SINGLE-MOLECULE FLUORESCENCE STUDIES OF DNA BENDING DURING PROKARYOTIC MISMATCH REPAIR INITIATION

Jacob Wayne Gauer

A dissertation submitted to the faculty at 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
2015

Approved by:

Dorothy Erie

Nancy Thompson

Matt Redinbo

Tom Kunkel

Barry Lentz

ABSTRACT

Jacob Wayne Gauer: SINGLE-MOLECULE FLUORESCENCE STUDIES OF DNA BENDING DURING PROKARYOTIC MISMATCH REPAIR INITIATION
(Under the direction of Dorothy Erie)

DNA mismatch repair (MMR) is a process that is responsible for repairing base-base mismatches and insertion/deletion loop errors incorporated during DNA replication. In humans, deficiencies in MMR are linked to cancers, including Lynch Syndrome. MMR is initiated by MutS in prokaryotes (or the MutS homologs in eukaryotes), which is responsible for recognizing the error in the DNA. Upon error recognition, MutS undergoes ATP-dependent conformational changes to form a sliding clamp state that moves along the length of the DNA. This state is thought to be important for downstream repair events, such as recruitment of the second protein in the pathway, MutL.

Recent single-molecule fluorescence studies have led to a refined model for error recognition and sliding clamp formation by *Thermus aquaticus* MutS. While it is well established that MutS bends DNA and that this DNA bending is dynamic in the absence of nucleotides, little evidence for the DNA bending status through MMR initiation exists in the presence of ATP.

In this work, the nucleotide dependence of *Thermus aquaticus* MutS-induced DNA bending throughout sliding clamp formation is characterized. The current model for MutS conformational changes is then modified to reflect the newfound DNA bending information. To this end, single-molecule fluorescence resonance energy transfer (smFRET) between two arms of specially designed DNA oligonucleotides is monitored. These substrates are designed such that changes in DNA bending would result in changes in FRET efficiency. Finally, a data analysis pipeline developed specifically for high throughput analysis of data of this type is presented.

To Aaron,

My favorite.

ACKNOWLEDGEMENTS

I did not want to do this. I never really saw myself earning a Ph.D. doing scientific research. Somehow, though, I have dragged, clawed, and whined my way through the last several years to make it here. I could not have done this alone, and I owe so many people for encouraging me, forcing me even, to stay the path and make it to the end.

I must start by recognizing my advisor, Dorothy. You have allowed me to define success in my own way, a trait that has been vital to perseverance. You have shown me that in a sea of never-ending disappointment and discouragement, nothing is ever as bad as it seems. At many times, I have been close to quitting, but you have always managed to pull me back from the edge to face the problems at hand. This thesis only exists because of your guidance. You are a wise, fervid, and, at times, weird mentor, and I aspire to become more like you.

To my husband, Aaron, words cannot express my gratitude. No one else has felt the impact of this experience more than you. Living in different states for years, dealing with my grumpiness, tolerating while I talk *at* you about my research – you have sacrificed for this thesis. I owe you. I love you. But seriously, let's never do this again.

I want to thank my colleagues, past and present, both in and out of the Erie lab, for tolerating my never-ending stream of whining. Despite having your own lives and your own problems, you have listened. You have offered advice. You have cared. And you have never stopped being helpful.

Kira, I have leaned on you more than anyone. When I think back on graduate school, after my initial shiver of dread, I will remember it as the time that started our friendship. Vanessa, I have been jealous of your success since the beginning. Your shoes were hard to fill, and I only hope I have come close. Jackie, I will miss your Christmas Tree Burning Party, your "pseudo-sister" stories, and our "laser

room chats". Remember me when you become Queen of the World. Zimeng, I cannot wait to see what you accomplish. You are someone I'll overhear people talking about in the future, and I'll get to say that I knew you before you were famous. Marc, you certainly are memorable. Legendary even. The stuff of stories passed on for generations. Danielle, few people make me smile every time I interact with them. Few people can beam optimism *during graduate school*. I don't know how you do it. Sarah, ***waves arms excitedly***. (I'm going to miss that.) Hunter, it amazes me that you met me at my most curmudgeon-y and thought, "Hey, why don't I try that?" I like how you don't back down from a challenge. Matt Satusky, I will miss your punny crossword talents. It will be difficult to find someone else who is willing to giggle at my terrible jokes. Sharonda, I have been and continue to be nothing but impressed by you. I am humbled by the way, even when faced with challenges, you make it work and carry on. I admire your diligence. Logan, it's still weird how our similar experiences somehow brought us to the Erie lab at the same time. I suspect it'll happen for no apparent reason again someday. See you then. Thao, you are super fun, and I am so glad that we met each other when we did. You made the final race to the end a good time. I hope I did the same for you. Matt Meiners, Cassandra, Jet, Rebecca, Karen, Minu, Adrienne, and Lior, you all have been there for me, commiserated with me, leaned on me, collaborated with me, and brought me cupcakes. I'm in your debt.

To everyone who contributed their scientific expertise to my thesis, thank you. As a person who often struggles at the bench, I have leaned on your guidance and, on occasion, your hands to help me complete these projects. Special thanks go to Keith, Ruoyi, Lauren, and Vanessa whose work provided the entire basis of my project.

I must also thank Brian Hogan and Laura Benton for showing me that, Ph.D. or not, the career I want is out there; I just have to go and get it. I have. Thank you.

Finally, to my family: Yes, I'm finally done with school. I promise I'll get a real job now.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°	Degrees
′	Prime
A	Adenine
Å	Angstrom
α	Leakage Constant
ADP	Adenosine Diphosphate
AFM	Atomic Force Microscopy
ATP	Adenosine Triphosphate
ATPγS	Adenosine 5′-[γ-thio]triphosphate
b	Biotin
B	Bent State
bp	Base Pairs
BSA	Bovine Serum Albumin
C	Cytosine
d	Penetration Depth
D	Free DNA State
dig	Digoxigenin
DNA	Deoxyribonucleic Acid
Δ	Deletion
E	Fluorescence Resonance Energy Transfer Efficiency
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
emCCD	Electron Multiplying Charge Coupled Device
Exo I	Exonuclease I
Exo VI	Exonuclease VI

Exo X	Exonuclease X
FEN I	Flap Endonuclease I
FRET	Fluorescence Resonance Energy Transfer
FRET-TACKLE	Fluorescence Resonance Energy Transfer Transition Analysis Coupled with Kinetic Lifetime Examination
G	Guanine
γ	Gamma Factor
HCl	Hydrochloric Acid
Htt	Huntingtin Gene
HTT	Huntingtin Protein
I	Intensity
I_A	Acceptor Intensity
I_D	Donor Intensity
IDL	Insertion/Deletion Loops
k	Exponential Decay Constant
K_D	Dissociation Constant
mg	Milligram
ml	Milliliter
mM	Millimolar
min	Minute
MMR	Mismatch Repair
MLH1	MutL Homolog 1
MSH2	Human MutS Homolog 2
Msh2	Yeast MutS Homolog 2
MSH3	Human MutS Homolog 3
Msh3	Yeast MutS Homolog 3

MSH6	Human MutS Homolog 6
Msh6	Yeast MutS Homolog 6
µl	Microliter
n	Refractive Index
N	Number of Molecules
N.A.	Numerical Aperture
nm	Nanometer
nM	Nanomolar
η	Detection Efficiency
OD	Optical Density
p	Exponential Weighting Term
PCNA	Proliferating Cell Nuclear Antigen
PDB ID	Protein Database Identification
PEG	Polyethylene Glycol
PMS2	Human Postmeiotic Segregation Increased 2
Pol III	DNA Polymerase III
Pol δ	DNA Polymerase δ
ppm	Parts per million
Q	Glutamine
RecJ	Exonuclease
RPA	Replication Protein A
r	Interfluorophore Distance
R_0	Förster Radius
S. cerevisiae	*Saccharomyces cerevisiae*
sec	Second
smFRET	Single-molecule Fluorescence Resonance Energy Transfer

SSB	Single-stranded Binding Protein
t	Time
T	Thymine
TAMRA	Carboxytetramethylrhodamine
Taq	*Thermus aquaticus*
T-Bulge	Single Thymine Insertion
TDP	Transition Density Plot
TIRF	Total Internal Reflection Fluorescence
Tris	Tris(hydroxymethyl)aminomethane Buffer
θ_C	Critical Angle
U	Enzymatic Units
U_1	Unbent State 1
U_2	Unbent State 2
URA	Uracil
φ	Quantum Yield
WT	Wild Type
w/v	Weight by Volume
YAPD	Yeast Extract, Adenine, Peptone, Dextrose

CHAPTER 1: DNA MISMATCH REPAIR AND SINGLE-MOLECULE FLUORESCENCE: THE STUFF OF NOBEL PRIZES

Introduction

Maintaining the integrity of genetic information during DNA replication is vital for all organisms. Due to the high fidelity of the DNA polymerases, including their 3' to 5' proofreading activity, DNA replication is inherently highly accurate. Unfortunately, bases are sometimes misincorporated resulting in non-Watson/Crick base pairs; the polymerases can also slip, leading to insertion/deletion loop errors (IDLs). Errors like these occur approximately once every 10^7 bases replicated (Iyer, Pluciennik, Burdett, & Modrich, 2006; Kunkel & Erie, 2005). In the human genome (approximately 10^9 base pairs), this error rate results in hundreds of mistakes per round of replication (Iyer et al., 2006). If these errors are left unresolved, the misincorporated bases will be read as the template strand in subsequent rounds of DNA replication, leading to mutations, genomic instability, and even cancer (Kunkel & Erie, 2005).

Unsurprisingly, several enzymatic pathways have evolved to post-replicatively resolve the errors introduced into the genome. One such pathway, DNA mismatch repair (MMR), increases the fidelity of DNA replication 100-1000 fold by correcting the base-base mismatches and IDLs introduced during DNA replication (Iyer et al., 2006; Kunkel & Erie, 2005; Modrich, 1987; Schofield & Hsieh, 2003). MMR can occur bi-directionally (i.e. 5' to 3' or 3' to 5') and is highly conserved across prokaryotes and eukaryotes. Notably, MMR is important to human health, as defects in MMR are linked to greater than 80% of hereditary non-polyposis colorectal cancer cases (Kaur et al., 2011).

