitro* or *in vivo*, only inference from homology models and co-immunoprecipitation of cell extracts (Xie, Sowden et al. 2004; Wang, Shinkura et al. 2006).

APOBEC2 & APOBEC3G Oligomeric States

The clear variation in oligomeric states for the three AID/APOBEC family members investigated here suggests that their diverse functionality is mediated, in part, by the oligomeric state of the protein. For example, it has been shown that APOBEC3G, exists as either a "high molecular weight species" (~700 kDa) or a "low molecular weight species" (~46-100 kDa) in CD4+ T cells (Chiu, Soros et al. 2005). In unactivated CD4+ T cells, which are resistant to HIV infection, APOBEC3G is present in its enzymatically active, low molecular weight form; activated CD4+ T cells, which are susceptible to HIV infection, have APOBEC3G in the enzymatically inactive high molecular weight form (Chiu, Soros et al. 2005). Since it is possible to convert the high molecular weight form of APOBEC3G to the low molecular weight form of APOBEC3G by treatment with RNase (Chiu, Soros et al. 2005), it follows that the oligomeric state of APOBEC3G, and likely any of the AID/APOBEC family members is regulated by the presence of oligonucleotide.

Our AFM studies of APOBEC2 reveal that at lower concentrations of protein (at least lower than the concentration required for crystallization), the protein exists as a dimer. Based on the crystal structure, this suggests that the dimer-dimer interface that forms the tetramer is weak (Figure 3.1B). Since APOBEC2 and AID have 44.6% sequence homology, one might expect that they would have similar oligomerization states. Since they don't, it is possible that the sequence homology is between residues buried at the core of the protein. This idea is supported by the fact that both proteins had different affinities for the mica surface (it was difficult to go over 20 nM AID protein

because it would coat the surface of the mica while it was possible to get up to 120 nM APOBEC2).

APOBEC3G is seen as a multimer of dimers; forming tetrameric, hexameric and even octameric species in solution and in the presence of equimolar ssDNA substrate. Since APOBEC3G is known to interact with viral proteins, it seems likely that one of these observed oligomeric states is preferential partner for cellular or viral proteins. Perhaps, suggested by the work done with activated and unactivated CD4+ T cells (Chiu, Soros et al. 2005), the formation of each oligomeric state is partially mediated by the interaction of APOBEC3G with either cellular or viral factors.

Future Directions

Since phosphorylation of AID seems to play a role in controlling the deamination activity of AID (Basu, Chaudhuri et al. 2005), it would be interesting to image the phosphorylated protein to see if the phosphorylation makes a difference in oligomerization state of AID in the presence and absence of ssDNA substrate. It would also be interesting to have a "bubble" substrate more amenable to AFM imaging (ie: long enough to be seen on its own by AFM; >400 nm) to look at the binding / localization of AID on the ssDNA region of the substrate.

Despite the lack of evidence of functionality, it would be interesting to investigate the oligomeric state of APOBEC2 in the presence of varying substrates, including ssDNA, RNA and free cytidine, to see if any of these possible substrates change the oligomerization state. Perhaps by simply going on a "fishing expedition" it may be possible to determine if any of these possible substrates have an effect on the

dimerization state of APOBEC2. However, lack of any results would not rule out binding of these substrates to APOBEC2, so these studies should be carried out in conjunction with other biochemical techniques.

Since APOBEC3G seems to have a preference for forming multimers of dimers, it would be interesting to make DNA where there is a double-stranded section that separates the two binding sites (-CCC-). This DNA substrate would allow us to see if DNA is being wrapped around protein to form the multimeric states – if the DNA is wrapping around the APOBEC3G, the decrease in DNA flexibility caused by the dsDNA region would likely discourage the wrapping of the DNA. Instead, there would likely be two proteins sitting next to each other if both binding sites are bound (kind of a dumbbell shaped molecule).

Materials and Methods

AID, APOBEC2 & APOBEC3G

Purified human AID was purchased from Enzymax (University of Kentucky). The proteins APOBEC2 (residues 41-224) and APOBEC3G were provided by Myron Goodman (USC).

DNA substrates

AID

The oligonucleotides used in the deamination assays and the predicted structures of the double-stranded oligonucleotides are depicted in Figure 3.2. The "L-oligo" was designed based on the primer used by Yu and colleagues (Yu, Huang et al.), while the

single-stranded DNA oligonucleotide (AID1 primer) was based on the CDR2 sequence of the human Ig heavy chain VH3-30-3. Oligonucleotides (100 μM) were 5'-end labeled with γ-^{32}P ATP (3000 Ci/mmol, Amersham Biosciences) using Optikinase (United States Biochemical Corp.) according to the manufacturer's instructions. The reaction was terminated by the addition of 20 mM EDTA, and the enzyme was heat denatured by incubation for 10 minutes at 65°C. After labeling, the free nucleotides were removed by gel filtration chromatography (Micro Bio-spin-6, Bio-Rad). For the generation of double-stranded substrates, labeled oligonucleotide was annealed with the complimentary strand ((AID1+AID2 = double strand), (AID1+AID8 = Bubble), and (AID1+AID9 = Loop) see Fig. 3.2) using 1.2 molar excess of cold oligonucleotides in the presence of 100 mM KCl, heated at 65°C for 5 minutes and slowly cooled to room temperature. The duplexes were analyzed on a native 10% polyacrylamide gel.

APOBEC3G

DNA substrate used with APOBEC3G (sequence: 5'-GTA TAT ATG TTG AGA CCC GTA GTA ATG AGA GAT TGA TTA GAT TAG TTT AAT GTG ATA TAT GTG TAT GAA AGA TAT AAG ACC CAG AGA GTA AAG TTG TTA AAT TTG TGT AGA TAT GTT (dT-Biotin)A-3') was provided by Myron Goodman (USC).

Atomic Force Microscopy

All AFM images were captured in air using either a Nanoscope III or IIIa (Digital Instruments, Santa Barbara, CA) microscope in tapping mode. Pointprobe Plus tapping mode silicon probes (Molecular Imaging, Tempe, AZ) with resonance frequencies of approximately 170 kHz were used for imaging. Images were collected at a speed of 2-3

Hz with an image size of 1 μm x 1 μm at 512 x 512 pixel resolution. Each experiment was repeated at least twice. Volume analysis was done as previously described (Ratcliff and Erie 2001; Yang, Wang et al. 2003). The program Kaleidagraph (Synergy Software, Reading PA) was used to generate statistical plots.

AID

H100-OAc imaging buffer (20 mM HEPES, pH 7.3; 10 mM MgOAc; 100mM NaOAc) was heated to 65°C and allowed to cool slowly. AID protein was diluted to final concentrations of 5, 10, or 20 nM in H100-OAc imaging buffer. For depositions in the presence of DNA, AID was incubated with various concentrations of DNA oligonucleotides (5 nM-50 nM) for 30 minutes at room temperature prior to deposition. After incubation, 10 μl of the AID sample (either in the presence or absence of DNA) was deposited onto freshly cleaved mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water, excess water was wicked from the surface using filter paper, and the surface was then dried using a stream of nitrogen.

APOBEC2

A3G imaging buffer (50 mM HEPES, pH 7.3; 5 mM $MgCl_2$) was heated to 65ºC and allowed to cool slowly. APOBEC2 was then diluted to 40, 80 or 120 nM in A3G imaging buffer and 1 mM DTT was added to the solution. Ten microliters of the protein sample was then deposited onto freshly cleaved mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water, excess water was wicked from the surface using filter paper, and the surface was then dried using a stream of nitrogen.

APOBEC3G

A3G imaging buffer (50 mM HEPES, pH 7.3; 5 mM $MgCl_2$) was heated to 65°C and allowed to cool slowly. APOBEC3G was then diluted to 20 - 50 nM in A3G imaging buffer, and 1 mM DTT was added to the sample. For depositions in the presence of DNA, APOBEC3G was incubated with an equimolar concentration of substrate DNA for 5 minutes before deposition. Ten microliters of the protein sample were then deposited onto freshly cleaved mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water, excess water was wicked from the surface using filter paper, and the surface was then dried using a stream of nitrogen.

Deamination assays

Typically, a reaction mixture (10 µl) containing 2-100 nM radioisotope-labeled oligonucleotide substrate, 5nM-2000nM purified human AID protein (Enzymax, University of Kentucky), 100ng of RNase A (Qiagen) and 1 unit of uracil DNA glycosylase (UDG, Invitrogen) was incubated for 15 min at 37 °C in a buffer containing: 25 mM Tris, pH 8.0, 50 mM NaCl, and 5 mM EDTA (see scheme in Fig. 3.3a).

To examine deamination at limiting and excess AID protein concentrations, deamination reactions were carried out at final AID protein concentrations ranging from 5 nM to 2 µM and oligo concentrations ranging from 2 nM to 100 nM. For a subset of samples, RNase A (400 ng) (Qiagen) was pre-incubated with AID for 30 minutes prior to addition of the oligo. The reaction was stopped by adding 1 ml of 2 M NaOH and heated for 5 min at 95° C. Ten mL of formamide was added and samples were heated at 95° C for 5 min and loaded on 10% denaturing polyacrylamide gels containing 7 M urea. Gels were run in 1X Tris borate-EDTA (TBE) buffer under constant voltage at ambient temperature.

Gels were dried and visualized by autoradiography using the phosphor-imager Typhoon 9400 (Amersham Biosciences) and the bands quantified by ImageQuant software (Molecular Dynamics). Deamination rates were determined by fitting the percentage of deamination with time to first order exponentials. All deamination reactions were repeated at least two times.

References

Arakawa, H., J. Hauschild, et al. (2002). "Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion." Science **295**(5558): 1301-6.

Bao, K. K., H. Wang, et al. (2003). "Functional oligomeric state of avian sarcoma virus integrase." J Biol Chem **278**(2): 1323-7.

Basu, U., J. Chaudhuri, et al. (2005). "The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation." Nature **438**(7067): 508-11.

Beale, R. C., S. K. Petersen-Mahrt, et al. (2004). "Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo." J Mol Biol **337**(3): 585-96.

Besmer, E., E. Market, et al. (2006). "The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination." Mol Cell Biol **26**(11): 4378-85.

Bransteitter, R., P. Pham, et al. (2004). "Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase." J Biol Chem **279**(49): 51612-21.

Bransteitter, R., P. Pham, et al. (2003). "Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase." Proc Natl Acad Sci U S A **100**(7): 4102-7.

Brar, S. S., M. Watson, et al. (2004). "Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks." J Biol Chem **279**(25): 26395-401.

Chaudhuri, J., C. Khuong, et al. (2004). "Replication protein A interacts with AID to promote deamination of somatic hypermutation targets." Nature **430**(7003): 992-8.

Chelico, L., P. Pham, et al. (2006). "APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA." Nat Struct Mol Biol **13**(5): 392-9.

Chiu, Y. L., V. B. Soros, et al. (2005). "Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells." Nature **435**(7038): 108-14.

Cho, D. S., W. Yang, et al. (2003). "Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA." J Biol Chem **278**(19): 17093-102.

Chou, S. H., K. H. Chin, et al. (2003). "Unusual DNA duplex and hairpin motifs." Nucleic Acids Res **31**(10): 2461-74.

Dickerson, S. K., E. Market, et al. (2003). "AID mediates hypermutation by deaminating single stranded DNA." J Exp Med **197**(10): 1291-6.

Dirr, H., P. Reinemer, et al. (1994). "X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function." Eur J Biochem **220**(3): 645-61.

Duquette, M. L., P. Pham, et al. (2005). "AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation." Oncogene **24**(38): 5791-8.

Gallo, A., L. P. Keegan, et al. (2003). "An ADAR that edits transcripts encoding ion channel subunits functions as a dimer." Embo J **22**(13): 3421-30.

Harris, R. S., J. E. Sale, et al. (2002). "AID is essential for immunoglobulin V gene conversion in a cultured B cell line." Curr Biol **12**(5): 435-8.

Honjo, T., M. Muramatsu, et al. (2004). "AID: how does it aid antibody diversity?" Immunity **20**(6): 659-68.

Ito, S., H. Nagaoka, et al. (2004). "Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1." Proc Natl Acad Sci U S A **101**(7): 1975-80.

Ko, T. P., J. J. Lin, et al. (2003). "Crystal structure of yeast cytosine deaminase. Insights into enzyme mechanism and evolution." J Biol Chem **278**(21): 19111-7.

Larijani, M., A. P. Petrov, et al. (2007). "AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner." Mol Cell Biol **27**(1): 20-30.

Lau, P. P., H. J. Zhu, et al. (1994). "Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene." Proc Natl Acad Sci U S A **91**(18): 8522-6.

Liao, W., S. H. Hong, et al. (1999). "APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family." Biochem Biophys Res Commun **260**(2): 398-404.

Martin, A. and M. D. Scharff (2002). "Somatic hypermutation of the AID transgene in B and non-B cells." Proc Natl Acad Sci U S A **99**(19): 12304-8.

McBride, K. M., V. Barreto, et al. (2004). "Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase." J Exp Med **199**(9): 1235-44.

Muramatsu, M., K. Kinoshita, et al. (2000). "Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme." Cell **102**(5): 553-63.

Nguyen, D. H., S. Gummuluru, et al. (2007). "Deamination-independent inhibition of hepatitis B virus reverse transcription by APOBEC3G." J Virol **81**(9): 4465-72.

Noguchi, C., N. Hiraga, et al. (2007). "Dual effect of APOBEC3G on Hepatitis B virus." J Gen Virol **88**(Pt 2): 432-40.

Okazaki, I. M., H. Hiai, et al. (2003). "Constitutive expression of AID leads to tumorigenesis." J Exp Med **197**(9): 1173-81.

Petersen-Mahrt, S. K., R. S. Harris, et al. (2002). "AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification." Nature **418**(6893): 99-103.

Pham, P., R. Bransteitter, et al. (2003). "Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation." Nature **424**(6944): 103-7.

Prochnow, C., R. Bransteitter, et al. (2007). "The APOBEC-2 crystal structure and functional implications for the deaminase AID." Nature **445**(7126): 447-51.

Rada, C., G. T. Williams, et al. (2002). "Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice." Curr Biol **12**(20): 1748-55.

Ratcliff, G. C. and D. A. Erie (2001). "A novel single-molecule study to determine protein--protein association constants." J Am Chem Soc **123**(24): 5632-5.

Revy, P., T. Muto, et al. (2000). "Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2)." Cell **102**(5): 565-75.

Rogozin, I. B. and M. Diaz (2004). "Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process." J Immunol **172**(6): 3382-4.

Ronai, D., M. D. Iglesias-Ussel, et al. (2007). "Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation." J Exp Med **204**(1): 181-90.

Segel, I. H. (1975). Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems. New York, Wiley.

Sheehy, A. M., N. C. Gaddis, et al. (2002). "Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein." Nature **418**(6898): 646-50.

Shen, H. M. and U. Storb (2004). "Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled." Proc Natl Acad Sci U S A **101**(35): 12997-3002.

Shinkura, R., S. Ito, et al. (2004). "Separate domains of AID are required for somatic hypermutation and class-switch recombination." Nat Immunol **5**(7): 707-12.

Sohail, A., J. Klapacz, et al. (2003). "Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations." Nucleic Acids Res **31**(12): 2990-4.

Ta, V. T., H. Nagaoka, et al. (2003). "AID mutant analyses indicate requirement for class-switch-specific cofactors." Nat Immunol **4**(9): 843-8.