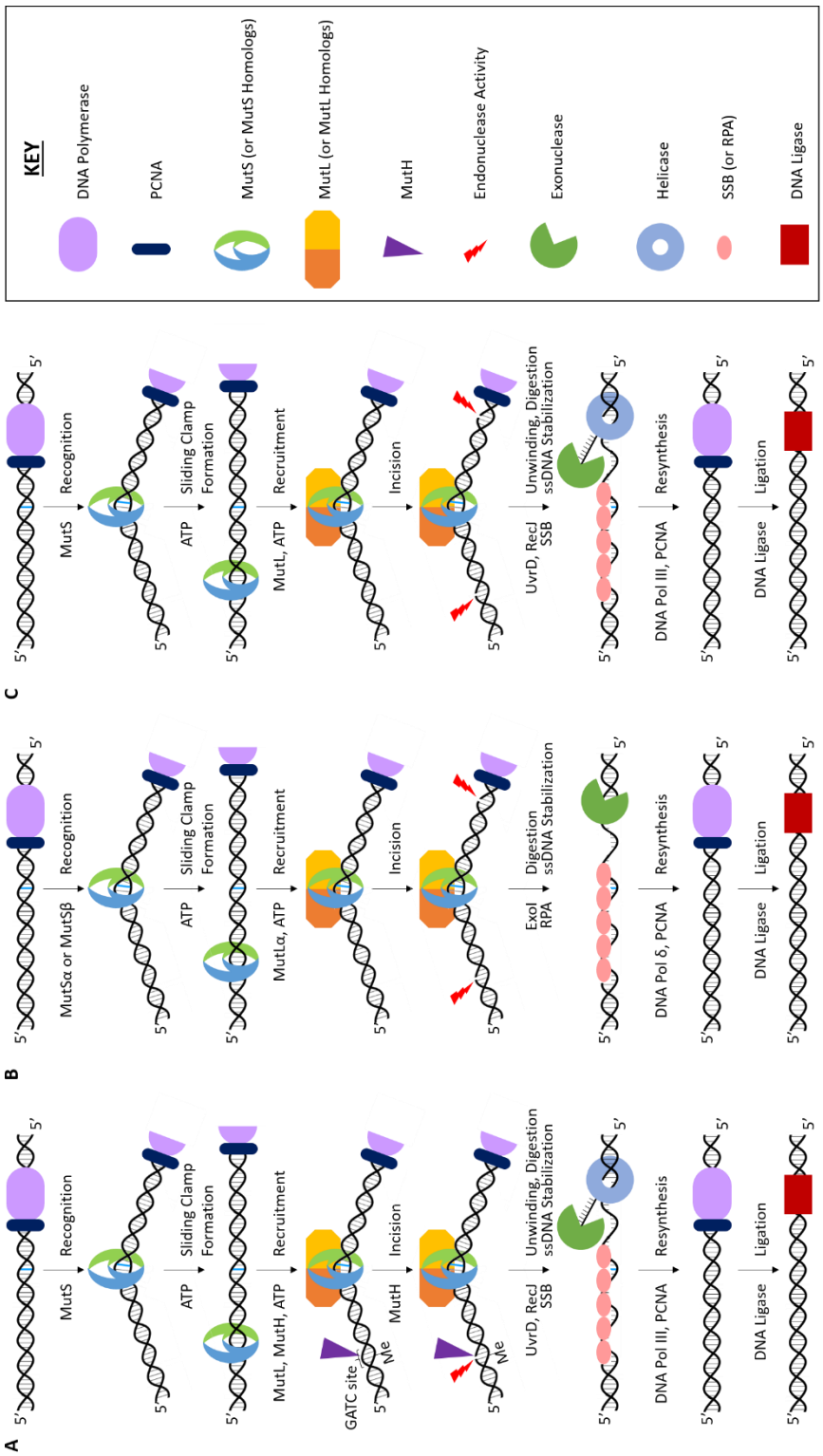

Figure 1.1: DNA Mismatch Repair.

The three panels diagram the steps of DNA mismatch repair in **A)** *Escherichia coli*, **B)** humans, and **C)** *Thermus aquaticus*.

DNA Mismatch Repair

Mismatch repair has been well characterized in *Escherichia coli* (*E. coli*). Several proteins are required to carry out MMR in *E. coli*, including MutS, MutL, MutH, UvrD, SSB, an exonuclease (Exo I, Exo VII, RecJ, or Exo X), DNA polymerase III (Pol III) and DNA ligase (Figure 1.1A). Following DNA replication, errors left behind by DNA polymerase are first recognized by a MutS homodimer. Upon introduction of ATP, MutS undergoes a conformational change into a sliding clamp conformation that encircles the DNA and is able to freely diffuse along the length of the DNA. After error recognition, MutS is also responsible for the ATP- and mismatch-dependent recruitment of a MutL homodimer. This MutS:MutL complex then induces the latent endonuclease activity in MutH, which nicks the daughter strand at a hemi-methylated GATC site on either side (i.e. either 5' or 3') of the error. At the hemi-methylated GATC site, the template strand is methylated while the newly synthesized strand is unmethylated. This difference in methylation status between the strands allows for discrimination between the correct (template) and incorrect (daughter) bases. The helicase UvrD then unwinds the DNA at the nick, and SSB binds to and stabilizes any single-stranded regions of DNA. An exonuclease then digests the daughter strand past the error, and the resulting single-stranded gap is resynthesized by Pol III. Finally, DNA ligase seals the nick (P Hsieh, 2001; Iyer et al., 2006; Jiricny, 2006; Kunkel & Erie, 2005; Schofield & Hsieh, 2003).

The initiation steps of MMR are highly conserved across several organisms, and homologs of several of the *E. coli* MMR proteins have been identified in other prokaryotes and in eukaryotes. In humans (Figure 1.1B), MMR is initiated by one of two isoforms of the heterodimeric MutS homologs. MutSα (a heterodimer of MSH2 and MSH6) is the isoform responsible for recognizing base-base mismatches and short IDLs, while MutSβ (a heterodimer of MSH2 and MSH3) recognizes longer IDLs. MutSα and MutSβ both undergo ATP-dependent conformational changes that allows them to recruit MutLα (a heterodimer of the MutL homologs MLH1 and PMS2) and/or form mobile "sliding clamps" (Erie & Weninger, 2014; Peggy Hsieh & Yamane, 2008; Iyer et al., 2006; Kunkel & Erie, 2005).

Interestingly, eukaryotes and most prokaryotes do not have a MutH homolog; instead, MutLα possesses the endonuclease activity that introduces a nick into the DNA on either side of the error. This activity is induced via an ATP-dependent interaction with MutSα or MutSβ (Constantin, Dzantiev, Kadyrov, & Modrich, 2005; F. A. Kadyrov et al., 2007; F. A. Kadyrov, Dzantiev, Constantin, & Modrich, 2006). If the nick introduced by MutLα occurs 5' of the error, the daughter strand can be excised by Exo I, leaving single stranded regions that are stabilized by the SSB homolog RPA until DNA Polymerase δ refills the gap (Constantin et al., 2005; Genschel, Bazemore, & Modrich, 2002; Longley, Pierce, & Modrich, 1997). Alternatively, the error-containing daughter strand may be removed by the DNA Polymerase itself via strand displacement synthesis. The resulting flap is then subsequently processed by FEN I (Kadyrov et al., 2009). In either case, the remaining nick is sealed by DNA ligase (Constantin et al., 2005).

In the absence of a MutH homolog, the eukaryotic mechanism for discrimination between the daughter and template DNA strands is achieved through an interaction between MutLα and Proliferating Cell Nuclear Antigen (PCNA), the processivity factor used in DNA replication. PCNA is loaded onto nicked DNA with a specific orientation, presumably at the replication fork. This orientation relative to the nicked strand positions a specific interaction surface of PCNA toward MutLα, thereby directing MutLα to preferentially nick the daughter strand (Pluciennik et al., 2010).

Thermus aquaticus (*Taq*) MMR (Figure 1.1C) shares many mechanistic similarities with both *E. coli* MMR and human MMR. Like *E. coli* MMR, *Taq* MutS and *Taq* MutL are asymmetric homodimers, meaning that both monomers within the dimer have identical primary structure but adopt slightly different conformations upon dimerization (Obmolova, Ban, Hsieh, & Yang, 2000). This asymmetry is somewhat mimicked in eukaryotes where homologs of MutS and MutL heterodimerize to form functional MMR proteins. It is also well established that *Taq* MutS also undergoes ATP-dependent conformational change to form a mobile "sliding clamp" state (Jeong et al., 2011; Qiu et al., 2012, 2015). Unlike *E. coli*, *Taq* does not possess a MutH homolog, though the remaining downstream events (i.e. error excision, resynthesis, and ligation) are thought to be similar. Due to its similarity to eukaryotic MMR, the *Taq*

MMR system is a more suitable model system for understanding human MMR, especially during MMR initiation by MutS and MutL. In addition, *Taq* MMR proteins are relatively easy to purify and study compared to their eukaryotic counterparts.

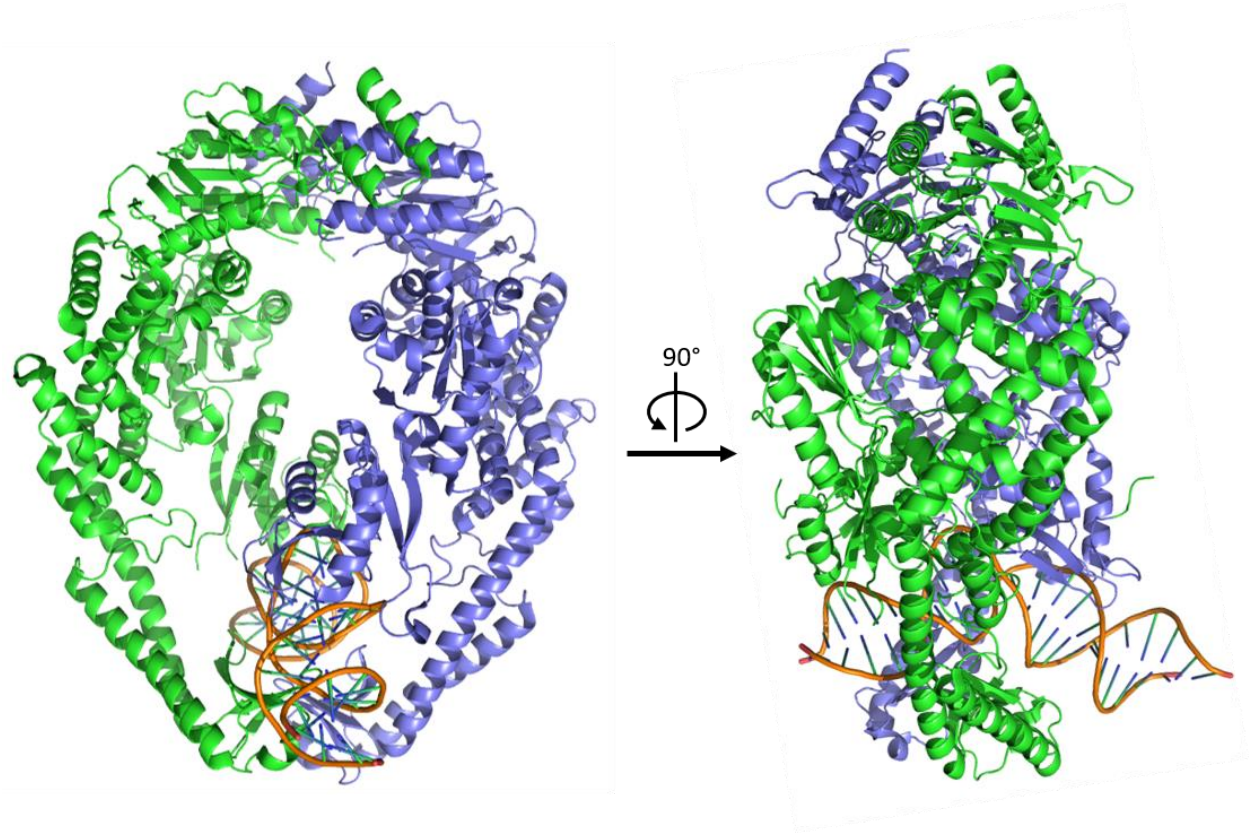

Figure 1.2: Crystal Structure of *Taq* MutS.