Tashiro, J., K. Kinoshita, et al. (2001). "Palindromic but not G-rich sequences are targets of class switch recombination." Int Immunol **13**(4): 495-505.

Tessmer, I., T. Moore, et al. (2005). "AFM studies on the role of the protein RdgC in bacterial DNA recombination." J Mol Biol **350**(2): 254-62.

Wang, J., R. Shinkura, et al. (2006). "Identification of a specific domain required for dimerization of activation-induced cytidine deaminase." J Biol Chem **281**(28): 19115-23.

Xie, K., M. P. Sowden, et al. (2004). "The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1." Proc Natl Acad Sci U S A **101**(21): 8114-9.

Xue, Y., G. C. Ratcliff, et al. (2002). "A minimal exonuclease domain of WRN forms a hexamer on DNA and possesses both 3'- 5' exonuclease and 5'-protruding strand endonuclease activities." Biochemistry **41**(9): 2901-12.

Yang, G., H. Obiakor, et al. (2005). "Activation-induced deaminase cloning, localization, and protein extraction from young VH-mutant rabbit appendix." Proc Natl Acad Sci U S A **102**(47): 17083-8.

Yang, Y., L. E. Sass, et al. (2005). "Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions." Nucleic Acids Res **33**(13): 4322-34.

Yang, Y., H. Wang, et al. (2003). "Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy." Methods **29**(2): 175-87.

Yoshikawa, K., I. M. Okazaki, et al. (2002). "AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts." Science **296**(5575): 2033-6.

Yu, K., F. T. Huang, et al. (2004). "DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine." J Biol Chem **279**(8): 6496-500.

Yu, K., D. Roy, et al. (2005). "Fine-structure analysis of activation-induced deaminase accessibility to class switch region R-loops." Mol Cell Biol **25**(5): 1730-6.

Zhang, K. L., B. Mangeat, et al. (2007). "Model Structure of Human APOBEC3G." PLoS ONE **2**: e378

CHAPTER 4

INTERACTIONS OF SALT, ATP, AND/OR DNA WITH yMUTLα

Introduction

The MMR system is best understood in *E. coli*, where it is carried out by the mismatch repair proteins MutS, MutL and MutH, in addition to a helicase, several exonucleases, and the rest of the replication machinery (for review, see (Au, Welsh et al. 1992; Modrich and Lahue 1996). Both MutS and MutL, which are homodimeric in *E. coli*, have multiple eukaryotic homologs that function in MMR, as well as other cellular processes, as heterodimers.

In eukaryotes, MMR is believed to be carried out by an array of proteins, including the MutS homolog MutSα (Msh2-Msh6), the MutL homolog MutLα (Mlh1-Pms2 (Pms1 in yeast), PCNA, ExoI, RPA, RFC and DNA Polymerase δ (Constantin, Dzantiev et al. 2005; Zhang, Yuan et al. 2005). The hydrolysis of ATP by both MutSα and MutLα is absolutely required for MMR (Iaccarino, Marra et al. 1998; Dufner, Marra et al. 2000; Tran and Liskay 2000; Drotschmann, Yang et al. 2002; Hall, Shcherbakova et al. 2002; Raschle, Dufner et al. 2002; Tomer, Buermeyer et al. 2002; Kijas, Studamire et al. 2003; Lamers, Winterwerp et al. 2003).

Much work has been done to study the ATPase activity of MutSα (Iaccarino, Marra et al. 1998; Dufner, Marra et al. 2000; Drotschmann, Yang et al. 2001; Drotschmann, Yang et al. 2002; Lamers, Winterwerp et al. 2003; Lamers, Georgijevic et al. 2004), and while there have been ATPase studies done on full length *E. coli* MutL, the only work on MutLα has been on isolated N-terminal domains (Ban and Yang 1998; Spampinato and Modrich 2000; Guarne, Junop et al. 2001; Hall, Shcherbakova et al. 2002). Interestingly, an N-terminal fragment of *E. coli* MutL, despite having ADP or the non-hydrolyzable ATP analog AMPPNP bound in the crystal structure (ref), is not able to hydrolyze ATP. In contrast, the N-terminal fragments of MutLα proteins are competent for ATP hydrolysis. A comparison of ATP hydrolysis rates of full-length *E. coli* MutL (with and without an attached His-tag) and for N-terminal fragments MutLα can be found in Table 4.1 (Ban and Yang 1998; Spampinato and Modrich 2000; Guarne, Junop et al. 2001; Hall, Shcherbakova et al. 2002).

Protein	K_m	k_{cat}	***reference***
His-$MutL_{37°C}$	90 mM	0.4 min^{-1}	Ban and Yang 1998
$MutL_{37°C}$	0.41 mM	0.9 min^{-1}	Spampinato and Modrich 2000
$yMlh1_{1-343, 30°C}$	69 mM	0.023 min^{-1}	Hall, Shcherbakova et al. 2002
$yPms1_{32-396, 30°C}$	1500 mM	0.15 min^{-1}	Hall, Shcherbakova et al. 2002
$hPms2_{1-365, rt}$	100 mM	0.2 min^{-1}	Guarne, Junop et al. 2001

Table 4.1: Comparison of ATP hydrolysis rates for full length MutL and N-terminal fragments of MutLα

Reaction temperatures are listed. For the N-terminal fragments, the residue numbers comprising the fragments are listed.

In addition to the information on the ATP binding pocket, the crystal structures of an N-terminal fragment of *E. coli* MutL reveal a putative binding site for single stranded DNA (ssDNA) (Ban, Junop, *et al.* 1999). The proposed binding site is lined with positively charged amino acid residues and is formed by the dimerization of MutL. However, the groove is too narrow to account for the binding of dsDNA to MutL homologues, which has been demonstrated by other researchers (Hall, Wang et al. 2001; Hall, Shcherbakova et al. 2003).

yMutLα does indeed bind to DNA (Hall, Wang et al. 2001) and has preference for long dsDNA over long ssDNA (Hall, Wang et al. 2001). Binding of yMutLα to the DNA occurs mostly in long tracts along either one or two strands of dsDNA (Hall, Wang et al. 2001), which suggests that binding to dsDNA is cooperative. Each N-terminal domain of the proteins that make up MutLα is able to bind to DNA in the absence of the other protein (Hall, Shcherbakova et al. 2003), suggesting that MutLα has two DNA binding sites.

Since yMutLα binding to DNA is sensitive to concentration of salt, it is believed that binding is a result of ionic interactions between regions of positive charge in yMutLα and the negative backbone of DNA (Hall, Wang et al. 2001; Hall, Shcherbakova et al. 2003). Indeed, mutations made to arginine residues in N-terminal fragments of Mlh1 and Pms1, as well as full length Mlh1 (in complex with wild-type Pms1) decreased the binding of the proteins to DNA (Hall, Shcherbakova et al. 2003).

Despite evidence pointing to two DNA binding sites in eukaryotic MutL proteins and the recently discovered endonucleolytic activity of MutLα, DNA binding has yet to be incorporated into any sort of MMR mechanism. Preferential binding of yMutLα to

long (>500 bp) dsDNA indicates positive cooperativity (Hall, Wang et al. 2001). The salt dependence of yMutLα binding to DNA suggests that DNA binding by yMutLα is primarily a result of ionic contacts between the protein and the DNA phosphate backbone (Hall, Wang et al. 2001; Hall, Shcherbakova et al. 2003). However, the mechanism of the formation of yMutLα protein tracts along DNA is unknown, and the relation of the protein tracts to the function of MutLα is largely unknown.

This chapter investigates the effects of Na^+ concentration and ATP on the formation of protein tracts along homoduplex DNA. Also presented are preliminary data on the interaction of yMutLα (at varying ATP concentrations) with nicked plasmid DNA. Additionally, since no work has been done to determine the rates of ATP hydrolysis using full-length MutLα, and since there are such large conformational changes upon adenine nucleotide binding (Chapter 2), including the probable formation of interactions between the N-terminal and C-terminal domains, it seems possible that these conformational changes could have an effect on the rate of ATP hydrolysis. For this reason, the ATP hydrolysis properties of full-length MutLα are also investigated, and presented here in this chapter. From all these data, predictions of the role of the binding of yMutLα to DNA in MMR and other cellular processes can be made.

Results

Effect of Na^+ on yMutLα protein tracts

yMutLα preferentially forms protein tracts along one strand of dsDNA

Previous AFM experiments showed that at low salt (25 mM Na^+) yMutLα bound to dsDNA in a cooperative manner, forming tracts primarily along two strands of dsDNA (2-ds tract, see Figure 4.1 for schematic of possible ways that yMutLα can bind to dsDNA) (Hall, Wang et al. 2001). Reanalysis of these data, as well as analysis of additional data collected under similar conditions, revealed that the DNA actually has more protein tracts along 1 strand of dsDNA (1-ds tract; Figure 4.1 and 4.2). These 1-ds tracts were likely missed because they tend to coat most of the DNA.

The 1-ds tracts were discovered when comparing the height of DNA that was thought to not have protein on it with the height of DNA alone (deposited in the absence of protein). The DNA containing 1-ds tracts had an average height of 0.94 ± 0.09 nm, while the DNA alone had an average height of 0.41 nm. Since this height difference was observed in scans from multiple days and depositions, the only logical explanation for this difference in height is that yMutLα is binding to the exposed DNA and forming 1-ds tracts.

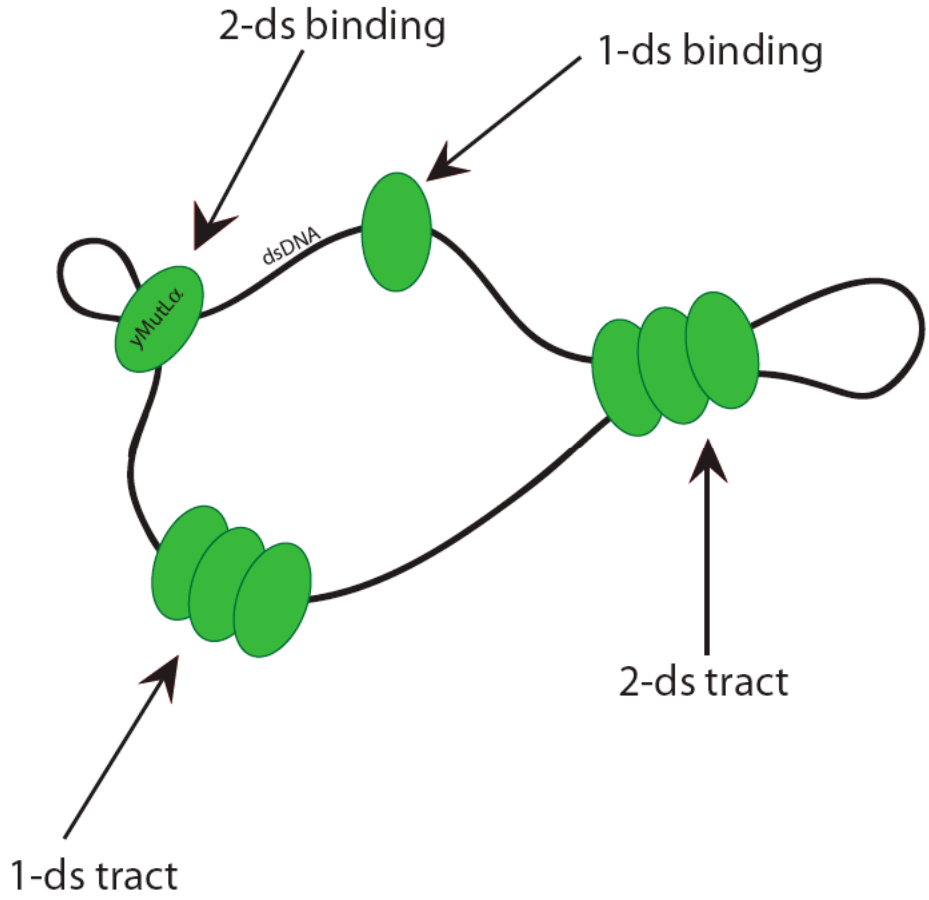


Figure 4.1: Cartoon of the four possible binding modes of yMutLα on dsDNA.

DNA is indicated by black line, yMutLα is representated as green oval. "1-ds binding" represents one yMutLα binding to one strand of dsDNA. "2-ds binding" represents one yMutLα binding to two strands of dsDNA. "1-ds tract" represents multiple yMutLα proteins binding and forming a protein tract along 1 strand of dsDNA. "2-ds tract" represents multiple yMutLα proteins binding and forming a protein tract along 2 strands of dsDNA.

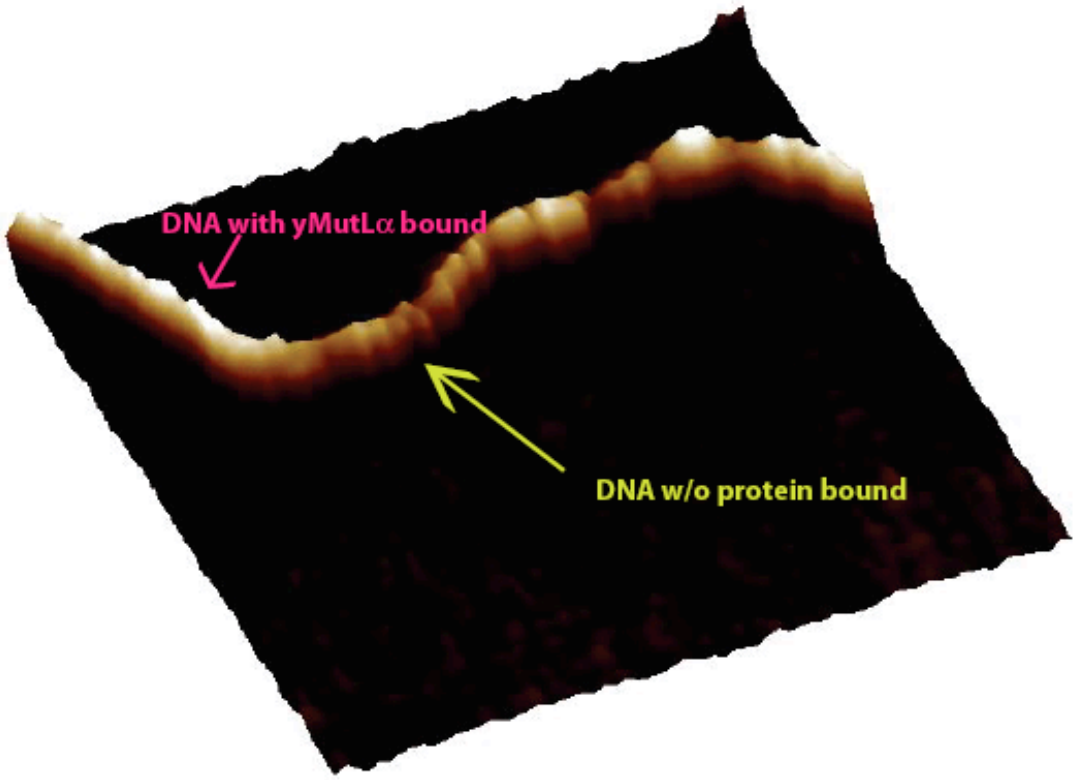


Figure 4.2: 3-D AFM image illustrating 1-ds tract formation.