Front view (left) and side view (right) of *Taq* MutS in complex with DNA containing a single thymine insertion (PDB ID: 1EWQ). The two subunits of the homodimer are colored blue and green, and the DNA is shown in orange (Obmolova et al., 2000).

Error Recognition by *Thermus aquaticus* MutS

MutS Structural Information

Errors in the genome are vastly outnumbered by properly base-paired (homoduplex) DNA, making error recognition an enormous challenge. This monumental task is undertaken by MutS and the MutS homologs. In *Taq*, MutS forms an obligate homodimer (hereafter referred to as MutS) that contains a DNA binding domain and two adenosine nucleotide binding sites. Upon binding DNA, the homodimer becomes asymmetric, and this asymmetry is conferred to the nucleotide binding sites. As is apparent in the crystal structure (Figure 1.2), MutS induces a significant bend in the DNA of approximately 60 degrees at the mismatch. This kink in the DNA is thought to be important for mismatch recognition, as the DNA near the error is expected to be more inherently flexible. By better tolerating protein-induced DNA bending, the position of the error becomes a localized energy minimum. Upon binding to the site of the error, the erroneous base is rotated 3Å out into the minor groove of the double helix and stacks with a conserved phenylalanine residue in the DNA binding domain of MutS (Obmolova et al., 2000). *Taq* MutS retains some affinity for homoduplex DNA (K_D = 20-35 μM), though, as a result of the specific interactions formed upon error recognition, MutS has a greater affinity for DNA containing a mismatch (K_D = 5-40 nM) (Yang, Sass, Du, Hsieh, & Erie, 2005).

The Role of DNA Bending in Error Recognition by MutS

An atomic force microscopy (AFM) studies in our lab further implicated DNA bending in mismatch recognition and MMR initiation by *Taq* MutS (Tessmer et al., 2008; Wang et al., 2003). In this study, *Taq* MutS interacting with DNA containing a single thymine IDL (T-bulge), a GT mismatch, or no error was visualized. The resulting images revealed two distinct populations of protein:DNA complexes with different degrees of DNA bending. In the "bent" population, the DNA bend angle was approximately 42°, while the DNA in the "unbent" population was fairly straight (a DNA bend angle of ~0°). Notably,

the “unbent” population was only found at the position of the GT mismatch or T-bulge, and the protein:DNA complexes observed on homoduplex DNA were all found to be bent. These results led to the development of a model wherein a mismatch-dependent “bent” to “unbent” transition was proposed to be vital for signaling downstream MMR events (Wang et al., 2003).

A single-molecule fluorescence study in our lab further explored DNA bending conformations adopted by *Taq* MutS:DNA complexes. Six different DNA-bending conformations were identified using DNA containing a GT mismatch, two of which correspond to the “bent” and “unbent” populations seen in the AFM studies. Importantly, this study revealed that MutS-induced DNA bending is highly dynamic, and certain transitions between states were preferentially observed, most notably the “bent” to “unbent” transition (Sass, Lanyi, Weninger, & Erie, 2010). Subsequent single-molecule fluorescence studies on DNA substrates containing either a T-bulge or a CC mismatch revealed that preservation of this “bent” to “unbent” transition correlated with repair efficiencies (Derocco, Sass, Qiu, Weninger, & Erie, 2014).

These data, taken together, suggest that error recognition by *Taq* MutS in the absence of ATP can occur through the following mechanism: MutS first binds to and bends homoduplex DNA and then slides along the DNA until it encounters the error. Once at the error, MutS bends the DNA into the initial recognition complex observed in the crystal structure. The MutS:DNA complex can then undergo a series of dynamic conformational changes that leads to the formation of the ultimate recognition complex in which the DNA is unbent (Wang et al., 2003).

ATP-induced Conformational Changes in Taq MutS

Following error recognition, *Taq* MutS undergoes a series of ATP-dependent conformational changes to signal for downstream MMR events. Like *E. coli* MutS and the human homologs, *Taq* MutS in the presence of ATP and DNA containing a mismatch can form a sliding clamp conformation. Once in this state, MutS can freely diffuse along the length of the DNA (Jeong et al., 2011; Qiu et al., 2012,

2015). While the existence of this state is well established, its purpose remains unclear. The sliding clamp is thought to be involved one or more downstream events of MMR, such as recruitment of MutL or an exonuclease. The purpose of the sliding clamp may also be to clear the error site for recognition by a second MutS and/or searching for the strand discrimination signal.

Formation of *Taq* MutS:MutL:DNA Ternary Complexes

Following error recognition, MMR initiation continues in all species with the ATP-dependent recruitment of MutL to the mismatched site by MutS. The *Taq* MutL homodimer (in humans, the heterodimer MutLα) is then responsible for nicking the error-containing daughter strand on either side of the error. To accomplish this task, the N-terminal domain of each MutL monomer has a DNA binding domain, while the C-terminal domain has endonuclease activity. These two domains are connected by a flexible linker that is thought to be intrinsically disordered. While X-ray crystallography has provided structural information for the N- and C-terminal domains of *E. coli* MutL (Ban, Junop, & Yang, 1999), a structure of full length MutL for any species has remain elusive, likely as a result of the disordered linker region.

Several models have been proposed to describe the nature of the MutS:MutL:DNA ternary complexes that vary in terms of size, structure, position on the DNA relative to the mismatch, and relative stoichiometries (Figure 1.3A) (Iyer et al., 2006). In the first model, MutS first recognizes the mismatch and then undergo ATP-dependent movement along the DNA in the sliding clamp state. This state then recruits MutL forming mobile MutS:MutL complexes. In the second model, MutS recognizes the error. ATP-dependent MutL recruitment by MutS to the site of the mismatch leads to MutL polymerization forming long MutL tracks along the DNA that are nucleated by MutS. In the third model, MutS:MutL complexes bend or loop the DNA. It is important to note that these models are not mutually exclusive. Recently, the structure of a cross-linked complex of *E. coli* MutS and the N-terminal domain of MutL was

solved that revealed large conformational changes in MutS (Groothuizen et al., 2015). In the *Taq* system, single-molecule fluorescence experiments using labeled MutS and labeled DNA reveal that MutL can trap MutS at the mismatch. These trapped complexes were shown to be in a rapid conformational equilibrium between two states. These data have led to a refined model for MMR initiation in *Taq* that accounts for the ATP-dependent conformational changes of MutS following error recognition, as well as what occurs upon introduction of MutL (Figure 1.3B) (Qiu et al., 2015). However, while it is clear that MutS can bend and straighten DNA in the absence of nucleotides, particularly when bound at the site of an error, little is known about the conformation of the DNA throughout this model of MMR initiation, particularly in the presence of adenosine nucleotides.

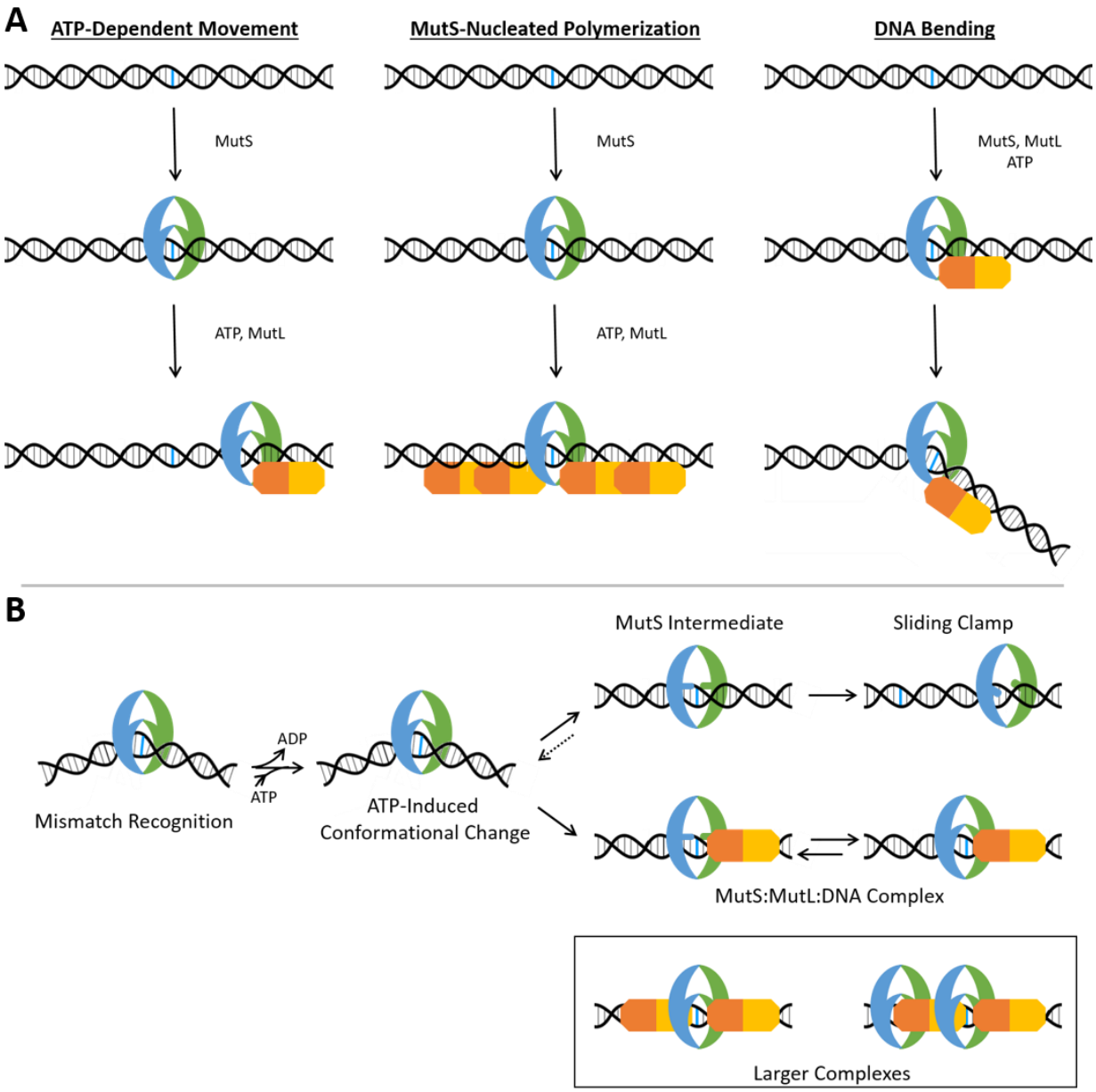

Figure 1.3: Models of MutS:MutL:DNA Ternary Complexes.

A) Three models depicting possible arrangements of MutS:MutL complexes on DNA during mismatch repair initiation. Adapted from (Iyer, Pluciennik, Burdett, & Modrich, 2006). **B)** A model of the ATP-dependent conformational changes in *Taq* MutS and *Taq* MutS:*Taq* MutL complexes on DNA during mismatch repair initiation. Adapted from (Qiu et al., 2015).

Single-molecule Fluorescence

Few tools are able to track the molecular events of complex biological processes; however, single-molecule fluorescence techniques are well suited for such studies. Unlike their bulk fluorescence counterparts, observations made using single-molecule techniques are not limited to the ensemble average. Rather, single-molecule methods are well suited for studying systems where the average does not represent any species. In addition, single-molecule methods are able to reveal relatively rare or transient events whose contributions to the ensemble would otherwise be hidden beneath the signals from more dominant populations in the ensemble. These rare or transient events are often vital to understanding the detailed molecular mechanisms behind complex, heterogeneous, and asynchronous molecular biology processes, such as DNA mismatch repair (Erie & Weninger, 2014; Gell, Brockwell, & Smith, 2006; Lakowicz, 2006).

Single-molecule Fluorescence Resonance Energy Transfer (smFRET)

Single-molecule fluorescence resonance energy transfer (smFRET) has emerged as a useful tool for detecting changes in the separation between two molecules of 10-100Å, a scale pertinent to molecular biological events. FRET was first described by Forster and Perrin and is defined as the non-radiative transfer of energy through long-range dipole-dipole interactions between two fluorophores dubbed the "donor" and "acceptor" (Förster, 1959). The efficiency of energy transfer from the donor fluorophore to the acceptor fluorophore depends on several factors, including: 1) the interfluorophore distance, 2) the relative orientations of the donor and acceptor fluorophores' transition dipoles, 3) the overlap between the donor fluorophore's emission spectrum and the acceptor fluorophore's excitation spectrum, and 4) the quantum yield of the donor fluorophore (i.e. the fraction of photons emitted by the donor fluorophore relative to the amount absorbed). FRET efficiency is defined mathematically as:

$$E = \frac{1}{\left(1+\frac{r^6}{R_0^6}\right)} \quad \textbf{Eq. 1.1}$$

where r is the interfluorophore distance and R_0 is known as the Forster radius. Due to the $1/r^6$ relationship shown in Eq. 1, FRET efficiency is exquisitely sensitive to changes in interfluorophore distance (Gell et al., 2006; Lakowicz, 2006).

The Förster radius R_0 is defined for a given pair of fluorophores as the interfluorophore distance at which the FRET efficiency is 50%. Importantly, R_0 represents the orientational-, spectral-, and quantum yield-dependence of FRET; thus, every FRET pair of fluorophores will have a unique Förster radius depending on these properties. Furthermore, since the fluorophore orientations, spectral properties, and quantum yields all depend on the local environment of the fluorophores, the Förster radius of a given FRET pair can vary depending on the labeling strategies and experimental conditions. Practically, however, these conditional variations are difficult to predict (Lakowicz, 2006).

When exciting only the donor fluorophore of a FRET pair, any emission from the acceptor dye will be indicative of FRET. Thus, the energy transfer efficiency can be described as fraction of the total detected intensity (i.e. the intensity of both the donor and acceptor) due to emission from the acceptor. Mathematically, this relationship is described as:

$$E = \frac{I_A}{I_D+I_A} \quad \textbf{Eq. 1.2}$$

where I_A and I_D are the intensities of the acceptor and donor fluorophore emissions, respectively.

This simplified expression for smFRET is problematic for several reasons. First, the donor and acceptor emission spectra may overlap. If this occurs donor emission, albeit minimal, can be detected at wavelengths assigned to acceptor emission. The acceptor intensity must be corrected to account for leakage of the donor signal into the acceptor channel. Second, there may be differences in the quantum yields and detection efficiencies of the dyes. These difference result in differential changes between the

donor and acceptor intensity when there are changes in transfer efficiency. Thus, a more exact description of smFRET is:

$$E = \frac{I_A - \alpha I_D}{\gamma I_D + (I_A - \alpha I_D)} \quad \textbf{Eq. 1.3}$$

where I_A and I_D are the raw, uncorrected intensities of the acceptor and donor fluorophore emissions, respectively. The α term accounts for leakage of a fraction of the donor signal into the acceptor signal. The factor γ corrects for the photophysical differences between the dyes and can be defined as:

$$\gamma = \frac{\varphi_A \eta_A}{\varphi_D \eta_D} \quad \textbf{Eq. 1.4}$$

where φ_A and φ_D represent the quantum yields of the acceptor and donor fluorophores, respectively, and η_A and η_D represent the detection efficiencies of the acceptor and donor fluorophores, respectively (Gell et al., 2006; Lakowicz, 2006; McCann, Choi, Zheng, Weninger, & Bowen, 2010).

Total Internal Reflection Fluorescence Microscopy

Emissions from single fluorophores are often sufficiently bright for modern detection technology; however, distinguishing the fluorophore emission from the background is much more challenging. Total internal reflection fluorescence (TIRF) microscopes are able to selectively excite only those fluorophores within ~200 nm of the surface of a slide by exploiting differences in the refractive index properties of the optical components and slides. Two optical setups for TIRF microscopes exist: through-prism and through-objective (Gell et al., 2006). This work was completed using a through-prism setup (Figure 1.4A).

When an excitation beam of light hits the interface of two materials with differing refractive indeces (n_1 and n_2), it can either be reflected or transmitted (possibly with some refraction). Total internal reflection (i.e. when all of the beam's intensity is reflected) occurs when the incident beam hits the interface at an angle that is greater than or equal to a critical angle (θ_C) defined as:

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad \textbf{Eq. 1.5}$$

At the point of reflection (Figure 1.4B), an electric field referred to as the evanescent field is created, which propagates through the interface. The intensity of this evanescent field decays exponentially as the distance from the interface increases according to the following relationship:

$$I(z) = I(0)e^{\left(-\frac{z}{d}\right)} \qquad \textbf{Eq. 1.6}$$

where *I(z)* is the intensity of the evanescent field at a distance *z* from the interface. *I(0)* is the intensity of evanescent field at the interface (i.e. where *z* = 0), and *d* is the penetration depth. The penetration depth of an evanescent field depends on several factors, including the refractive indices of the two materials (n_1 and n_2) and the wavelength and incidence angle of the excitation light (Gell et al., 2006; Lakowicz, 2006).

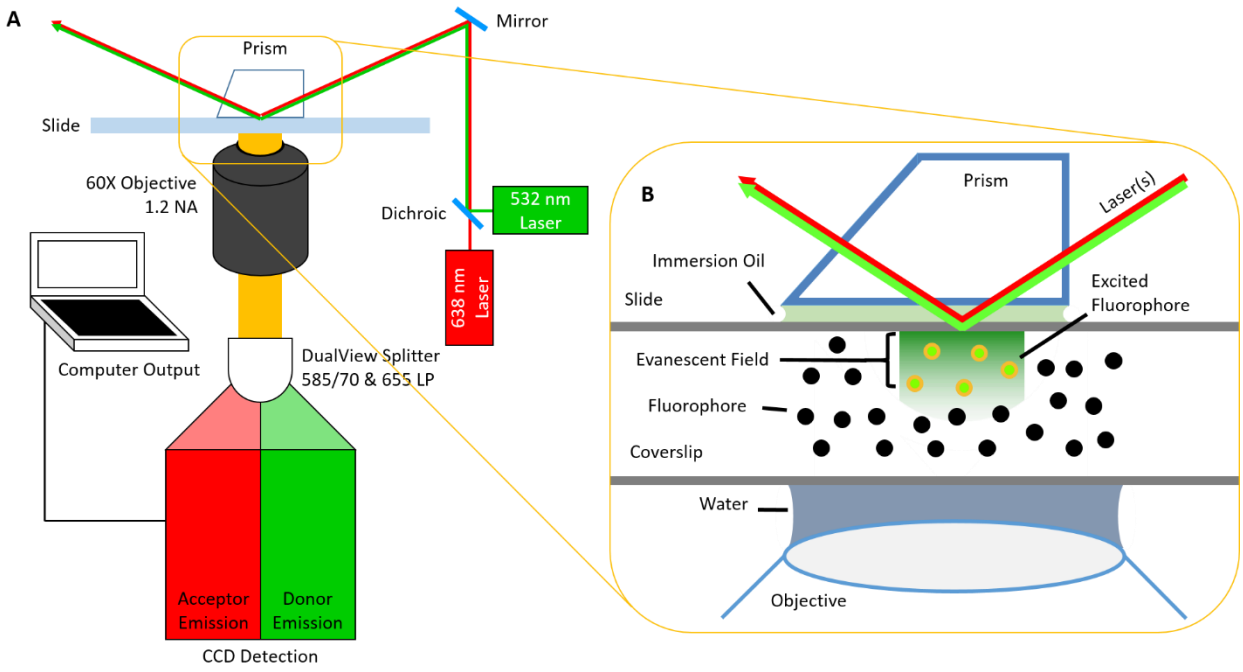

Figure 1.4: TIRF Microscopy.

A) Our optical setup: 532 nm and 638 nm lasers are used for TIRF excitation. The image is collected through a 60X water immersion, 1.2 numerical aperture objective and then split by a Dualview optical splitter (645 nm dichroic mirror). The donor and acceptor signals then pass through optical filters (donor: 585/70 bandpass; acceptor: 655 longpass) before detection by an emCCD camera. **B)** Total internal reflection at the quartz:water interface produces an evanescent wave that excites only those molecules within ~200 nm of the slide surface.

In practice, a quartz (n_1 = 1.55) slide and prism are used, and biological samples are typically studied in aqueous environments (n_2 = 1.33). 532 and 638 nm are common wavelengths of lasers used for excitation in smFRET experiments. Given these parameters and a 60° angle of incidence, the predicted penetration depths for the two excitation beams are 233 and 278 nm for 532 and 638 nm laser beams, respectively. Remembering that the evanescent wave decays exponentially as distance from the quartz:water interface increases, only those fluorophores very close to the slide (~200 nm) are sufficiently excited. By limiting off-target excitation in this way, background fluorescence is minimized.

Thesis Statement

This thesis aims to characterize the nucleotide dependence of *Thermus aquaticus* MutS-induced DNA bending throughout sliding clamp formation during mismatch repair initiation. The current model for MutS conformational changes that occur upon error recognition will be expanded to include DNA bending information. Largely, this aim will be achieved through the use of single-molecule fluorescence resonance energy transfer experiments designed to be sensitive to changes in DNA bending. Finally, a data analysis pipeline has been developed to streamline the analysis as well as optimize detection of small, yet significant changes in FRET.