Pink arrow points to DNA with yMutLα bound (average height of protein-bound DNA is 0.9 ± 0.1 nm); lime arrow points to DNA without yMutLα bound (average height of unbound DNA is 0.4 ± 0.1 nm)

Increasing Na^+ concentration does not affect yMutLα protein tracts along DNA

Filter binding studies of yMutLα to DNA show that at 25 mM Na^+, approximately 70% of DNA is bound by yMutLα. In these studies, at 100 mM Na^+, the percent of bound DNA has dropped to 50%, and at 125 mM Na^+, ~20% of DNA is bound by yMutLα (Hall, Wang et al. 2001). Since electrostatic interactions are believed to be involved with yMutLα binding to DNA, AFM imaging was conducted at 100 and 125 mM Na^+ in addition to the imaging done at 25 mM Na^+ to see if an increase in Na^+ would effect the formation of yMutLα protein tracts.

To reach this end, the potential protein tracts were verified by visual inspection and measurement of height before measuring the length of the tract. A height greater than 0.4 ± 0.1 nm (the average height of DNA alone) along one strand of DNA, coupled with the appearance of protein indicated a 1-ds tract; a height greater than 0.8 ± 0.1 nm (the average height of two overlapping strands of DNA alone) along two strands of DNA coupled with the appearance of protein indicated a 2-ds tract.

At 25 mM Na^+, 78 % of the protein tracts are along 1 strand of dsDNA, and 22% of the protein tracts are along 2 strands of DNA (Figure 4.3A). Despite the fact that there are more 1-ds tracts than 2-ds tracts, both types of protein tracts have similar distributions of length (Figure 4.3B), with an average tract length of 58 nm for 1-ds DNA tracts and 49 nm for 2-ds tracts.

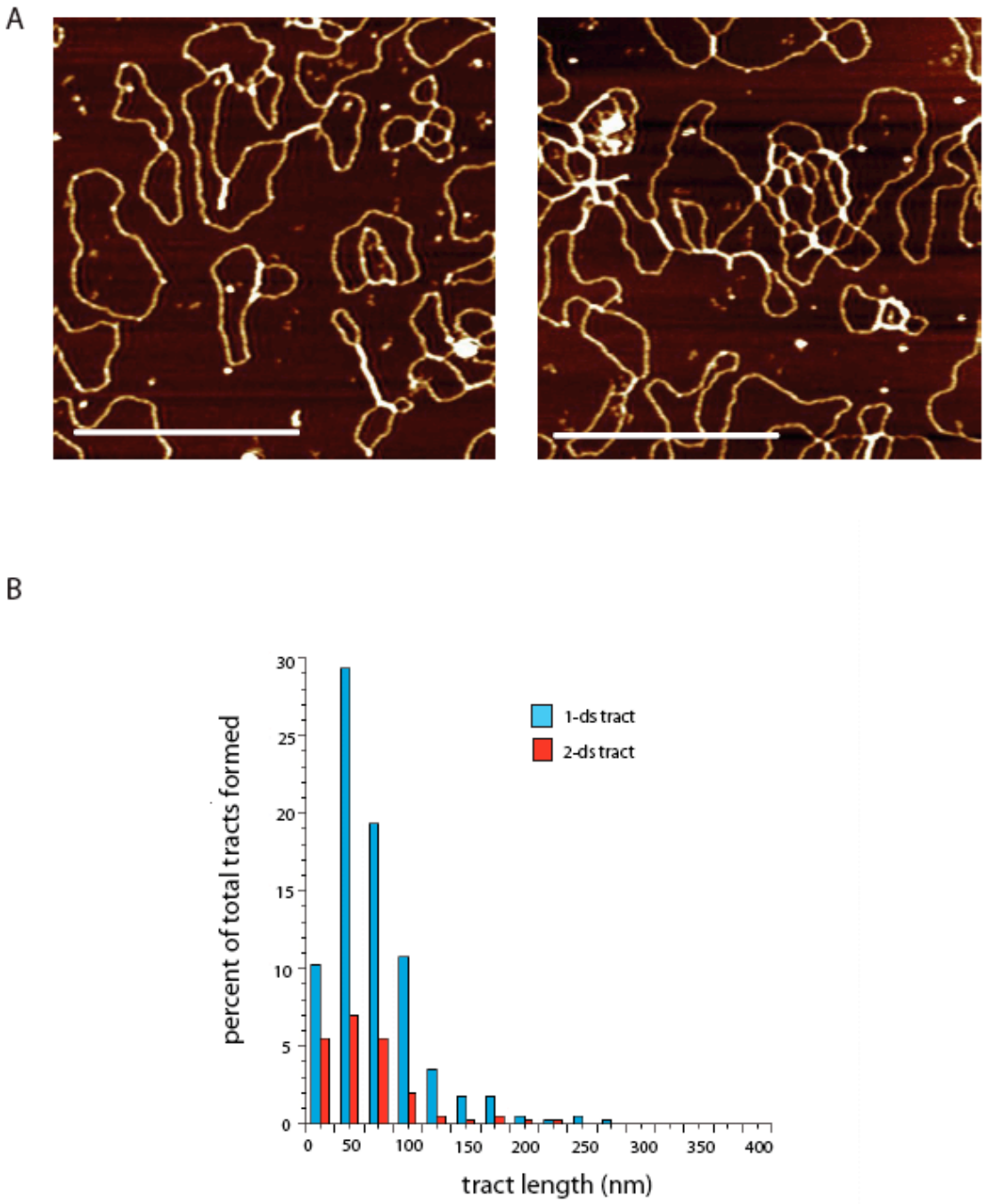


Figure 4.3: yMutLα and DNA in the presence of 25 mM Na^+.

(A) AFM images of yMutLα and DNA at 25 mM Na^+. Images are 1 x 1 μM; scale bar represents 500 nm. (B) distribution of 1-ds (n = 312 tracts) and 2-ds (n = 87) tract lengths. Lengths are plotted as percent of total tracts measured.

It should be noted that these data are potentially biased since only protein-DNA complexes that were entirely in the AFM image were measured. There were many complexes containing yMutLα and multiple-DNAs that were not measured for this reason, and it is possible that the numbers obtained are potentially biased towards 1-ds tracts, since 2-ds tracts are often formed by interactions between strands of DNA from two different DNA molecules.

Surprisingly, the AFM data on yMutLα tract formation along DNA at 100 mM Na^+ and 125 mM Na^+ does not appear to reflect a decrease in protein tract formation as Na^+ concentration increases (Figures 4.4A and 4.5A). Perhaps these differences may be attributed to a difference in buffer (filter binding studies were conducted using Tris buffer; AFM studies presented here are done using HEPES buffer).

At 100 mM Na^+, like at 25 mM Na^+, 78% of protein tracts are 1-ds tracts and 22% of protein tracts are 2-ds tracts (Figure 4.4B). While the average tract length is approximately the same for 2-ds tracts (59 nm) the average length has shifted to 84 nm for the 1-ds tracts. This increase in the length of 1-ds tracts suggests that 100 mM Na^+ helps to facilitate the formation of 1-ds tracts.

At 125 mM Na, 80% of protein tracts are 1-ds tracts and 20% of protein tracts are 2-ds tracts. Despite this slight increase in 1-ds tracts (and slight decrease in 2-ds tracts), the length distributions (Figure 4.5B) are similar to those observed at 100 mM Na^+. The average tract length for 2-ds tracts is 55 nm, and the average tract length for 1-ds tracts is slightly shorter than the tract length at 100 mM Na^+ at 71 nm.

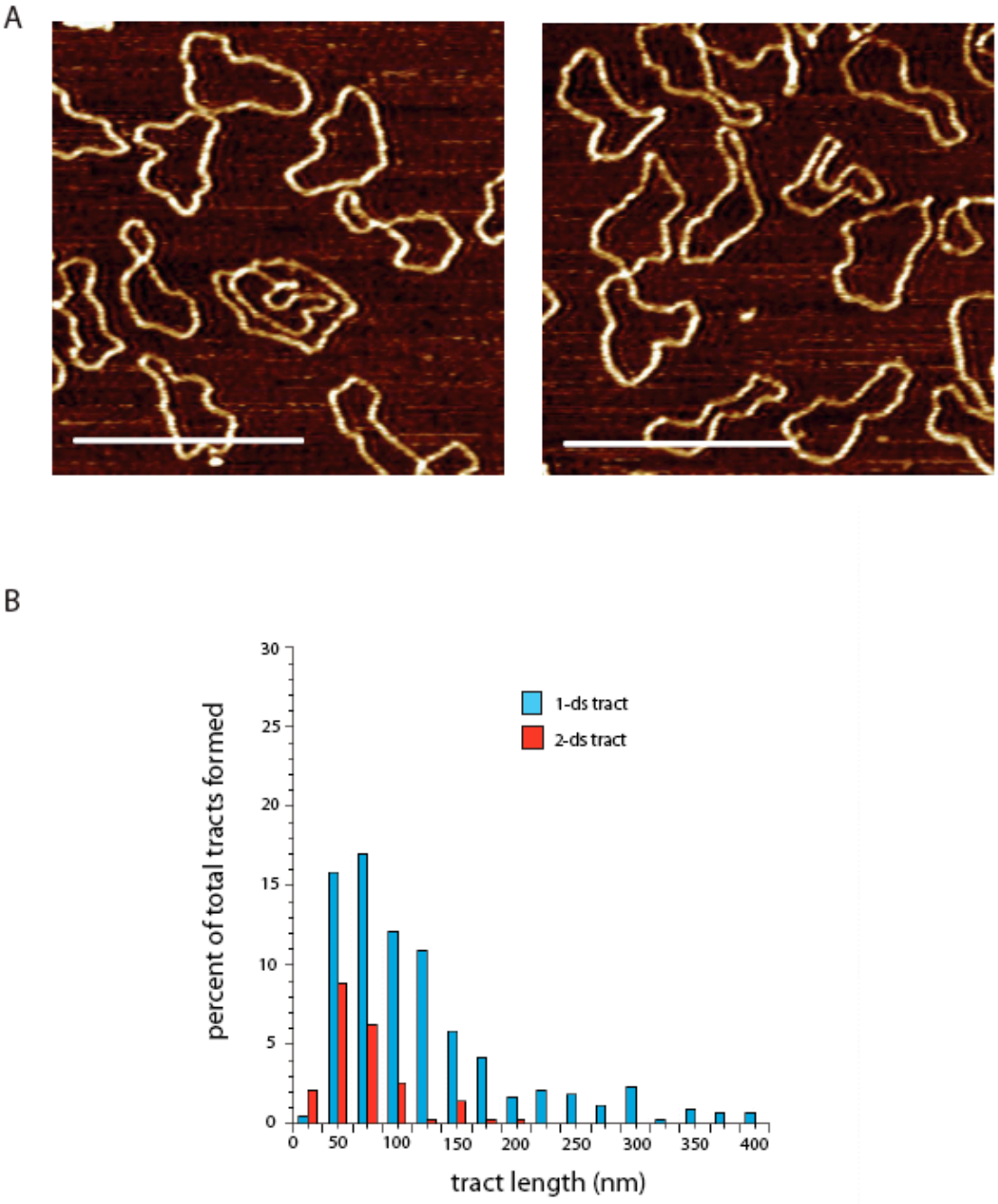


Figure 4.4: yMutLα and DNA in the presence of 100 mM Na^+.

(A) AFM images of yMutLα and DNA at 100 mM Na^+. Images are 1 x 1 μM; scale bar represents 500 nm. (B) distribution of 1-ds (n = 341) and 2-ds (n = 94) tract lengths. Lengths are plotted as percent of total tracts measured.

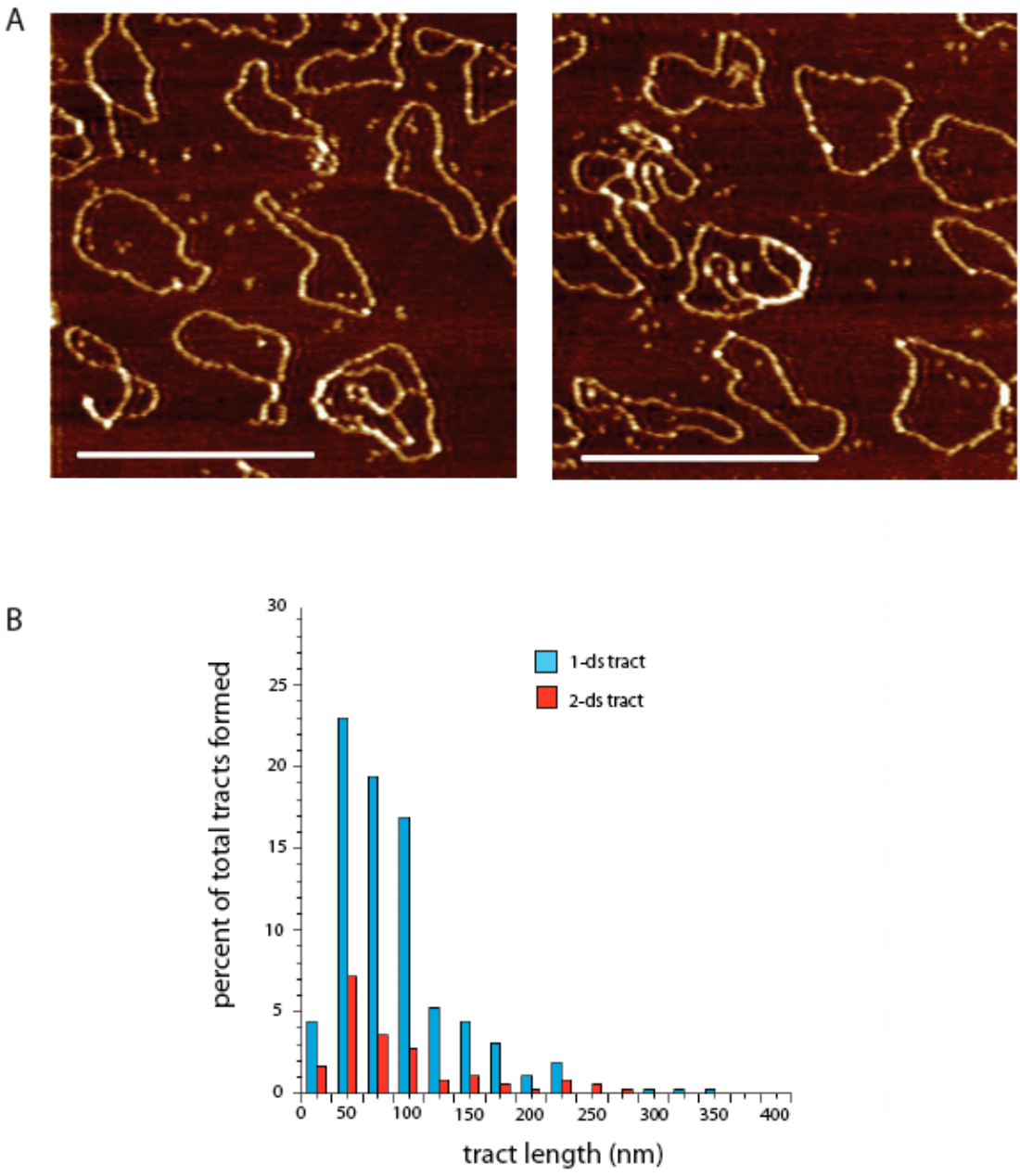


Figure 4.5: yMutLα and DNA in the presence of 125 mM Na^+.

(A) AFM images of yMutLα and DNA at 125 mM Na^+. Images are 1 x 1 μm; scale bar represents 500 nm. (B) distribution of 1-ds (n = 291) and 2-ds (n = 71) tract lengths. Lengths are plotted as percent of total tracts measured.

Effect of ATP on yMutLα binding to DNA at 25 mM Na^+

To date, there have been no published studies of yMutLα binding to DNA in the presence of ATP. Since ATP is ubiquitous in the cell, and has been shown to induce conformational changes in yMutLα (Chapter 2), it seems possible that DNA binding by yMutLα could be affected by the presence of ATP. To test this hypothesis, yMutLα and DNA were imaged by AFM in the absence and presence of 0.1 and 5 mM ATP, and the four possible binding modes (Figure 4.1) were tallied.

With no added adenine nucleotide present, all DNA seen in AFM images has protein bound to it (Figure 4.6). There are no observed single proteins bound to either one strand (1-ds binding) or two strands (2-ds binding) of DNA (Table 4.2). In the presence of 0.1 mM ATP, the binding of yMutLα to DNA decreases (Figure 4.7), and 86% of the DNA observed in the images has protein bound to it. Of the DNA that has protein bound to it, there is a decrease in the number of tracts, with tracts making up only 55% of all binding events. There is also a decrease in the frequency of 1-ds tracts going from 78% of the total protein tracts (in the absence of added ATP) to 1% of the total protein tracts (Table 4.2). There are also single binding events in the presence of 0.1 mM ATP, with yMutLα binding to either one strand (1-ds binding) or two strands (2-ds binding) of DNA. Of those single binding events, 68% are 1-ds binding events, and 32% are 2-ds binding events.

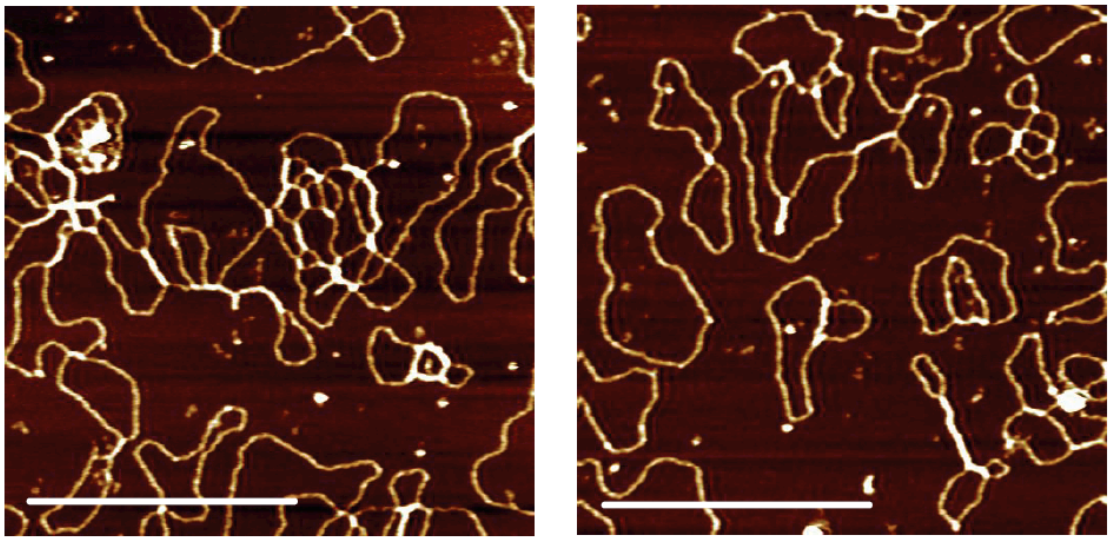


Figure 4.6: AFM images of yMutLα & DNA in the absence of added adenine nucleotide.

Scans are 1 x 1 μm; scale bar = 500 nm.

yMutLα + DNA +	*Percentage of DNAs with protein bound*	*% of tracts that are 1-ds tracts*	*% of tracts that are 2-ds tracts*	*% tracts / total binding events*	*% of binding events that are 1-ds binding*	*% of binding events that are 2-ds binding*
0 mM ATP	100%	78 ± 4%	22 ± 4%	100%	not observed	not observed
0.1 mM ATP	86 ± 8%	1 ± 6%	99 ± 6%	55 ± 2%	68 ± 6%	32 ± 6%
5 mM ATP	65 ± 1%	4 ± 1%	96 ± 1%	43 ± 2%	34 ± 8%	66 ± 8%

Table 4.2: Statistics for yMutLα binding to DNA in the absence and presence of ATP

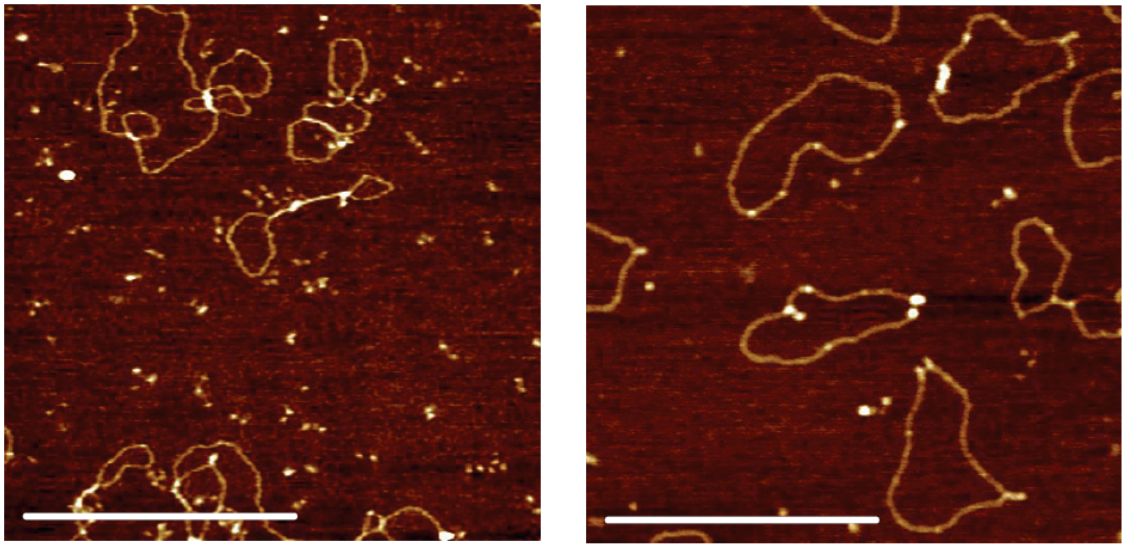


Figure 4.7: AFM images of yMutLα & DNA in the presence of 0.1 mM ATP.

Scans are 1 x 1 μm; scale bar = 500 nm.

In the presence of 5 mM ATP, binding of yMutLα to DNA decreases further (Figure 4.8), with only 65% of DNA having protein bound. Protein tracts (both 1-ds and 2-ds tracts) comprise only 43% of all yMutLα binding events. The distribution of the protein tracts is similar to the distribution seen with 0.1 mM ATP (Table 4.2), with 4% of the total protein tracts being 1-ds tracts and 96% of them being 2-ds tracts. The single binding events, however, "switch"; in the presence of 5 mM ATP, 34% are 1-ds binding events, and 66% are 2-ds binding events (Table 4.2).

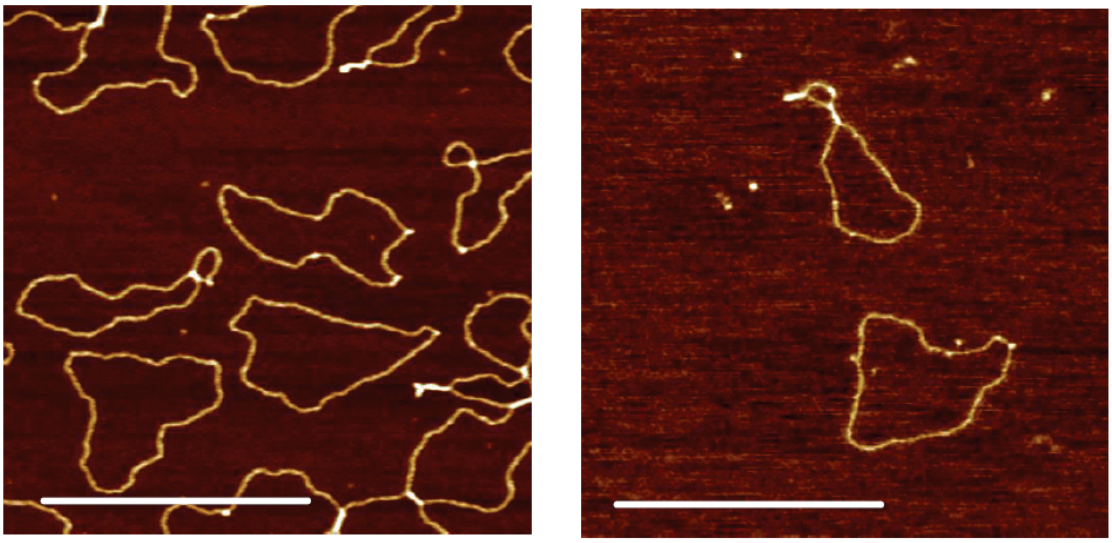


Figure 4.8: AFM images of yMutLα & DNA in the presence of 5 mM ATP.

Scans are 1 x 1 μm; scale bar = 500 nm.

Binding of yMutLα to nicked DNA in the presence and absence of ATP

The endonucleolytic activity in MutLα requires, among other things, the presence of a nick in the DNA (Kadyrov, Dzantiev et al. 2006). To see if yMutLα bound to nicked DNA differently it binds to covalently closed circular DNA, nicked DNA was imaged with yMutLα in the presence and absence of 5 mM ATP.

Preliminary data suggest that yMutLα interacts differently with nicked DNA than non-nicked, dsDNA in the presence and absence of ATP (Figure 4.9A and B). In the absence of ATP, the extensive tract formation seen with covalently closed circular DNA is not seen. There are single binding events and protein tracts formed, but these are in addition to large protein complexes on the DNA that cause looping out of the DNA (Figure 4.9C). Since this is preliminary data, there are no statistics on how frequently these large, looped protein complexes form or if they preferentially form under one set of conditions.

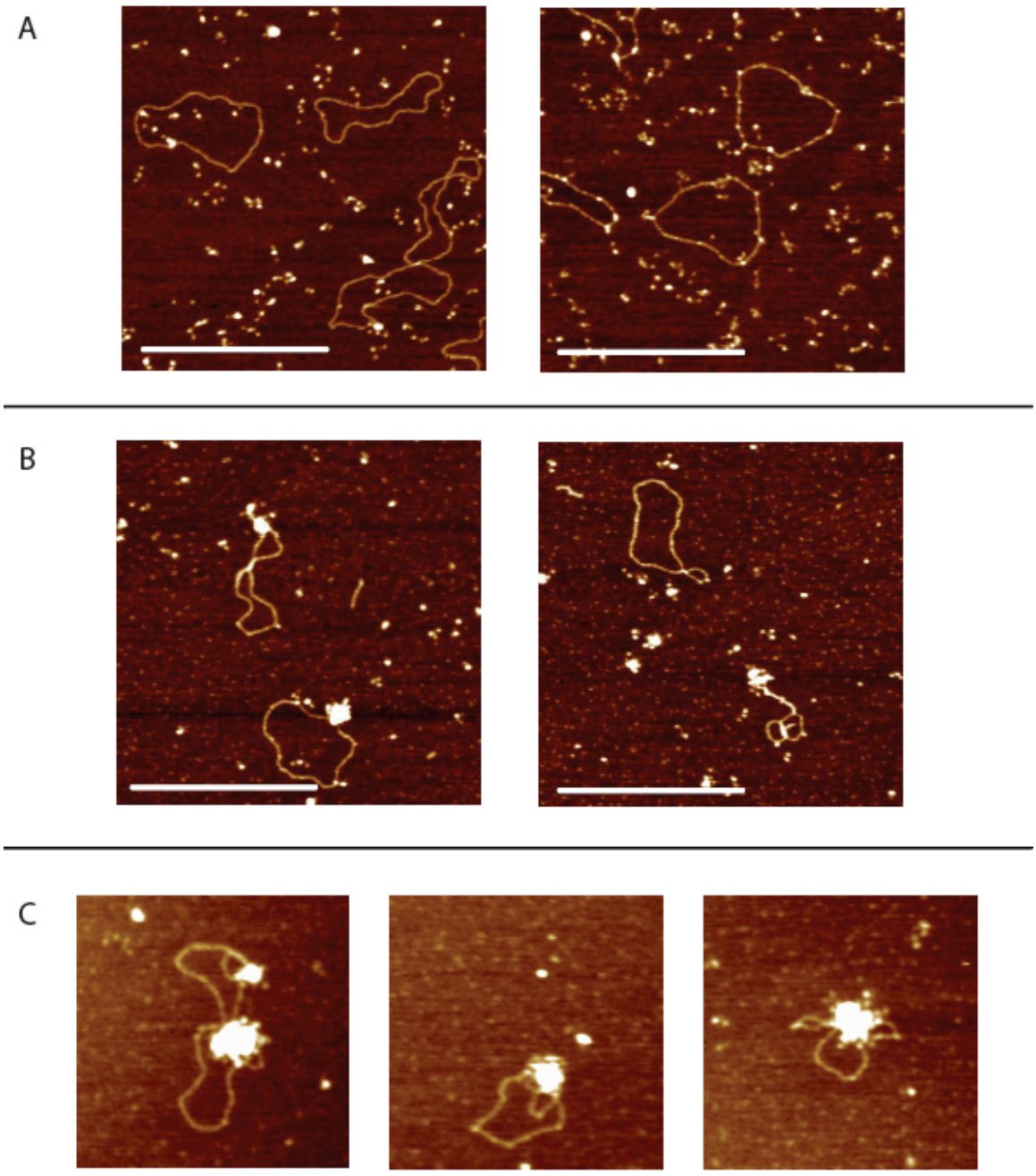


Figure 4.9

Figure 4.9: yMutLα and nicked DNA in the presence and absence of 5 mM ATP.

(A) yMutLα + nicked DNA in the absence of added adenine nucleotide (B) yMutLα + nicked DNA in the presence of 5mM ATP Scan size for both A and B is 1 x 1 μm; scale bar is 500 nm (C) assortment of looped structures seen in the presence and absence of ATP. Scan size is 250 x 250 nm.

Full-length yMutLα hydrolyzes ATP faster than N-terminal fragments

Even though yMutLα is an ATPase, there have been no published studies of ATP hydrolysis using full length protein. Analysis of the formation of ADP in an ATPase assay, spotted on TLC plates, such as that in Figure 4.10 shows that at an ATP concentration of 100 μM, ATP is hydrolyzed by yMutLα at a rate of ~2.4 min^{-1} (Figure 4.11, Table 4.3). This rate, based on preliminary data, is 2-6 times larger than what has been previously reported for full-length *E. coli* MutL or the N-terminal fragment of any MutL homolog. Unfortunately, because yMutLα often became catalytically inactive before the end of the time course, these rates were only reproducible twice over the full 60-minutes of the reaction (experiment was repeated 6 times) (Figure 4.11, solid lines are linear fits to full 60 minutes. Looking at the first ten minutes of the reaction (Figure 4.11, dashed lines are linear fits to first 10 minutes), however you can see that the protein is still functional and actively hydrolyzing ATP, and the rate of ATP hydrolysis is slightly faster (Table 4.3, in parentheses).