Chapter 2 of this thesis outlines key components of the methodology used to study DNA bending using smFRET, including oligonucleotide design and data analysis strategies. Chapter 3 then applies these approaches to study DNA bending by *Taq* MutS during sliding clamp formation. Chapter 4 presents several unfinished extensions of the project presented in Chapter 3. Finally, Chapter 5 describes a brief side-project studying the link between MMR deficiencies and proteotoxic stress in yeast. The subsequent appendices include protocols for operating (Appendix A) and assembling (Appendix B) the two-color TIRF microscope used for single-molecule fluorescence studies in our lab, as well as a protocol for using the data analysis software written to complete this work (Appendix C).

CHAPTER 2: A GUIDE TO MONITORING PROTEIN-INDUCED DNA BENDING BY smFRET

Introduction

Protein-induced DNA bending is used as a signaling mechanism in many molecular biological processes, such as DNA mismatch repair (Iyer et al., 2006; Kunkel & Erie, 2005). Thus, measuring the extent of DNA bending, identifying DNA bending states, and characterizing the dynamics of exchange between these states can provide mechanistic insight into these processes. Several structural methods for observing protein-induced DNA bending exist, including atomic force microscopy and x-ray crystallography, but these methods are limited to static images of the protein:DNA complexes. Single-molecule fluorescence resonance energy transfer (smFRET), however, is able to sensitively detect changes in DNA bending in real time. This technique provides kinetic and conformational information that can elucidate the molecular details of these processes (Derocco et al., 2014; Erie & Weninger, 2014; Sass et al., 2010).

While smFRET can detect even small changes in DNA bending, the experimental design is vital to success. In addition, an efficient data analysis pipeline is essential to discern patterns from thousands of molecules worth of data. Here, key considerations for designing the labeling scheme of DNA oligonucleotides used in the smFRET experiments are outlined, and a data analysis pipeline for the resulting data is presented.

Designing Fluorescently Labelled DNA Oligonucleotides

When monitoring protein-induced DNA bending, the oligonucleotide labeling strategy must be thoughtfully designed such that the fluorescent properties of the fluorophores are sensitive to changes in the DNA conformation. Also, fluorescently labelled oligonucleotides are quite expensive (e.g. several hundred dollars per 100 nanomoles). Therefore, it is important to take care when deciding where and how to label the DNA. The specific fluorophores used, the position of the fluorophores on the DNA, and the fluorophore attachment chemistries have significant impacts on the sensitivity of the DNA substrates' fluorescent properties. Each of these parameters is discussed below.

Selecting the Fluorophores

There are several fluorescent tags that are commercially available for labeling oligonucleotides (e.g. TAMRA, the Cy dyes, Alexa dyes, and many others). When choosing which specific fluorophores to use, it is important to select dyes that have high quantum yields so they can be more easily detected. Also, the dyes' excitation and emission spectra should be compatible with your optical setup such that independent excitation and emission detection for each dye is possible. For FRET between the dyes to be possible, the emission spectrum of the donor fluorophore must overlap with the excitation spectrum of the acceptor fluorophore. Donor-acceptor dye pairs commonly used to label DNA in smFRET experiments include: Alexa 555-Alexa 647, TAMRA-Cy5, and Cy3-Cy5 (Derocco et al., 2014; Qiu et al., 2012, 2015; Sass et al., 2010).

When deciding between the many options for fluorophores, there are several unknown parameters that complicate dye selection. Notably, the structures of several commercially available fluorophores are proprietary information; thus, there is no way to know the exact nature of the perturbations being introduced into the DNA structure. Also, fluorophores that behave well in bulk fluorescence experiments may be prone to blinking or bleaching (i.e. transient or permanent dark states) in single-molecule experiments, making

them poor choices for monitoring DNA bending. Finally, some dyes are prone to interactions with the DNA (e.g. stacking with the nitrogenous bases), which may change their fluorescence properties in ways that do not depend on DNA bending or any other property. Due to these factors, it is often best to use a pair fluorophores which have proven successful in other experimental contexts.

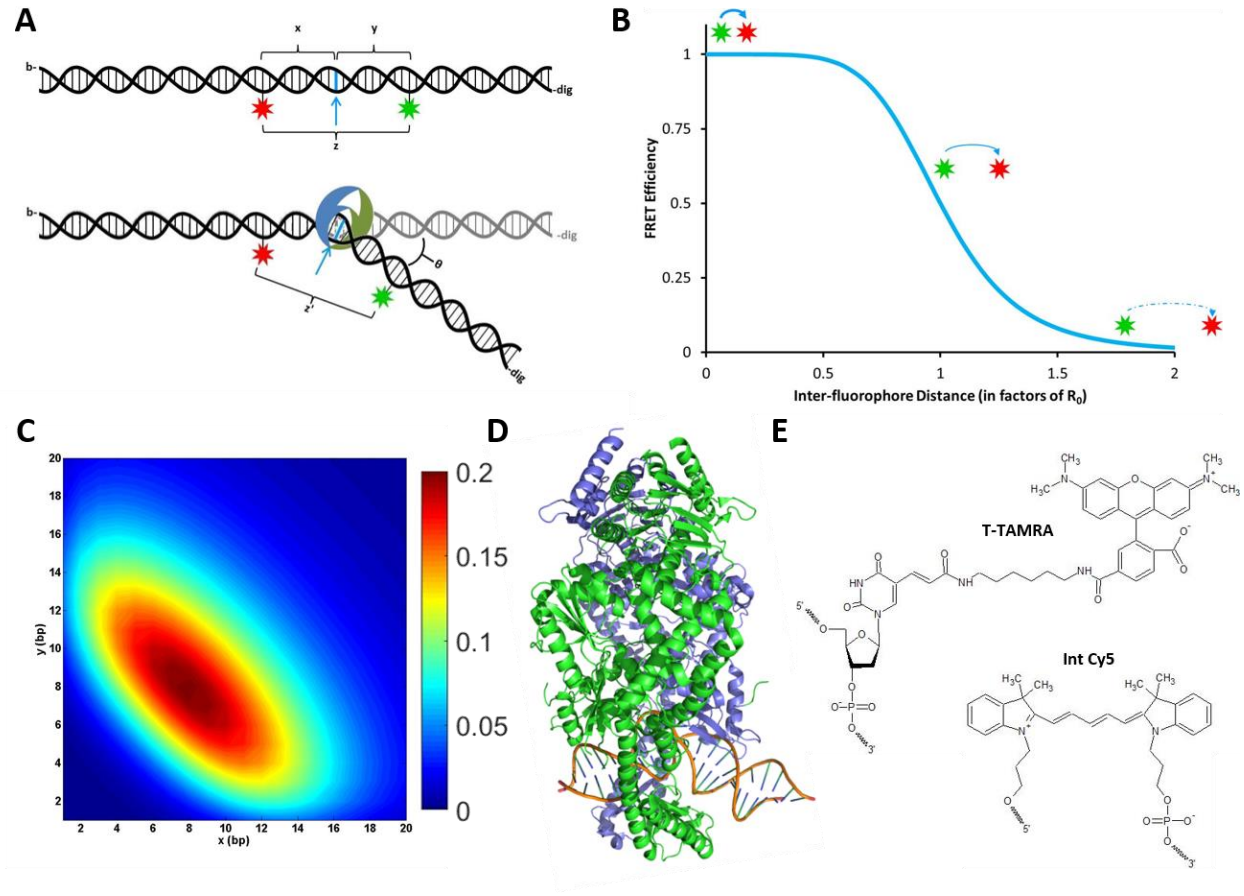

Figure 2.1: Considerations for Designing Fluorescent Oligonucleotides to Study DNA Bending

A) A schematic of a doubly-labeled oligonucleotide sensitive to changes in DNA bending. Donor (green) and acceptor (red) fluorophores flank either side of a site of protein-induced DNA bending (blue arrow). **B)** FRET curve showing the relationship between inter-fluorophore distance and FRET efficiency. **C)** Heat map showing the predicted change in FRET (according to the color scale) for oligonucleotides labeled in different positions (x and y from panel A). These simulations used the following parameters: $\theta = 60°$, $R_0 = 5$ nm, 1 bp = 0.34 nm. **D)** Crystal structure of *Thermus aquaticus* MutS showing DNA bending (PDB ID: 1EWQ). **E)** Two different attachment chemistries for labeling DNA internally. Shown are a thymine-linked TAMRA and a backbone Cy5.

Optimizing the Fluorophore Positions

Consider an oligonucleotide that is bent upon protein binding (Figure 2.1A) (Derocco et al., 2014; Sass et al., 2010). Such protein-induced bend in the DNA brings the two flanking arms of the DNA (x and y) nearer to each other. Thus, if a DNA substrate is labelled with a donor and acceptor fluorophore on either side of a site of protein-induced DNA bending (Figure 2.1A), the induced DNA bend angle (θ) will decrease the interfluorophore distance (from z to z′) thereby increasing the FRET efficiency between the two dyes. The sensitivity and reliability of this increase in FRET can be optimized by considering several parameters, including the predicted bend angle, the placement of the dyes relative to the site of DNA bending, and the protein footprint.

Changes in interfluorophore distance cause changes in FRET efficiency (E) following the equation:

$$E = \frac{1}{1+\left(\frac{r}{R_0}\right)^6} \qquad \textbf{Eq. 2.1}$$

where r is the interfluorophore distance and R_0 is the Förster radius of the donor:acceptor fluorophore pair (Figure 2.1B) (Gell et al., 2006; Lakowicz, 2006).

More significant bend angles (θ) will produce larger changes in interfluorophore distance, and thus, larger changes in FRET efficiency. Placing the fluorophores far from the site of bending (large values of x and y), will increase the change in interfluorophore distance before and after bending (z-z′); however, because FRET efficiency is related to $(1/r)^6$, large changes in interfluorophore distance will not always result in large changes in FRET efficiency. If $z >> R_0$ or if $z << R_0$, the FRET efficiency will not change significantly upon bending. Instead, placing the dyes such that the FRET efficiency is near 0.2-0.4 (i.e. $z > R_0$) in the absence of DNA bending is best. From this position, even small changes in DNA bending will decrease the interfluorophore distance, causing large changes in FRET efficiency.

If the arms of the DNA are assumed to be rigid, the interfluorophore distances before and after bending (z and z′) can be calculated. Likewise, the change in FRET efficiency can also be predicted for

every combination of labeling positions (x and y), base-by-base along the DNA (Figure 2.1C). In the resulting heat map, warmer colors indicate regions of the greatest change in FRET upon bending. A clear maximum is observed when both fluorophores are placed 8 base pairs from either side of the site of DNA bending.