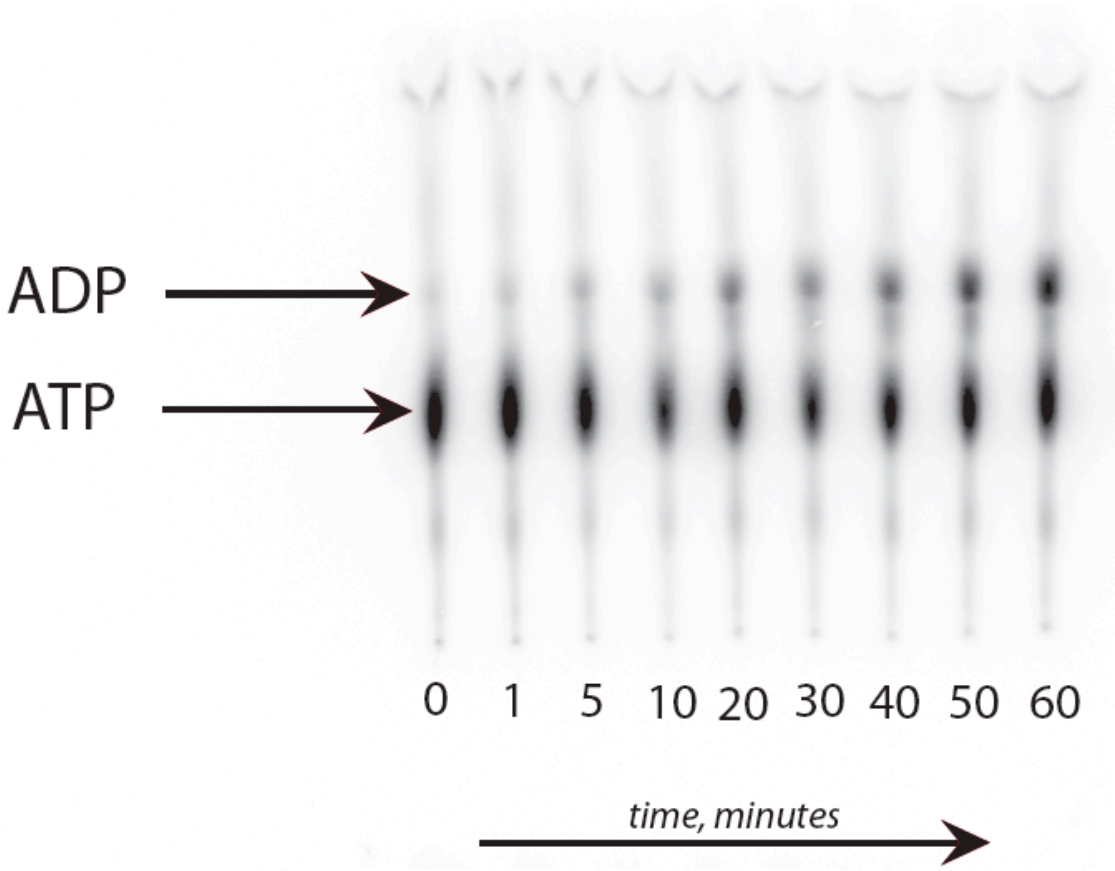


Figure 4.10: TLC plate monitoring ATP hydrolysis by yMutLα.

Radiolabelled α-^{32}P-ADP and α-^{32}P-ATP spots are marked.

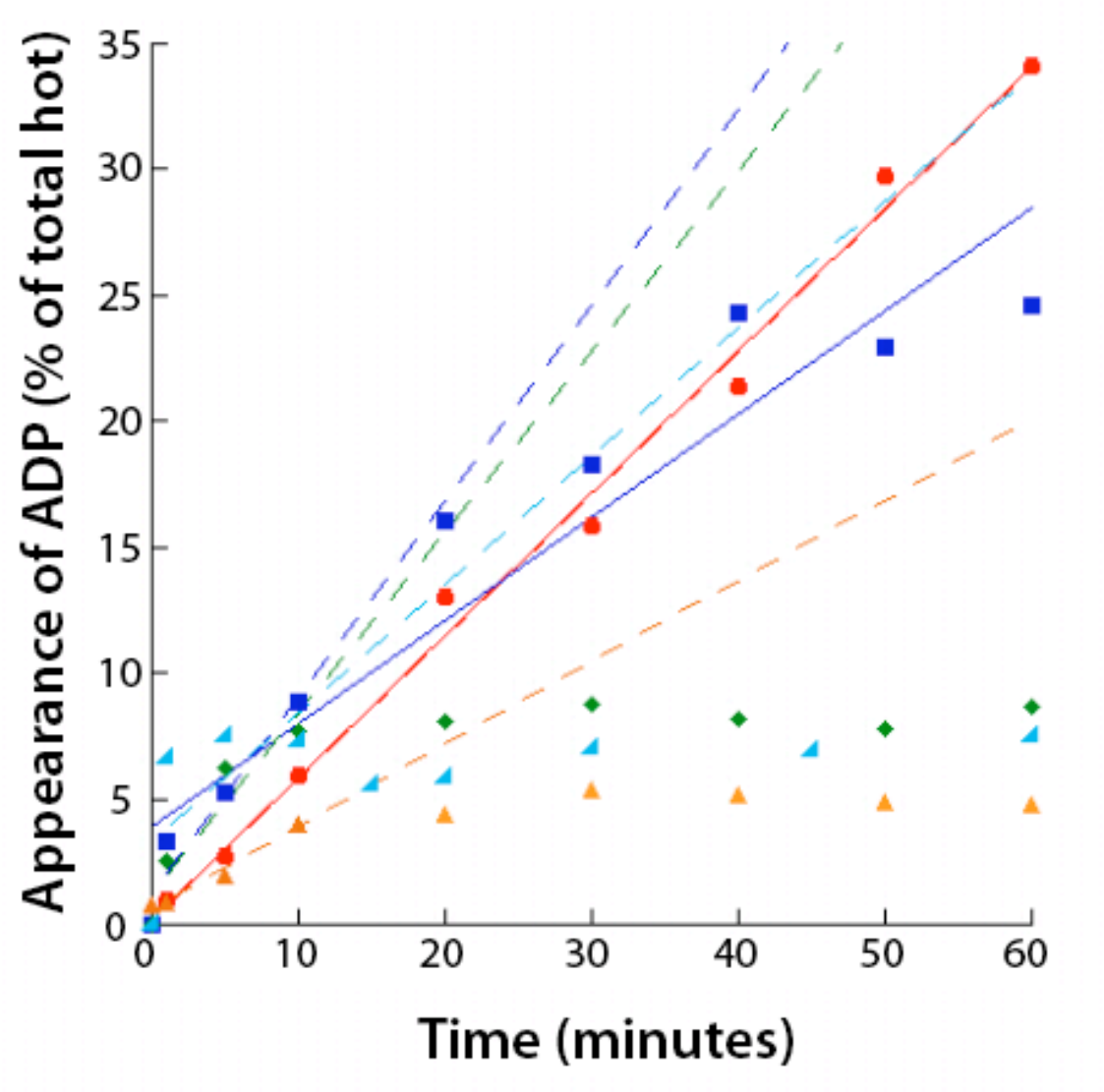


Figure 4.11: Appearance of ADP (as percent of total hot in the reaction) versus time for all ATPase reactions done with 100 μM ATP.

Linear fits are shown for the full 60 minutes of the reaction (solid lines) and for the first 10 minutes of the reaction (dashed lines).

Experiment	***Average Rate of ATP hydrolysis (min^{-1})*** *Full 60 min reaction (first ten minutes)*
yMutLα + 10μM ATP	0.17 (0.46)
yMutLα + 50 μM ATP	0.95 (0.91)
yMutLα + 100 μM ATP	2.4 (2.9 ± 0.89)

Table 4.3*: Calculated ATP hydrolysis rates for different concentrations of ATP.*

Data are from the reactions that appeared to go to completion. Numbers in parentheses represent the average rate calculated when using only the first 10 minutes of data. Standard deviations are only reported for those values that result from 3 or more rates.

At 50 μM ATP, the rate of ATP hydrolysis decreases to 0.93 min^{-1} (Figure 4.12, Table 4.3). Much like the experiments done at 100 μM ATP, yMutLα frequently became catalytically inactive ~10-15 minutes into the 60 minute time course. Fitting the first 10 minutes of all data to a line gives an ATP hydrolysis rate of 0.91 min^{-1} (Table 4.3, in parentheses). At 10 μM ATP, the average rate of ATP hydrolysis decreases further to 0.17 min^{-1} (Figure 4.12, Table 4.3). Interestingly, the first 10 minutes of the 10 μM ATP reaction has a hydrolysis rate of 0.46 min^{-1} (Table 4.3), suggesting that both of the reactions done at 10 μM ATP contained yMutLα that became catalytically inactive over the course of the experiment.

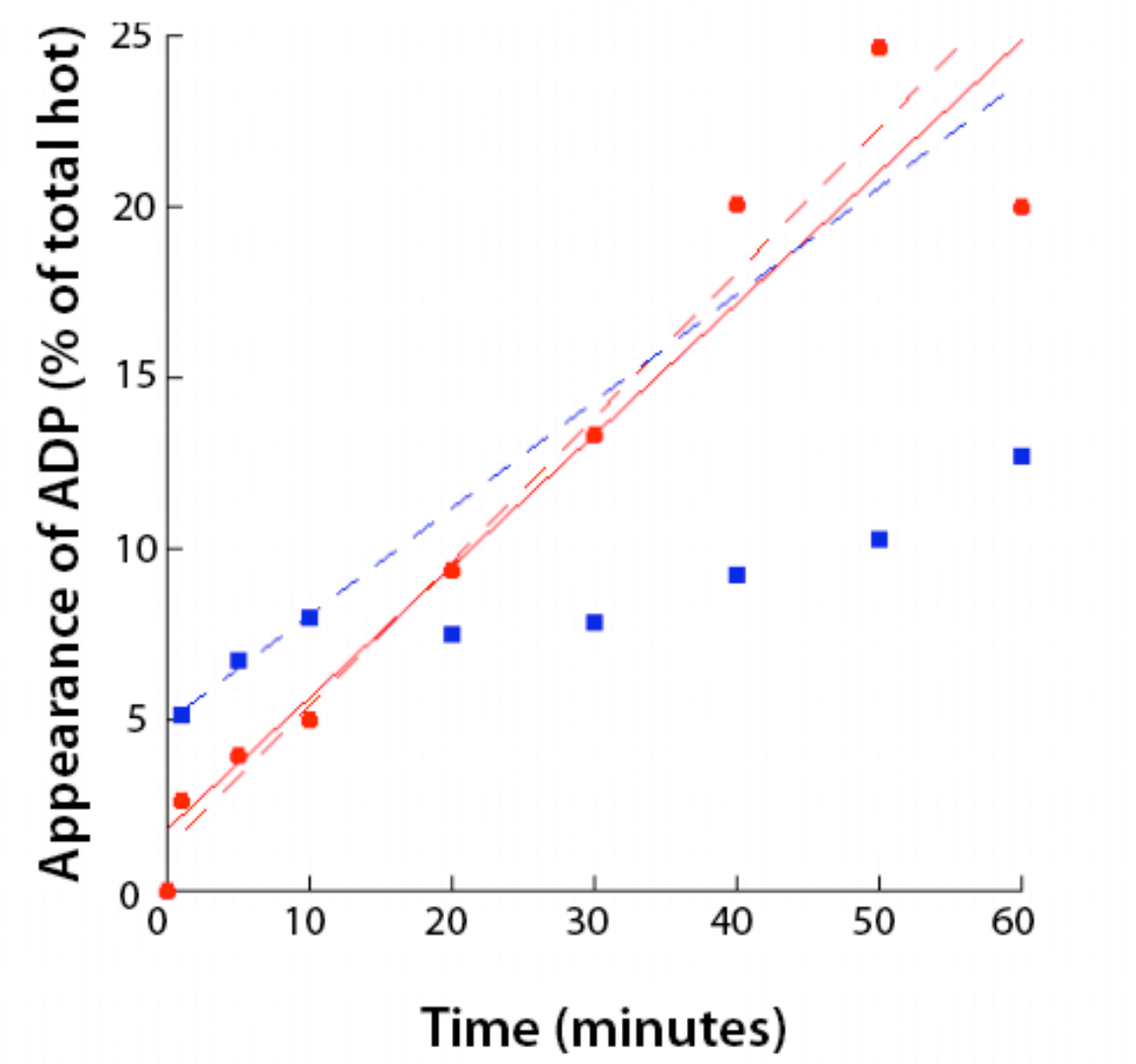


Figure 4.12: Appearance of ADP (as percent of total hot in the reaction) versus time for all ATPase reactions done with 50 μM ATP.

Linear fits are shown for the full 60 minutes of the reaction (solid line) and for the first 10 minutes of the reaction (dashed lines).

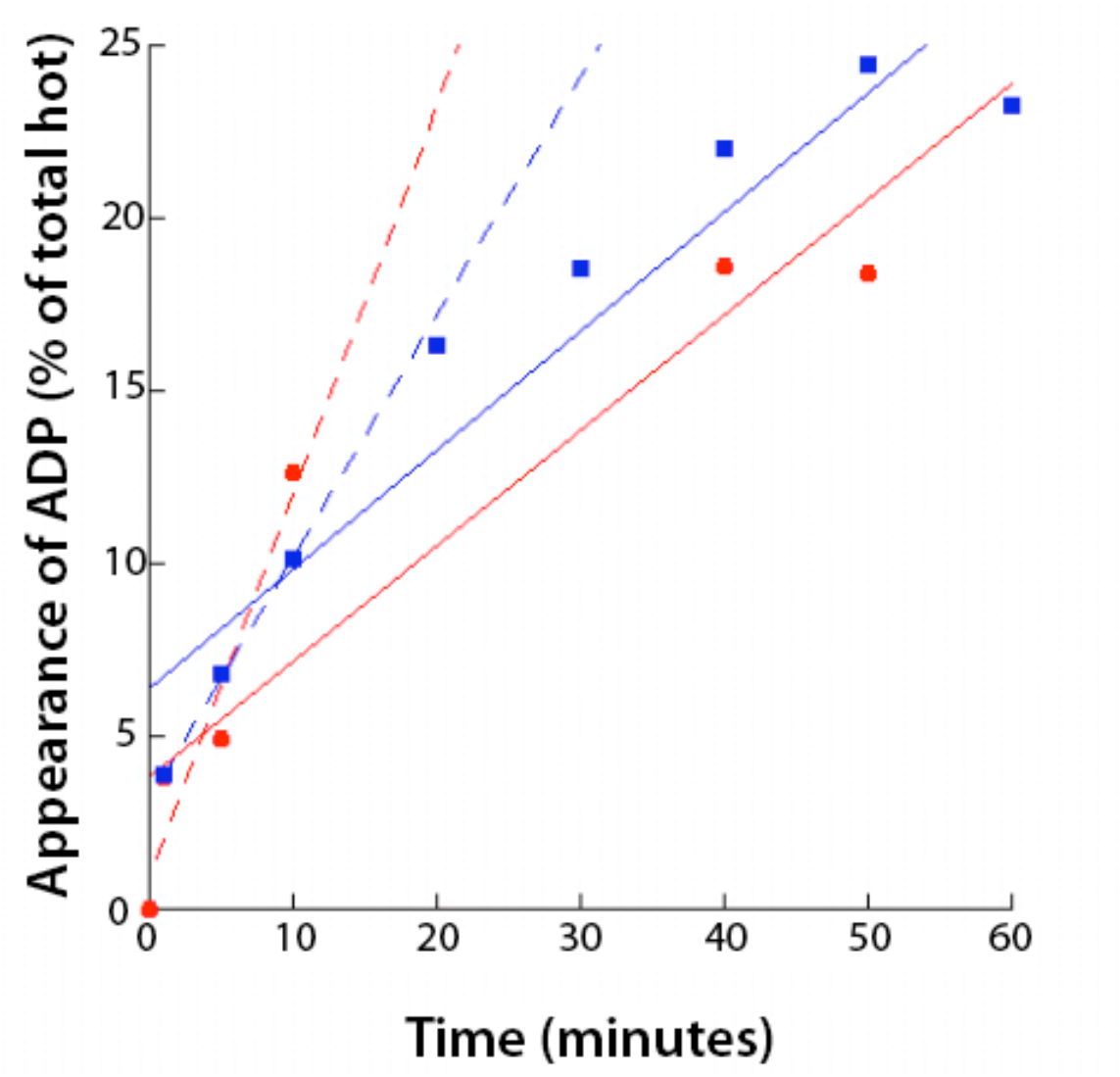


Figure 4.13: Plot of the appearance of ADP (as percent of total hot in the reaction) versus time for all ATPase reactions done with 10 μM ATP.

Linear fits are shown for the full 60 minutes of the reaction (solid lines) and for the first 10 minutes of the reaction (dashed lines).

Nicked DNA may increase the rate of ATP hydrolysis of yMutLα

Given that the ATPase activity of MutLα is required for the endonuclease activity of the protein, and that yMutLα will bind to nicked DNA, forming looped structures, it seemed plausible that the presence of nicked DNA might affect the rate of ATP hydrolysis. To investigate this, the ATPase assays of yMutLα at 100 μM ATP were repeated in the presence and absence of 6 nM nicked pUC18 (the same concentration used in AFM studies). While the –DNA reaction (Figure 4.14, blue markers) was reproduced twice, the + DNA reaction (Figure 4.14, red markers) only worked once without the protein becoming catalytically inactive. Looking at the first 10 minutes of the reactions (both in the presence and absence of DNA), however, shows that nicked DNA may not effect the ATPase activity of MutLα. Calculated rates of ATP hydrolysis (using the data from the 1st ten minutes minutes) are 3.3 min^{-1} in the absence of nicked DNA and 3.0 min^{-1} in the presence of nicked DNA.

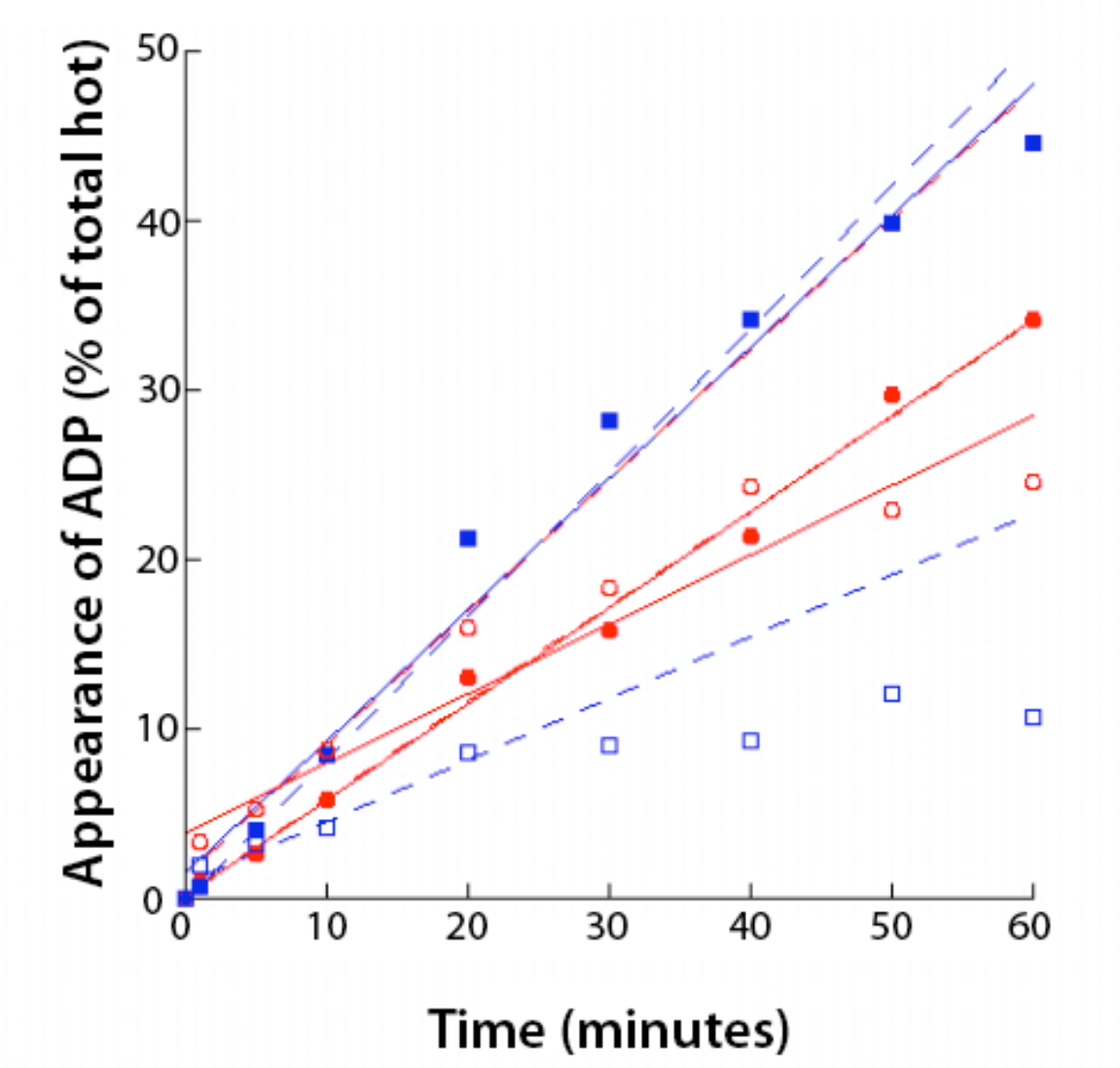


Figure 4.14: Plot of the appearance of ADP (as percent of total hot in the reaction) versus time for all ATPase reactions done with 100 μM ATP ± nicked DNA.

Blue squares represent experiments done in the absence of nicked DNA, and red circles indicate experiments done in the presence of nicked DNA. Linear fits for the full 60 minutes are shown for three of the four experiments (solid lines).Linear fits to the first 10 minutes are shown as dashed lines.

Discussion

yMutLα protein tract formation isn't sensitive to Na^+

Previous filter binding studies showed that increasing salt concentration lead to a decrease in the binding of yMutLα to DNA (Hall, Wang et al. 2001). The AFM results presented here show that increasing Na^+ concentration does not lead to an apparent decrease in yMutLα binding to DNA (as seen by AFM). These AFM results suggest that perhaps the Na^+ ion either does not affect binding to DNA as much as expected, or perhaps another factor contributed to the filter binding results. The buffer used in the filter binding studies contained Tris, pH 8.0, NaOAc, DTT, and MgCl. The buffer used in the AFM studies contained HEPES, pH 7.3, NaOAc and MgOAc. Perhaps the buffer used, the pH of the buffer used or the counter ion to the Mg may have a synergistic effect on the process of DNA binding; working together to enhance or prevent yMutLα from binding to DNA at higher salt.

yMutLα protein binding is affected by the ATP concentration

The experiments conducted in the presence of ATP, taken in conjunction with the yMutLα conformational changes observed in the absence of DNA, show that conditions that modulate the yMutLα conformation (Chapter 2) also affect its preferred mode of binding to DNA (Figure 4.15). In the absence of added adenine nucleotide, yMutLα exists in an extended, v-shaped conformation. This fact, in conjunction with the AFM images of yMutLα in the presence of DNA (with no added adenine nucleotide), suggests that the extended conformation of yMutLα will form both 1-ds and 2-ds tracts along DNA.

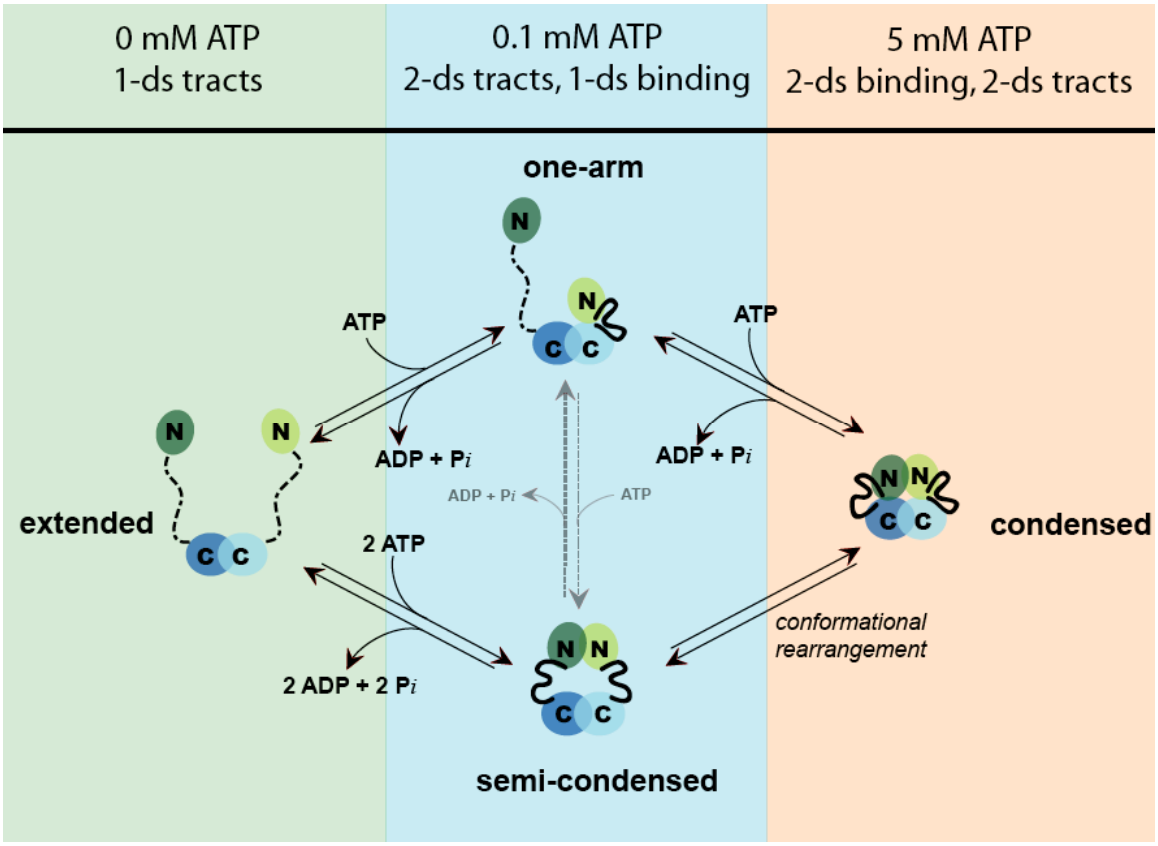


Figure 4.15: Relationship between DNA binding and proposed ATPase cycle for MutLα

In light of the fact that the DNA binding domains of yMutLα are in the N-termini, this makes sense; since the N-termini are likely extended away from the dimerization domain and each other each DNA binding domain is free to interact with sections of DNA that are either close (eventually forming 1-ds tracts) or far away (eventually forming 2-ds tracts).

At 0.1 mM ATP, overall yMutLα binding to DNA has decreased slightly (14%), and 1-ds tracts have all but disappeared (Figure 4.15). In the absence of DNA, at this ATP concentration the extended state of yMutLα is still dominant, but an increase in the one-armed state of yMutLα is seen (Chapter 2). It is incredibly interesting that 2-ds tracts are still formed; their presence suggests that ATP, while decreasing the total number of protein tracts formed (Table 4.2), increases the formation of 2-ds tracts (at least when compared to 1-ds tracts). Also interesting is the observation that of the yMutLα single binding events observed, ~2/3 of them are to 1 strand of DNA (1-ds binding); it suggests that the one-armed conformation of yMutLα prefers to bind to bind to only one strand of DNA, perhaps because one of the N-terminal domains is already making contacts with the linker arm or the dimerized C-terminal domains it is unavailable for binding to DNA.

At 5 mM ATP, yMutLα binding to DNA has decreased even more than at 0.1 mM ATP, with only 65% of DNA having protein bound (Figure 4.15). The protein tracts now make up only 43% of yMutLα binding to DNA, however the ratio of 1-ds tract to 2-ds tract is still roughly the same (Table 4.2). Of the binding events (one yMutLα bound to one or two strands of dsDNA) observed, ~2/3 of them are now 2-ds binding events. These data suggest that ATP decreases yMutLα 1-ds tract formation dramatically and

increases single binding events. In the absence of DNA, over 70% of yMutLα molecules are in the condensed state. Since both N-termini are believed to be associated with the dimerized C-termini in this state, it seems reasonable to think that the DNA binding domains are exposed to solvent in this conformation.

It could be inferred from these data and from the ATPase cycle proposed in Chapter 2 that the one-armed state of yMutLα binds preferentially to one strand of DNA, while the condensed state of yMutLα binds preferentially to 2 strands of DNA. These shifts in yMutLα binding to DNA at different ATP concentrations also suggest that DNA binding may be modulated by the conformational changes of yMutLα, which may play a role not only in MMR, but also in its other cellular processes, such as meiotic crossing over.

Hydrolysis of ATP by full-length yMutLα

The preliminary data presented in this chapter on the hydrolysis of ATP by full-length MutLα illustrates that the rate of ATP hydrolysis is faster than previously reported. The most likely explanation for this is that the previous data were on N-terminal fragments, while this data uses the full protein. Since it has been demonstrated that MutLα undergoes dramatic conformational changes upon ATP binding (Chapter 2), it seems reasonable that these conformational changes, and new inter-protein contacts formed as a result of the conformational changes, are important for facilitating ATP hydrolysis by the full-length protein.

Another possible explanation for the difference between the hydrolysis rates previously measured and those presented here could be that the previous experiments were conducted catalytically inactive protein, since the calculated rates of ATP

hydrolysis, for those experiments done with full length yMutLα that became inactive over the course of the experiment were close to (if not lower than) the previously reported rates of hydrolysis.

Since the rate of ATP hydrolysis appears to be higher than what has been previously reported, it suggests that MutLα might cycle through its conformational changes faster (Chapter 2). Proceeding through the ATPase cycle faster may help facilitate the timely repair of mismatched bases by the MMR process, although, if certain conformational changes are stabilized by binding to other MMR proteins (such as ExoI, PCNA or MutSα), perhaps interaction with those proteins with yMutLα would slow the rate of ATP hydrolysis. Likely a complex series of events keeps the "pace" for MMR.

Future Directions

It would be interesting to look at the effect of buffer identity and pH on yMutLα tract formation to see if the buffer identity (Tris vs. HEPES) or its pH has an effect on protein tract formation. It would also be beneficial to repeat the filter binding assays of yMutLα to DNA in the buffer used for AFM imaging to see if the results are similar and as an attempt to explain the differences between AFM data and filter binding results.

The multi-protein, looped DNA complexes seen when yMutLα is imaged with nicked DNA (both in the presence and absence of ATP) suggest that the interactions of yMutLα with nicked DNA are extensive. These preliminary experiments need further work to determine the frequency and extent to which these looped complexes of nicked DNA and yMuLα form. Following that work, it would be interesting to investigate what

would happen to these looped complexes if other proteins, like ExoI and PCNA were added.