Note that this simulation is limited by several factors. The calculations require prior knowledge of the expected bend angle as well as the Förster radius for the FRET pair used. The bend angle may be known if structural information (i.e. a crystal structure) of your protein:DNA complexes of interest is available. Förster radii are available for several FRET pairs, but these values are determined using fluorophores free in solution. Attachment to DNA can change the properties of the dyes, so it is impossible to know how the exact Förster radius of a given FRET pair in these contexts. Also, this simulation does not account for DNA twisting that may occur during DNA bending, nor does it consider slight differences in interfluorophore distance caused by placing both fluorophores the same DNA strand or on opposite strands of the helix. These simplifications are used, in part, because protein-induced twisting and its effects on changes in interfluorophore separation are difficult to predict.

Finally, it is crucial to position the labels such that they are not within the binding footprint of the protein being studied because the fluorescent properties of fluorophores are sensitive to changes in their local chemical environment. Protein-dye interactions often change the fluorescent properties of fluorophores, resulting in bending-independent changes in fluorescence intensity and/or FRET. Thus, labeling positions very near the site of DNA bending (i.e. low values of x and y) are often not ideal. For example, in the crystal structure of the *Thermus aquaticus* DNA repair protein MutS (Figure 2.1D), protein:DNA interactions occur as far as 8 nucleotides from the site of MutS-induced bending (Obmolova et al., 2000). Thus, fluorescent labels should be placed outside this range. If the greatest predicted change in FRET is found to be at labeling positions within the protein's footprint, changing to a different FRET pair with a greater R_0 is recommended.

Choosing the Fluorophore Attachment Chemistry

Several chemical strategies exist to link fluorophores to the DNA (Figure 2.1E). Labeling the 5′- or 3′-ends of the DNA is one option; however, being limited to only the ends of the DNA may be problematic. While internally labeling DNA oligonucleotides provides much more flexibility in choosing labeling positions, fewer commercially available options for fluorophores exist. There are two major types of internal labels: 1) fluorophores are covalently attached to thymine bases via a flexible carbon linker extending from the major groove in the DNA, and 2) certain dyes (e.g. the Cy dyes) can be incorporated directly into the DNA backbone. Note that these backbone fluorophores are rigidly locked into a specific orientation relative to the DNA backbone as they are covalently attached at both ends. Conversely, fluorophores on a flexible linker have more conformational freedom. As the Förster radius of a given FRET pair of fluorophores depends, in part, on the relative orientation of the transition dipoles of the two fluorophores (Lakowicz, 2006), limiting both fluorophore moieties' conformational freedom may lead to unpredictable changes in FRET upon DNA bending due to changes in their relative orientations. For more predictable changes in FRET, at least one freely rotating dye (i.e. those attached to flexible linkers) is recommended.

Other Considerations

To study protein induced DNA bending using single-molecule microscopy, the so-designed DNA oligonucleotides must be immobilized onto a slide. To this end, labeling the 5′ end with biotin provides an anchor to attach the oligonucleotide to the slide surface via an interaction with streptavidin. Streptavidin adheres well to slide surfaces, and this adherence can be further encourage by pre-treating the surface with biotinylated bovine serum albumin (BSA) (Derocco et al., 2014; Roy, Hohng, & Ha, 2008; Sass et al., 2010).

In several studies, blocking the ends of the DNA may be experimentally useful, such as when studying proteins that are topologically linked around the DNA. To prevent dissociation of such proteins from the end of the DNA oligonucleotides, the ends must be blocked. One end of the DNA is blocked by the slide surface. By attaching a digoxigenin moiety to the 5′-end of the unbiotinylated strand, end-blocking can be achieved using an anti-digoxigenin antibody (Qiu et al., 2012, 2015).

Optical Setup and Data Collection

Biotinylated fluorescently-labeled DNA oligonucleotides can be immobilized onto the surface of a quartz slide pre-functionalized with biotinylated BSA and streptavidin. Importantly, the slide surface coverage must not be overcrowded so that emissions from single DNA molecules can be resolved. These samples can then be imaged using a through-prism total internal reflection single molecule fluorescence microscope. For oligonucleotides labeled with TAMRA (donor) and Cy5 (acceptor), an appropriate optical setup (Figure 2.2A) is as follows: Donor and acceptor fluorophore excitation is achieved using 532 nm and 638 nm lasers respectively. The fluorophore emission is collected through a 60X water immersion, 1.2 N.A. objective, and the image is split by a DualView optical splitter with a 645 nm dichroic mirror. The donor and acceptor signals then pass through optical filters (i.e. a 585/70 bandpass filter for TAMRA, and a 655 longpass filter for Cy5) before detection by an emCCD camera. To observe changes in DNA bending over time, movies of approximately 1000 frames are collected using the following excitation sequence: 1) Excitation of the acceptor dye (~ 1 sec) to locate DNA molecules; 2) Excitation of the donor dye (~2 min) to monitor changes in FRET; 3) Excitation of the acceptor dye (~5 sec) to reveal whether the acceptor has photobleached (Derocco et al., 2014; Qiu et al., 2012, 2015; Sass et al., 2010).

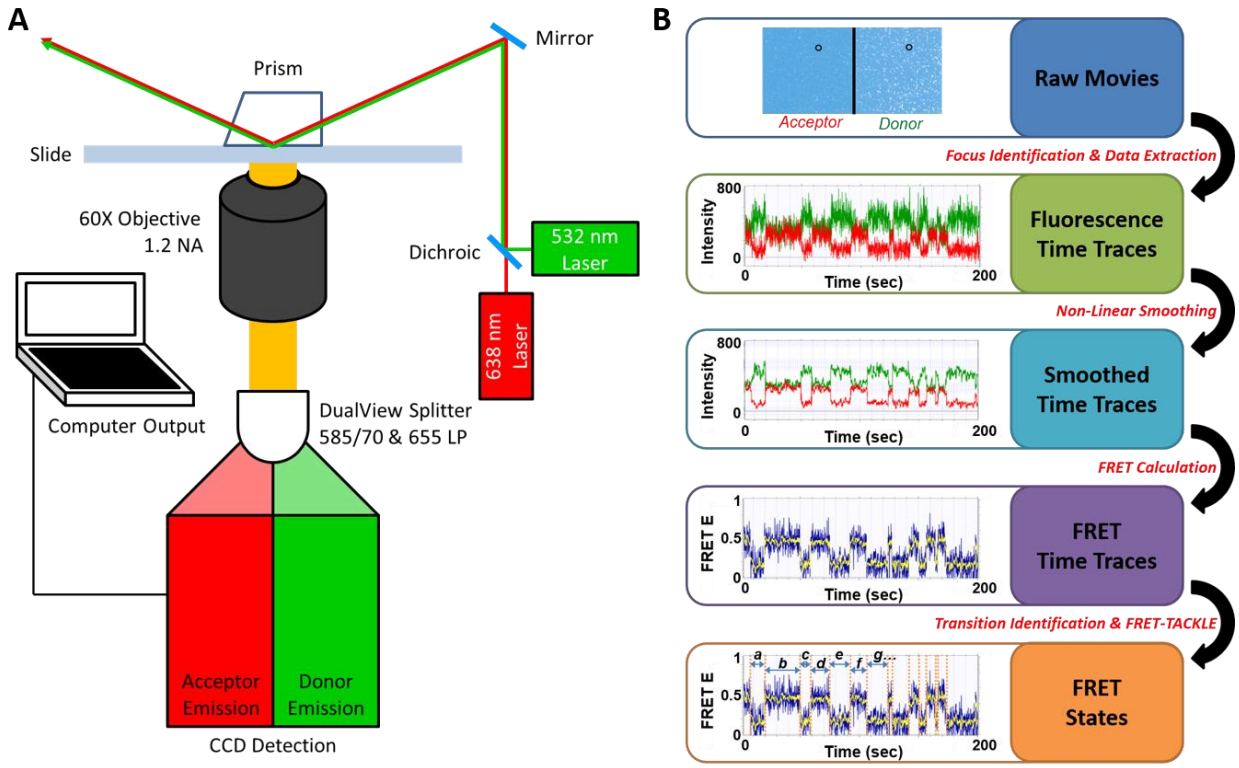

Figure 2.2: Optical Setup and Data Analysis Pipeline.

A) An example total internal reflection fluorescence microscopy set up suitable for the described DNA bending measurements. **B)** A schematic of the data analysis pipeline.

Data Analysis

Given proper slide coverage by a sample of labelled oligonucleotide and protein, a single movie collected as described above may contain fluorescence information for 200-500 DNA molecules. While already a large amount of information, the amount of data to analyze multiplies with varying experimental conditions and replicate experiments. Thus, an automated computational analysis approach is vital. However, automated analyses require verification to identify unexpected patterns and check for systematic errors. Below, a computational analysis pipeline is presented that systematically applies an analysis routine and outputs the results for a user interface:

Extracting the Fluorescence Time Traces of Individual DNA Molecules

Data acquisition by the aforementioned setup produces a movie made up of approximately 1000 frames. Each frame contains the fluorescence emission intensity gathered over a period of time for 512 X 512 pixels using a DualView splitter, the emissions from the donor and acceptor dyes are projected to the right and left halves, respectively, of each frame (Figure 2.2B, Top). So long as the slides are not overloaded with DNA substrate, emissions from each individual DNA molecule appears as two distinct foci (a focus from the acceptor emission on the left half and a focus from the donor emission on the right).

By initially exciting only the acceptor dye, local maxima on the left half of the image can be identified. The pixel coordinates of these maxima represent the positions of the DNA molecules within the image. These same pixel coordinates mapped to the other half of the frame should locate the emissions from the donor dye; however, the two halves of each frame must first be aligned to one another to account for misalignments in their optical paths. To determine the relative offset between the two halves of the image, beads with a broad emission spectra are imaged. These images allow an experimentally determined offset to be determined, which is then applied to the experimental movies.

Using an approach developed by the Weninger lab, the sum of the fluorescence intensities of a 3 X 3 pixel square centered on the local maxima pixel coordinates for the donor and acceptor dyes are determined for each DNA molecule. The local background intensity is subtracted, and those foci below a threshold intensity are discarded. Foci too close to the camera edges or another focus are also discarded. For those foci that remain, the intensity is determined for each from the movie for both the donor and acceptor, ultimately resulting in a “time trace” of the donor and acceptor fluorescence emission intensities as a function of time (Figure 2.2B, Second Row) (Derocco et al., 2014; Qiu et al., 2012, 2015; Sass et al., 2010).

Correcting the Donor and Acceptor Signals

Emissions from the donor and acceptor must be corrected to account for differences in the photophysical properties of the dyes. First, it is important to note that the donor and acceptor fluorescence emission spectra may not be completely distinct. That is to say that despite using optical filters and dichroic mirrors to separate emissions from the two dyes, the donor dye may still emit at wavelengths assigned to the acceptor. Even though these emissions are likely not large, they may still introduce significant error. This "leakage" into the acceptor detection window is proportional to the donor intensity. Thus, a proportionality constant can be empirically determined and used to subtract the "leaked" donor signal from the acceptor intensity at each time point using the following equation:

$$I_{A,cor} = I_{A,raw} - \alpha I_{D,raw} \quad \textbf{Eq. 2.2}$$

where α represents the leakage proportionality constant, $I_{A,raw}$ and $I_{A,cor}$ are the raw and corrected acceptor intensities, respectively, and $I_{D,raw}$ is the raw donor intensity at each time point (Gell et al., 2006; Lakowicz, 2006).