Another interesting path to take would be to repeat the experiments testing the effect of ATP on binding of yMutLα to DNA, only changing the order of addition. In the original experiments, the ATP was always added to the yMutLα before the addition of the DNA. This order of addition potentially gave the preferred conformational state a chance to form before the DNA was added. Repeating the experiment with the reverse order of addition (yMutLα plus DNA first, then ATP) would determine if the pre-formation of yMutLα conformation is required for the changes seen in yMutLα binding to DNA.

The most interesting experiments to do based on the data on yMutLα binding to DNA are also the most difficult. Ideally, one could imagine yMutLα with DNA in the presence of the other mismatch repair proteins it interacts with. However, all of those known proteins (MutSα, PCNA and ExoI) also interact with some form of DNA, which may make interpretation of the data complicated.

Based on the preliminary data on the hydrolysis of ATP by yMutLα in this chapter, there are many exciting avenues to take with these experiments. The first would be to repeat these preliminary experiments, and add in some higher concentrations of ATP in an attempt to get kinetic parameters for the hydrolysis reaction. In addition, it would be interesting to assay the ATPase activity of MutLα in the presence of other MMR proteins. However, given the fact that two of the three proteins that MutLα is known to interact with in the MMR pathway also hydrolyze ATP, those proteins would

only complicate the assay. Perhaps a better place to start would be to see if the ATPase activity of MutLα is changed in the presence of PCNA.

The drugs radicicol and geldanamycin have been shown to inhibit the ATPase activity of type II topoisomerases and Hsp90s (Soti, Racz et al. 2002; Immormino, Dollins et al. 2004; Chu, Maynard et al. 2006; Corbett and Berger 2006; Gadelle, Graille et al. 2006), both members of the GHL ATPase family. Since MutLα is also a member of that ATPase family, it seems reasonable to expect that the drugs would also inhibit the ATPase activity of MutLα. This could be tested by incubating MutLα with either radicicol or geldanamycin for a period of time before addition of ATP to start the reaction.

It would also be interesting to look at the ATPase activity of Mlh1 alone, since it is possible to purify it. Thus far, no purification of full-length yPms1 or hPms2 alone has been reported. The ATP hydrolysis experiments using Mlh1 alone could help to elucidate the roles of both proteins in the hydrolysis of ATP.

Materials and Methods

Chemicals

ATP was purchased from Sigma. Radiolabelled ATP (α-^{32}P ATP, 800 Ci/mmol) was purchased from MP Biomedicals.

Protein Expression and Purification

yMutLα was purified as previously described (Hall and Kunkel 2001), with the exception of HEPES, pH 7.3 being used in the dialysis buffer instead of Tris, pH 7.3.

DNA Substrates

Relaxed pUC18

Relaxed pUC18 was generated by reacting pUC18 with Topoisomerase II (USB) for 1 hours at 30°C. The reaction was terminated by addition of 7 mM EDTA and 0.77% SDS. To check for DNA relaxation, a portion of each reaction was run on a 1% agarose gel in the absence of ethidium bromide, stained in a solution of ethidium bromide (0.5 μg/mL) for 30 minutes, and destained for 30 minutes in ddH_2O. Reactions that went to completion were then purified using a GFX PCR purification kit (Amersham) and stored at -20°C until being used in imaging.

Relaxed pBluescript

Relaxed pBluescript was a gift from Jode Plank at Duke.

Nicked pUC18

Nicked DNA was generated by digesting pUC18 with the nicking enzyme Nt.BstNB I (NEB) for 3 h at 55 °C, followed by enzyme inactivation at 80 °C for 20 min. DNA was separated from the inactivated nickase by filtering it through Micropure-EZ enzyme filters (Microcon). To check the state of the nicked DNA, it was imaged by AFM in buffer (25 mM HEPES, pH 7.3; 100 mM NaOAc; 10 mM MgOAc). If the DNA appeared relaxed on the surface, then it was assumed that there was at least one nick present.

Atomic Force Microscopy

AFM images were captured in air using either a Nanoscope III or IIIa (Digital Instruments, Santa Barbara, CA) microscope in tapping mode. Pointprobe Plus tapping

mode silicon probes (Agilent Technology, Tempe, AZ) with resonance frequencies of approximately 170 kHz were used for imaging. Images were collected at a speed of 2-3 Hz with image sizes of 1.5 - 2 μm at 512 x 512 pixel resolution.

Images were analyzed using Nanoscope III v5.12r3 software (Veeco, Santa Barbara, CA), Image SxM, v 1.69 (Barrett 2006) and NIH ImageJ (Rasband 1997-2006) with OpenAFM and Cell Counter plug-ins. Statistical plots were generated with Kaleidagraph (Synergy Software, Reading, PA).

Salt-Dependence of tract formation

To determine the effect of Na+ on yMutLα protein tract formation, yMutLα was diluted to a final concentration of 30 nM in one of three imaging buffers in the presence of relaxed pUC18. Each buffer contained 20 mM HEPES, pH 7.3, 5 mM MgOAc and then NaOAc at either 25, 100 or 125 mM. A volume of 10 μl of the sample was deposited onto freshly cleaved ruby mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water; excess water was blotted from the surface, and the surface was then dried using a stream of nitrogen.

Effect of ATP on yMutLα binding to DNA

To determine the effect of ATP on the binding of yMutLα to DNA, yMutLα was diluted to a final concentration of 30 nM in imaging buffer (20 mM HEPES, pH 7.3, 10 mM MgOAc, 25mM NaOAc) in the presence of 2 nM relaxed pBluescript DNA. A volume of 10 μl of the sample was deposited onto freshly cleaved ruby mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure

water; excess water was blotted from the surface, and the surface was then dried using a stream of nitrogen.

For experiments where ATP was present, ATP and yMutLα were mixed and then diluted in imaging buffer to a final adenine nucleotide concentration of 0.1 mM or 5 mM and protein concentration of 30 nM before the addition of relaxed DNA to a final concentration of 2 nM. The sample was then deposited the same as for MutLα protein and DNA alone.

yMutLα binding to nicked DNA in the absence and presence of ATP

To determine the binding of yMutLα to nicked DNA and to see if ATP would effect this binding in any way, yMutLα was diluted to a final concentration of 30 nM in imaging buffer (20 mM HEPES, pH 7.3, 10 mM MgOAc, 25mM NaOAc) in the presence of 6 nM nicked pUC18 DNA. A volume of 10μL of the sample was deposited onto the freshly cleaved ruby mica (Spruce Pine Mica Company, Spruce Pine, NC). The sample was rinsed immediately with nanopure water; excess water was blotted from the surface and the surface was then dried using a stream of nitrogen.

For experiments where ATP was present, ATP and yMutLα were mixed and then diluted in imaging buffer to a final adenine nucleotide concentration of 0.1 mM or 5 mM and protein concentration of 30 nM before the addition of relaxed DNA to a final concentration of 2 nM. The sample was then deposited the same as for MutLα protein and DNA alone.

TLC Plate Preparation

Before using the PEI cellulose TLC plates (Selecto Scientific), spot locations were marked and the plates were pre-run in ddH_2O to remove any residual material from the surface. The plates were then dried flat overnight.

ATPase assays

Determination of the Rate of ATP Hydrolysis

yMutLα (200nM) was mixed with reaction buffer (25 mM HEPES, pH 7.3; 100 mM NaOAc; 10 mM MgOAc) for a final reaction volume of 20 μL. Labeled ATP (α-^{32}P (10, 50 and 100μM final ATP concentration)) was added to initiate the ATPase reaction. At 0, 1, 5, 10, 20, 30, 40, 50 and 60 minutes, 1 μL of reaction was removed and added to 4 μL of formamide to quench the reaction. Two μL of each quenched time point was then spotted on pre-run PEI cellulose TLC plates (Selecto Scientific) and ran in 0.3 M KPO_4 buffer, pH 7.0. TLC plates were exposed to phosphor imager screens for 1 hour and read on a Storm Phosphorimager (Molecular Dynamics). Lanes were quantified using Image Quant software (Molecular Dymanics) and kinetic analysis was done with Kaleidagraph (Synergy, PA).

Effect of DNA on Rate of ATP Hydrolysis

yMutLα (200nM) was mixed with reaction buffer (25 mM HEPES, pH 7.3; 100 mM NaOAc; 10 mM MgOAc) in the presence and absence of 6 nM nicked pUC18, for a final reaction volume of 20 μL. Labeled ATP (α-^{32}P (100 μM)) was added to initiate the ATPase reaction. At 0, 1, 5, 10, 20, 30, 40, 50 and 60 minutes, 1 μL of the reaction was

removed and added to 4 μL of formamide to quench the reaction. Two μL of each quenched time point was then spotted on pre-run PEI cellulose TLC plates (Selecto Scientific) and ran in 0.3 M KPO_4 buffer. TLC plates were exposed to phosphor imager screens for 1 hour and read on a Storm Phosphorimager (Molecular Dynamics). Lanes were quantified using Image Quant software (Molecular Dymanics) and kinetic analysis was done with Kaleidagraph (Synergy, PA).

Calculation of ATP hydrolysis rates

ATP hydrolysis rates were calculated using the slope of the linear fit of the data of the appearance of ADP versus time. The slope, which reports the appearance of ADP as percent/minute, is directly related to the disappearance of ATP, since as ATP disappears, it is being converted to ADP. Multiplying the slope by the concentration of ATP used gives a rate in concentration/min, which, when the concentration of yMutLα is taken into account, can be converted into number of ATPs hydrolyzed per minute.

References

Au, K. G., K. Welsh, et al. (1992). "Initiation of methyl-directed mismatch repair." J Biol Chem **267**(17): 12142-8.

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

Barrett, S. D. (2006). Image SXM.

Chu, F., J. C. Maynard, et al. (2006). "Identification of novel quaternary domain interactions in the Hsp90 chaperone, GRP94." Protein Sci **15**(6): 1260-9.

Constantin, N., L. Dzantiev, et al. (2005). "Human mismatch repair: Reconstitution of a nick-directed bidirectional reaction." J Biol Chem.

Corbett, K. D. and J. M. Berger (2006). "Structural basis for topoisomerase VI inhibition by the anti-Hsp90 drug radicicol." Nucleic Acids Res **34**(15): 4269-77.

Drotschmann, K., W. Yang, et al. (2001). "Asymmetric recognition of DNA local distortion. Structure-based functional studies of eukaryotic Msh2-Msh6." J Biol Chem **276**(49): 46225-9.

Drotschmann, K., W. Yang, et al. (2002). "Evidence for sequential action of two ATPase active sites in yeast Msh2-Msh6." DNA Repair (Amst) **1**(9): 743-53.

Dufner, P., G. Marra, et al. (2000). "Mismatch recognition and DNA-dependent stimulation of the ATPase activity of hMutSalpha is abolished by a single mutation in the hMSH6 subunit." J Biol Chem **275**(47): 36550-5.

Gadelle, D., M. Graille, et al. (2006). "The HSP90 and DNA topoisomerase VI inhibitor radicicol also inhibits human type II DNA topoisomerase." Biochem Pharmacol **72**(10): 1207-16.

Guarne, A., M. S. Junop, et al. (2001). "Structure and function of the N-terminal 40 kDa fragment of human PMS2: a monomeric GHL ATPase." Embo J **20**(19): 5521-31.

Hall, M. C. and T. A. Kunkel (2001). "Purification of eukaryotic MutL homologs from Saccharomyces cerevisiae using self-affinity technology." Protein Expr Purif **21**(2): 333-42.

Hall, M. C., P. V. Shcherbakova, et al. (2003). "DNA binding by yeast Mlh1 and Pms1: implications for DNA mismatch repair." Nucleic Acids Res **31**(8): 2025-34.

Hall, M. C., P. V. Shcherbakova, et al. (2002). "Differential ATP binding and intrinsic ATP hydrolysis by amino-terminal domains of the yeast Mlh1 and Pms1 proteins." J Biol Chem **277**(5): 3673-9.

Hall, M. C., H. Wang, et al. (2001). "High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer." J Mol Biol **312**(4): 637-47.

Iaccarino, I., G. Marra, et al. (1998). "hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSa." EMBO J **17**(9): 2677-2686.

Immormino, R. M., D. E. Dollins, et al. (2004). "Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone." J Biol Chem **279**(44): 46162-71.

Kadyrov, F. A., L. Dzantiev, et al. (2006). "Endonucleolytic function of MutLalpha in human mismatch repair." Cell **126**(2): 297-308.

Kijas, A. W., B. Studamire, et al. (2003). "Msh2 separation of function mutations confer defects in the initiation steps of mismatch repair." J Mol Biol **331**(1): 123-38.

Lamers, M. H., D. Georgijevic, et al. (2004). "ATP increases the affinity between MutS ATPase domains. Implications for ATP hydrolysis and conformational changes." J Biol Chem **279**(42): 43879-85.

Lamers, M. H., H. H. Winterwerp, et al. (2003). "The alternating ATPase domains of MutS control DNA mismatch repair." Embo J **22**(3): 746-56.

Modrich, P. and R. Lahue (1996). "Mismatch repair in replication fidelity, genetic recombination, and cancer biology." Annu Rev Biochem **65**: 101-33.

Rasband, W. S. (1997-2006). ImageJ. US National Institutes of Health, Bethesda, Maryland, USA

Raschle, M., P. Dufner, et al. (2002). "Mutations within the hMLH1 and hPMS2 subunits of the human MutLalpha mismatch repair factor affect its ATPase activity, but not its ability to interact with hMutSalpha." J Biol Chem **277**(24): 21810-20.

Soti, C., A. Racz, et al. (2002). "A Nucleotide-dependent molecular switch controls ATP binding at the C-terminal domain of Hsp90. N-terminal nucleotide binding unmasks a C-terminal binding pocket." J Biol Chem **277**(9): 7066-75.

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

Tomer, G., A. B. Buermeyer, et al. (2002). "Contribution of human mlh1 and pms2 ATPase activities to DNA mismatch repair." J Biol Chem **277**(24): 21801-9.

Tran, P. T. and R. M. Liskay (2000). "Functional studies on the candidate ATPase domains of Saccharomyces cerevisiae MutLalpha." Mol Cell Biol **20**(17): 6390-8.

Zhang, Y., F. Yuan, et al. (2005). "Reconstitution of 5'-directed human mismatch repair in a purified system." Cell **122**(5): 693-705.

APPENDIX A

MATERIALS & METHODS

Buffers for AFM imaging

All buffers for AFM imaging (see Table A.1 for recipes) were made from 1 M stock solutions of each component (e.g.: 1 M NaOAc, 1 M MgOAc, 1 M HEPES, etc.) and diluted to the appropriate volume with nanopure water. Typically, 10 mL of buffer was made at a time, filtered at least twice with 0.02 μm Anotop syringe filters (Whatman), aliquoted into 1.5mL eppindorf tubes and stored at 4°C. Before use in imaging, buffers were heated to 65°C for at least 10 minutes and cooled slowly to room temperature.

Mica Substrates

Ruby mica (Spruce Pine Mica, Spruce Pine, NC) was affixed to metal discs (Ted Pella) using Tempfix adhesive (Ted Pella). To affix mica, the metal discs were heated directly on a hot plate, the heat-sensitive glue was dabbed onto the center of the disc and the piece of mica was placed directly on top of the glue. Pressure was applied to spread the glue and the discs were removed from the hot plate with tweezers and allowed to cool. Immediately before sample deposition, the top layer of mica was cleaved off with the use of transparent tape.

AFM

AFM experiments were conducted using a Nanoscope III or IIIa AFM (Veeco, Santa Barbara, CA) operated in tapping mode in air. All imaging samples were prepared similarly. The buffer and DNA (if DNA is being used in the experiment) were first heated to 65°C and incubated for 10 minutes before being slowly cooled to room temperature. If any protein is being used in the experiment it was added at this point and the sample was then diluted with buffer to an appropriate concentration for the experiment. A volume of 10-20 μL was then deposited either directly onto freshly cleaved mica or onto a small square of parafilm, after which the mica was lowered to the sample drop. The sample on the mica was then rinsed immediately with nanopure water; excess water was wicked from the surface using a piece of filter paper. The sample was then dried under a gentle stream of nitrogen gas, and placed in the AFM for imaging.

Pointprobe Plus tapping mode silicon probes (Agilent Technology, Tempe, AZ) with resonance frequencies of approximately 170 kHz were used for imaging. Images were collected at a speed of 2-3 Hz with an image size of 1 - 2 μm at 512 x 512 pixel resolution.

Nanoscope III v5.12r3 software was used for image analysis. Volume analysis was done as previously described (Ratcliff and Erie 2001; Yang, Wang et al. 2003) using Image SXM v1.69 software (Barrett 2006). Image analysis was done using NIH ImageJ v 1.37 (Rasband 1997-2006) with the plugins OpenAFM (to open AFM images in ImageJ) and CellCounter (to tally shapes of molecules on the surface). Statistical plots were generated using Kaleidagraph (Synergy Software, Reading, PA).

Buffer Name	Purpose	Components
HEPES 25-OAc (H25-OAc)	yMutLα protein tract formation yMutLα adenine nucleotide induced conformational changes	25 mM HEPES, pH 7.3 5 mM MgOAc 25 mM NaOAc
HEPES 25-OAc-10Mg (H25-OAc-10Mg)	yMutLα adenine nucleotide induced conformational changes interactions of yMutLα with nicked DNA	25 mM HEPES, pH 7.3 10 mM MgOAc 25 mM NaOAc
HEPES 100-OAc (H100-OAc)	yMutLα protein tract formation yMutLα adenine nucleotide induced conformational changes interactions of yMutLα with nicked DNA AID oligomerization state	20 mM HEPES, pH 7.3 10 mM MgOAc 100 mM NaOAc
HEPES 125-OAc (H125-OAc)	yMutLα protein tract formation	20 mM HEPES, pH 7.3 10 mM MgOAc 125 mM NaOAc
HEPES 150-OAc (H150-OAc)	yMutLα protein tract formation	20 mM HEPES, pH 7.3 10 mM MgOAc 150 mM NaOAc
EcoRV buffer (1-D diff)	1-dimensional diffusion project using EcoRV	10 mM HEPES, pH 7.3 50 mM NaCl 10 mM $CaCl_2$
APOBEC3G (A3G)	APOBEC3G oligomerization state APOBEC2 oligomerization state	50 mM HEPES, pH 7.3 5 mM $MgCl_2$

Table A.1: AFM Buffers

yMutLα Protein Purification

This is a detailed version of a published procedure (Hall and Kunkel 2001), along with the adjustments necessary to do the first column in batch, so that the purification of yMlh1-Pms1 from yeast can be repeated with minimal assistance. Recipes for all buffers and media can be found in tables A.2, A.3 and A.4.

Transformation:

The *Saccharomyces cerevisiae* strain BJ2168 was transformed with plasmids pMH1 (Mlh1 and Trp coding sequences) and pMH8 (Pms1 and Leu coding sequences) (both plasmids are from the Kunkel Lab at NIEHS), following standard yeast transformation procedure using LiOAc (Ausubel 1987). Briefly, 5 mL of YPD media was inoculated with a single colony of BJ2168 and grown overnight at 30°C. To this growth, 20 mL of YPDA was added, and growth continued at 30°C for another 6 hours. An aliquot of 1.5 mL of the culture was removed, pelleted in a microcentrifuge, and the supernatant was discarded. Carrier DNA (200 μg salmon sperm DNA) and ≤ 2.5 μg of each transforming DNA were added to the tube containing the pelleted yeast cells. Fresh PEG solution (1.2 mL) was added to the cells and DNA and the whole reaction was incubated for 30 minutes at 30°C. The tube was then placed in a 42°C water bath for exactly 15 minutes. Following heat shock, the cells were spun down and the supernatant was discarded. The cell pellet was resuspended in 200 μL of TE, and 100 μL of the transformation as well as 100 mL each of 1:5 and 1:10 dilutions were plated on selective plates ($SC_{-trp\ -leu}$) and allowed to grow for 3 days at 30°C.

Growth & Induction:

Overnight cultures (usually four) were started from the plate of transformed yeast and grown in 10 mL of selective media ($SCGE_{-trp\ -leu}$) for 24h with shaking at 30°C. From these 10 mL cultures, 4 x 1L cultures were started, and grown in $SCGE_{-trp-leu}$ for 24h with shaking (225 rpm) at 30°C (resulting in a final culture density of OD_{595} ~1). The cell culture was pelleted at 6000 x *g* for 5 minutes in four 500 mL centrifuge bottles, and the supernatant decanted. The cell pellet was then resuspended in a minimal amount (~45 mL per centrifuge bottle, 180 mL total) of YP-Gal, and to induce the galactose-regulated production of Mlh1 and Pms1 the resuspended cells (~12 mL per flask) were then used to inoculate 16 flasks each containing 1L of YP-Gal. Cultures were grown with shaking (225 rpm) at 30°C until the OD_{595} reached between 1.2 -1.3 (16-18 hours).

Cells were harvested by centrifugation at 6000 x *g* for 5 minutes in 6 x 500 mL centrifuge bottles. Cell pellets were washed with a minimal amount (~35 mL per bottle) of sterile water and consolidated into one centrifuge bottle. Cells were then spun for 15 minutes at 6000 x *g*. The water was decanted and cell pellet frozen at -80°C until ready to proceed with the protein purification.

Typical cell pellets from 16L of culture were around 100g, and normally lead to 0.8 mg of purified yMutLα. Unfortunately, yMutLα responds poorly to centrifugal ultrafiltration, so to obtain more concentrated protein, two growths (a total of 32 L of culture eluted with the same volume off the final column) are recommended.

Purification:

Unless stated, all procedures are done at 4°C.

Cell pellet was removed from storage at -80°C and thawed overnight on ice in the cold room.

Six protease inhibitor tablets (Roche, cOmplete EDTA-free) and PMSF (0.5 mM final concentration) were added to ~200 mL of chitin column buffer and allowed to stir until everything was dissolved. The thawed cell pellet was resuspended in a minimal amount of this ~200 mL of chitin column buffer. Half of the cell pellet was then left on ice while the other half and an equal volume of 0.5 mm glass beads (Biospec Products) were placed in a bead beater equipped with jacket. The bead beater jacket was packed with ice and water and the cells were beat for a total of 8 minutes, in 1-minute increments with a 5-minute rest period in between each beating. The lysed cells were separated from the beads using a pipette, and the lysis procedure was repeated using the remaining half of the resuspended cells. Lysed cells and beads were then centrifuged at 6000 x *g* to separate the cell debris from the supernatant. The pellet was discarded and the supernatant was centrifuged in an ultracentrifuge at 12,500 x *g* for 30 minutes at 4^{o}C.

While the ultracentrifuge was running, 30 mL of Chitin Beads (NEB) were washed 5 times, with 20 mL each time, of Chitin Column buffer. The beads were spun in 50 mL polypropylene conical tubes at 2500 x *g* for 5 minutes to pellet. After the washes, the beads were resuspended in 20 mL of chitin column buffer and divided between six new 50 mL conical tubes.

Following ultracentrifugation the supernatant, containing yMutLα, was removed and a 5% Polymin P solution was added gradually (20 μL Polymin P for every 1 mL of supernatant) to precipitate DNA and RNA. The solution was allowed to stir for 30 minutes, and then centrifuged at 10,000 x *g* for 10 minutes. The supernatant was

distributed between the 6 tubes containing chitin beads, and the remaining volume was filled with chitin column buffer. The tubes were rotated at 4°C for 1 hour, in order for yMutLα to bind to the chitin beads.

After the hour-long binding period, the beads were washed 4 times with 20 mL of chitin column buffer each time and spun at 2500 x *g* for 5 minutes. The chitin beads were then washed twice with 20 mL each of ATP wash buffer and spun at 2500 rpm for 5 minutes. The beads were finally resuspended in minimal ATP wash buffer and poured into the column with flow adaptor (BioRad; 1.7 cm ID). The column was allowed to settle before 20 mL of DTT-containing cleavage buffer was loaded onto the chitin column at a rate of 1 mL/min and allowed to sit on the column overnight (approximately 12 hours).

Free Mlh1 and Pms1 were eluted from the chitin column with 30 mL of buffer A + 100 mM NaCl at a rate of 0.5 mL/min directly onto a 5 mL HiTrap Heparin-Sepharose column (GE biosciences) pre-equilibrated with buffer A plus 100 mM NaCl (Heparin column was attached directly to the outlet of chitin column). The chitin column was then disconnected and the heparin column was washed with 15 mL buffer A + 100 mM NaCl (rate 0.5 mL/min). Mlh1 and Pms1 were then eluted into 1 mL fractions by a 50 mL linear gradient of NaCl (100-1000 mM) in buffer A. Mlh1 and Pms1 eluted around ~500 mM NaCl. Elution fractions containing Mlh1 and Pms1 were pooled.

The conductivity of the pooled fractions was measured, and adjusted to a final NaCl concentration of 100 mM using equation A.1, where V_{pool} is the volume of pooled fractions in mL, C_{pool} is the conductivity of pooled fractions in μS, *x* is the volume of buffer A + no salt to add, $C_{A\text{-}NS}$ is the conductivity of buffer A + no salt and C_{fin} is the

desired final conductivity. The sample was then loaded at 0.5 mL/min using a Superloop onto a 1 mL DEAE-Sepharose column (GE biosciences) equilibrated with Buffer A + 100 mM NaCl.

Equation A.1 $((V_{pool})(C_{pool})) + x(C_{A\text{-}NS}) = ((x+V_{pool})(C_{fin}))$

The DEAE column was washed with 4 mL of buffer A + 100 mM NaCl at 0.5 mL/min, and then Mlh1-Pms1 was eluted with a 10 mL linear gradient of buffer A + NaCl (100 mM – 1000 mM) at 0.5 mL/min. Mlh1-Pms1 typically eluted at ~ 180 mM NaCl. Bradford analysis (Bradford 1976) was performed on fractions in and around the peak before those fractions were pooled and dialyzed overnight against 1L dialysis buffer.

Following dialysis overnight (~12h), final protein concentration was determined via Bradford analysis (Bradford 1976). The protein was then aliquoted (100 μL aliquots), flash frozen in liquid N_2 and stored in liquid N_2 storage until needed, at which point the aliquot was thawed on ice and either used in its entirety or aliquoted into smaller volumes, flash frozen in liquid N_2 and stored at -80°C.

YPD	10 g yeast extract 20 g BactoPeptone 900 mL ddH2O autoclave 20 min (liquid cycle) & let cool 45 min – add 100 mL sterile 20% glucose
YPDA	10 g yeast extract 20 g BactoPeptone 885 mL ddH2O autoclave 20 min (liquid cycle) & let cool 45 min – add 100 mL sterile 20% glucose and 15 ml of a 0.2% adenine hemisulfate solution
PEG solution	8 mL of 50% w/v PEG 3500 or 4000 1 mL of 10x TE buffer, pH 7.5 1 mL of 10x (1M) LiOAc, pH 7.5

Table A.2 yMutLα transformation solutions and media

Amino Acid Mix (-trp, -leu)	1 g adenine 1 g uracil 1 g histidine 1 g arginine 1 g methionine 2.52 g phenylalanine 3 g tyrosine 3 g lysine 4 g isoleucine 5 g aspartic acid 5 g glutamic acid 7.24 g valine 10 g threonine 20 g serine mix well and store in dark glass or foil covered bottle
$SC_{-trp-leu}$ plates	6.7 g yeast nitrogen base, w/o AAS 1.4 g amino acid mix (-trp, -leu) 20 g agar 900 mL ddH_2O autoclave 20 min (liquid cycle) & let cool 45 min – add 100 mL sterile 20% glucose before pouring plates
$SCGE_{-trp-leu}$ media	6.7 g yeast nitrogen base, w/o AAS 60 mL 50% glycerol 20 mL 200 proof Ethanol 1.4 g amino acid mix (-trp, -leu) 915 mL ddH_2O autoclave 20 min (liquid cycle) & let cool 45 min then add 5 mL sterile 20% glucose
YP-Gal	20 g bactopeptone 10g yeast extract 950 mL ddH2O autoclave 20 min (liquid cycle) & let cool 45 min, then add 50 mL sterile 40% galactose

Table A.3: yMutLα media

5% Polymin P (Polyethylenimine) solution (50 mL)	5.35g Polymin P to ~30 mL ddH2O pH to 7.7 with 12 N HCl ddH_2O to 50 mL filter sterilize
Chitin Column Buffer (1L)	25 mM Tris-HCl, pH 8.0 500 mM NaCl 10% glycerol 0.1 mM EDTA ddH_2O to 1L filter sterilize
ATP Wash Buffer (150 mL)	25 mM Tris-HCl, pH 8.0 200 mM KCl 10 % glycerol 3 mM $MgCl_2$ ddH_2O to 150 mL filter sterilize **before use: add ATP to 100 μM final concentration**
Cleavage Buffer (100 mL)	25 mM Tris-HCl, pH 8.0 100 mM NaCl 10 % glycerol 0.1 mM EDTA ddH_2O to 95 mL filter sterilize **before use: add DTT to 50 mM final concentration**
2x Buffer A (DEAE column) (250 mL)	50 mM Tris-HCl, pH 8.0 20 % glycerol 0.2 mM EDTA 40 mM DTT ddH_2O to 250 mL filter sterilize **keep in cold room once made**
Buffer A + 100 mM NaCl (250 mL)	125 mL 2x Buffer A 5 mL 5M NaCl ddH_2O to 250 mL filter sterilize

Buffer A + 1M NaCl (100 mL)	50 mL 2x buffer A 20 mL 5M NaCl ddH_2O to 100 mL filter sterilize
1x Buffer A (no salt) (100 mL)	50 mL 2x Buffer A ddH_2O to 100 mL filter sterilize
Dialysis Buffer (1L)	25 mM Tris-HCl, pH 8.0 10 % glycerol 0.1 mM EDTA 100 mM NaCl 1 mM DTT ddH_2O to 1 L

Table A.4: yMutLα Protein Purification Solutions

References

Ausubel, F. M. (1987). Current protocols in molecular biology. New York, Published by Greene Pub. Associates and Wiley-Interscience : J. Wiley.

Barrett, S. D. (2006). Image SXM.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem 72: 248-54.

Hall, M. C. and T. A. Kunkel (2001). "Purification of eukaryotic MutL homologs from Saccharomyces cerevisiae using self-affinity technology." Protein Expr Purif 21(2): 333-42.

Rasband, W. S. (1997-2006). ImageJ. US National Institutes of Health, Bethesda, Maryland, USA

Ratcliff, G. C. and D. A. Erie (2001). "A novel single-molecule study to determine protein--protein association constants." J Am Chem Soc 123(24): 5632-5.

Yang, Y., H. Wang, et al. (2003). "Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy." Methods 29(2): 175-87.