Additionally, the donor and acceptor dyes may have different quantum yields and/or different detection efficiencies. These differences can be corrected using a method described by McCann et al. using an empirically determined gamma factor by the following equation:

$$I_{D,cor} = \gamma I_{D,raw} \quad \textbf{Eq. 2.3}$$

where γ represents the empirically determined correction factor. $I_{A,raw}$ and $I_{A,cor}$ are the raw and corrected acceptor intensities, and $I_{D,raw}$ and $I_{D,cor}$ are the raw and corrected donor intensities (McCann et al., 2010).

Smoothing the Donor and Acceptor Time Traces Using the Chung-Kennedy Filter

Due to the sensitivity of single-molecule detection, the resulting fluorescence intensity time traces have relatively high noise. This noise is problematic when monitoring small changes in FRET. Consequently, if the change in DNA bending is small, it can be difficult to detect such changes in noisy data. Several methods exist to smooth noisy data, such as box-car averaging, but most of these smoothing methods do not preserve sharp edges, making the transition even more difficult to detect.

To overcome this limitation, Chung and Kennedy developed a non-linear smoothing algorithm designed specifically to smooth data while preserving edges (Chung & Kennedy, 1991; Haran, 2004). This algorithm smoothes data containing transitions (Figure 2.3A) by first determining the averages for windows of data of various size on either side of a given data point (dubbed the "forward" and "backward" average windows, Figure 2.3B). To preserve edges, "forward" and "backward" averages that contain transitions are given less weight in the overall average. The statistical weights assigned to the "forward" and "backward" averages are determined as follows: The standard deviations for windows of data on either side of the data point being smoothed (referred to as the "forward" and "backward" predictor windows, Figure 2.3C) are determined. The inverse of these standard deviations raised to a user-defined exponential term p are then used to calculate the statistical weights.

For example, consider a data point (at time t) with a transition occurring within the forward predictor window but not in the backward predictor window (Figure 2.3A). The forward predictor (Figure 2.3C, green) window would have a large standard deviation relative to the backward window (Figure 2.3C, red). Thus, the averages from the forward average windows (Figure 2.3B, green) would be assigned lower weight in the overall average, while the averages from backward average windows (Figure 2.3B, red) would be assigned a larger statistical weight. Smoothing is this way results in preservation of the transition in the forward predictor window.

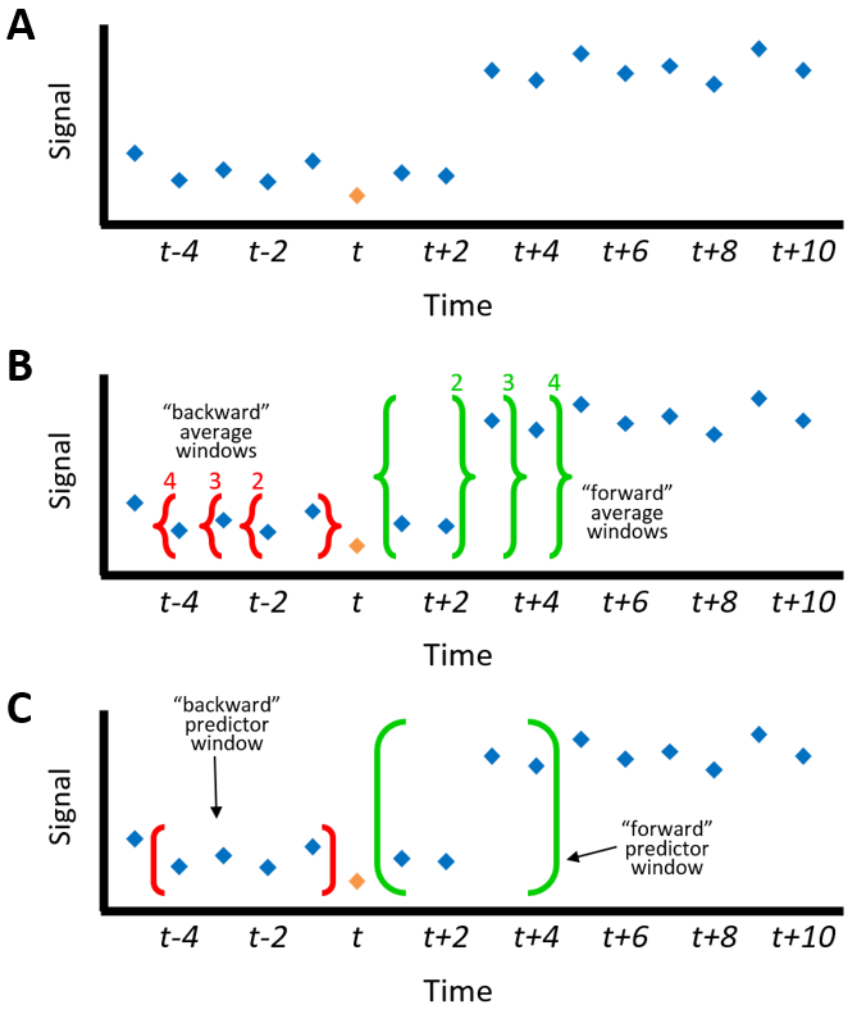

Figure 2.3: Chung-Kennedy Smoothing Algorithm.

A) An example of data containing a transition between $t + 2$ and $t + 3$. **B)** A schematic depicting three sizes of "forward" (green) and "backward" (red) average windows. **C)** A schematic depicting 4-point "forward" (green) and "backward" predictor windows.

Note that in this approach, the average windows and the predictor windows need not be the same size, nor is it necessary to use only one window size for each type of window. In fact, transitions that occur on different timescales can be preferentially preserved using windows of various sizes. Furthermore, the statistical weight term can be raised to the power of an empirically determined value (p) to exaggerate the different contributions of the forward and backward window averages. Thus, the input parameters of the

average window size(s), predictor window size(s), and exponent term(s) can be defined by the user, and the optimal values in a given application can be empirically determined. Over filtering (i.e. introduction of false transitions) can be minimized by optimizing these input parameters.

When properly applied to the donor and acceptor time traces, the signal-to-noise ratio is significantly improved (Figure 2.2B, Third Row), and transitions in the data are clearer. Notably, in single-molecule FRET experiments, transitions in the donor and acceptor traces are expected to be anti-correlated (i.e. if the donor intensity increases, the acceptor intensity should simultaneously decrease, and vice versa). Because transitions in the donor and acceptor traces are expected to be simultaneous, the Chung-Kennedy smoothing algorithm can be improved by using the sum of the predictor window standard deviations for the both the donor and acceptor to determine the statistical weights. As a result, simultaneous donor and acceptor intensity transitions will be more strongly preserved compared to uncorrelated donor and acceptor changes in intensity (Haran, 2004).

Screening Time Traces for Data Quality

To expedite analysis of large amounts of single-molecule data, time traces of insufficient quality should be discarded from further analysis. There are several reasons a time trace or portions of a time trace may not be worth analyzing. For example, the detected fluorescence emission intensities may be too low to reliably detect changes in intensity, or they may be too high to represent emission from only one molecule. More commonly, either the donor or acceptor fluorophore (or both) may permanently or temporarily loose its fluorescence properties (i.e. bleach or blink) during data acquisition.

To identify unanalyzable regions of time traces, the fluorescence intensity of each data point can first be checked against user-defined minimum and maximum thresholds. To assess for bleaches or blinks in the remaining data, the average and standard deviation of a moving seven-point window can be used to

determine a 95% confidence interval for each data point. If this interval includes zero, the data point can be considered as part of a bleaching or blinking event and discarded from analysis.

Time traces made up entirely of data points that cannot be analyzed can be discarded out right. Those containing both analyzable and unanalyzable regions, however, may contain useful information (e.g. prior to photobleaching). While this screening process greatly increases the efficiency of data processing, false-positives and/or false-negatives may occur. For example, individual data points in analyzable regions may be falsely identified as unusable by this method. For this reason, the quality of each data point can be iteratively compared to its neighboring data points and, in the event of discrepancies, changed to conform to its neighbors.

Calculating FRET and Identifying Transitions in the FRET Time Traces

Regions of the remaining donor and acceptor time traces that are of sufficient quality to analyze can then be used to calculate single-molecule FRET efficiency using the following equation:

$$E = \frac{I_A}{I_A + I_D} \qquad \textbf{Eq. 2.4}$$

where I_A and I_D represent the corrected and smoothed fluorescence emission intensities of the acceptor and donor fluorophores, respectively (Derocco et al., 2014; Gell et al., 2006; Lakowicz, 2006; Qiu et al., 2012, 2015; Sass et al., 2010). The resulting FRET time trace (Figure 2.2B, Fourth Row) contains several pieces of information: 1) the FRET at any given point in time is a readout of the interfluorophore distance, which represents the different DNA bending states; 2) the dwell time of a given FRET state provides kinetic information, such as the characteristic lifetimes associated with each DNA bending state; and 3) the transitions provide insight into the preferred pathways of conformational changes. To extract this information, it is crucial to identify the time point at which the FRET time trace undergoes a transition.

Several methods exist to detect transitions in data, though each has its own limitations. False-positives (i.e. finding transitions that do not exist) and false-negatives (i.e. missing transitions that should be detected) can both hinder interpretation of the data. Applying the same transition detection method at different levels of stringency can minimize these mistakes, as transitions that withstand more stringent thresholds are more likely to be "real". Comparing the results of multiple methods, can also help to overcome the limitations of using just one technique, as different detection methods will be better suited for detecting different types of changes (e.g. short- vs. long-lived states) and will have different limitation. Edges detected by multiple methods at multiple stringency levels can then be assigned a higher confidence score. Presented here are two transition detection methods that can be independently applied at multiple thresholds to smFRET data.

Method 1: The Gaussian Kernel Method (Sass et al., 2010)

Mathematically, transitions in continuous functions can be identified by finding inflection points (i.e. maxima and minima in the first derivative of the time trace). Unfortunately, the FRET time traces are made up of discrete data points that have many apparent inflection points due to the significant noise in the signal. These issues can be circumvented by first convolving the FRET time traces with a Gaussian kernel of various widths and subsequently detecting inflection points in the convolved data. To ensure only "real" transitions are kept, a threshold can be incorporated. By changing the rigor of the threshold, the remaining "real" transitions can also be scored for confidence.

Method 2: The Chung-Kennedy Method

The previously described smoothing algorithm developed by Chung and Kennedy uses increases in the predictor windows' standard deviations to calculate statistical weights (Chung & Kennedy, 1991; Haran, 2004). These increases in standard deviation can also be used to detect transitions in the FRET time traces as follows:

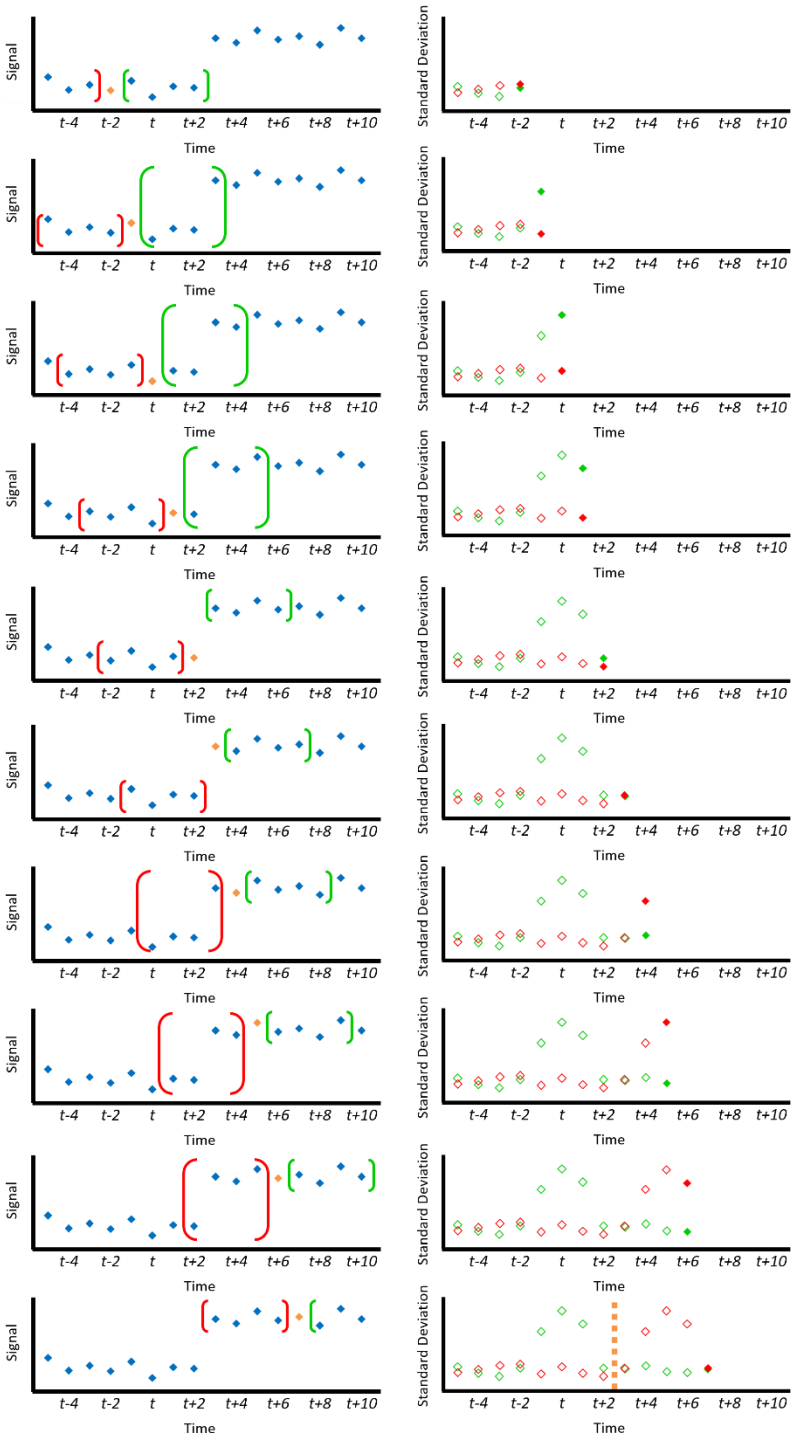

Figure 2.4: A Novel Transition Detection Method Based on the Chung-Kennedy Filter.

"Forward" and "backward" predictor windows (left column, green and red brackets, respectively) are shown for successive time points (blue diamonds). Standard deviations for the depicted predictor windows (right column) reveal two identical peaks. The orange dotted line depicts the midpoint between these peaks (between $t + 2$ and $t + 3$) and identifies the time of the transition in the data.

Consider a data point at time t with a transition occurring between time $t + 2$ and $t + 3$ (Figure 2.4). Using a predictor window of 4 data points, the standard deviation of both the forward and backward predictor windows can be calculated at each value of t. (Figure 2.4 depicts these calculations for $t - 2$ to $t + 7$.) In this example, the standard deviation of the forward predictor window at time t is at a local maximum. This also occurs in backward predictor window at time $t + 5$ because these windows contain the same range of data points. Thus, transitions can be detected by finding the midpoint between local maxima in the standard deviations of the forward and backward predictor windows. To ensure only "real" transitions are kept, only the highest percentile local maxima (e.g. 95^{th}-99^{th} percentile) are considered. By changing this percentile, the transitions can then be scored for confidence.

Notably, this method's ability to detect transitions is depends on the noise in the FRET time traces, as those traces with low signal-to-noise will have inherently high standard deviations, confounding the results. The most common source of false-positives is bleaching and blinking events where either the donor or acceptor fluorescence intensity is approximately zero. Very low donor or acceptor intensity can cause the FRET time trace to be very noisy, oscillating between 1 and 0. To circumvent this issue, regions of data previously identified as bleaches and blinks can be assigned a constant FRET (e.g. 0 or -1), which eliminates their standard deviations and allows this transition detection method to function properly.

Alignment and Confirmation of Transitions in the FRET Time Traces

Typically, the timing of each transition identified by both methods are in agreement. However, the two described methods occasionally produce slightly disparate results. In these instances, any transitions occurring within 0.3 seconds can be empirically aligned taking the weighted average of the time detected by each method. The weights used in this average can be determined using the confidence scores provided by each method.

Both of the methods described here require user input to choose the appropriate thresholds. In the event that transitions are being missed (false-negatives), the thresholds ought to be made less stringent. More often, though, too many "unreal" transitions are identified (false-positives). To test the significance of each transition, the FRET efficiency between each transition can be averaged, and the averages of adjacent FRET states can be subjected to a t-test. If two states are deemed not statistically significantly different from one another (p level 0.05), then the transition between the two states can be identified as a false-positive and discarded.

User Interaction and FRET-TACKLE

The resulting set of transitions (Figure 2.2B, Bottom Row) represents a best approximation of the simultaneous donor and acceptor changes in intensity. In the analysis described so far, the user is only required for the initial input parameters, such as smoothing windows and transition detection thresholds; the rest of the analysis can be completed by a computer in batch. Upon completion of the batch analysis, the computationally discovered transitions can be verified by the user, and any remaining false-positives in each molecule's FRET time trace can be discarded by hand. This process, though tedious, can be crucial to recognizing patterns or detecting systematic errors in the computational approach.

Once the transitions have been determined, FRET-TACKLE (FRET Transition Analysis Coupled with Kinetic Lifetime Evaluation) analysis can extract the pertinent mechanistic information from the compiled data (Derocco et al., 2014; Sass et al., 2010). In this method, distinct molecular conformations are identified by their characteristic FRET and kinetic lifetime properties. This approach allows molecular states with the same extent of DNA bending (i.e. the same FRET) but distinct kinetics to be distinguished, and vice versa.

Conclusion

Using the design strategies and analysis approach outlined here, smFRET data monitoring protein-induced DNA bending can be produced and interpreted. As previously discussed, DNA oligonucleotides used in these experiments must be thoughtfully labelled with the proper fluorophores in the proper positions using the proper attachment chemistries. A mistake in any of these considerations not only produces an oligonucleotide that is insensitive to changes in DNA bending, but is often costly. Furthermore, an automated analysis approach can allow the plethoric amounts of data to be more efficiently analyzed. By systematically locating foci, extracting and smoothing intensity time traces, determining which data are usable, and detecting and verifying transitions in the data, little user input is needed. Final analysis of the distribution of FRET efficiencies, the characteristic lifetimes, and the preferred transitions reveals conformational, kinetic, and pathway preference information for the protein:DNA complexes being studied, which can provide mechanistic insights. Examples of this approach successfully applied to understanding DNA bending by MutS during DNA mismatch repair can be found in this thesis (Chapter 3).

CHAPTER 3: CHANGES IN DNA BENDING CORRELATE WITH MUTS CONFORMATIONAL CHANGES DURING SLIDING CLAMP FORMATION

Introduction

Errors introduced during DNA replication must be corrected to maintain genomic stability. DNA mismatch repair is a biochemical pathway that increases the fidelity of DNA replication 100-fold by correcting misincorporated bases or insertion/deletion loop errors. The proteins involved in mismatch repair proteins are also involved in a range of other biochemical processes, including DNA damage response and double strand break repair. Unsurprisingly, mutations in the mismatch repair proteins are associated with several types of cancer, including certain hereditary nonpolyposis colorectal cancers (Iyer et al., 2006; Kunkel & Erie, 2005).

Mismatch repair is initiated by the protein MutS, which is responsible for recognizing the replicative errors in the DNA. Though prokaryotic MutS homodimerizes and the eukaryotic MutS homologs function as heterodimers, the mechanism of error recognition appears to be conserved (Iyer et al., 2006; Kunkel & Erie, 2005). Functional MutS dimers (referred to hereafter as simply "MutS") possess DNA-binding and ATPase activities. MutS can bind to homoduplex DNA and slide along the DNA in search of an error. Upon encountering an error, MutS forms specific interactions with the mispaired base(s) and can then undergo a series of nucleotide-dependent conformational changes to signal for downstream repair events (Alani et al., 2003; Jeong et al., 2011; Obmolova et al., 2000; Qiu et al., 2012, 2015). Most notably, both prokaryotic and eukaryotic MutS form a mobile "sliding clamp" state that can move along the length of the DNA. Forming this state is known to be both ATP- and mismatch-containing DNA-dependent, and it is believed to be important in mismatch repair signaling.

Recent studies using single-molecule fluorescence resonance energy transfer (smFRET) successfully characterized the conformational and kinetic properties of *Taq* MutS during sliding clamp formation (Jeong et al., 2011; Qiu et al., 2012, 2015). In these studies, MutS was tagged with a donor fluorophore and the DNA containing an error tagged was tagged with an acceptor fluorophore such that when MutS bound to the error, FRET between the dyes could occur. In the presence of saturating ATP, *Taq* MutS bound to the error in a state with high FRET. Qiu et al. observed that a subset of these mismatch-binding events (~20%) underwent a preferred pathway of changes in FRET (high FRET → intermediate FRET → zero FRET). These changes were attributed to sliding clamp formation, as the final zero FRET could only occur if MutS had moved away from the acceptor dye (i.e. far from the mismatch) while still bound to the DNA. In addition, kinetic analysis of the total time MutS spent bound to T-bulge substrates yielded a characteristic lifetime of 11.7 sec (Qiu et al., 2012, 2015). In additional experiments, the DNA binding domains of each monomer of MutS was labeled with one dye from a FRET pair. Using this method, conformational changes within the MutS dimer were identified during sliding clamp formation. Importantly, the lifetimes associated with the states identified in the intramolecular FRET experiments correlated with those observed in the protein-to-DNA FRET experiments (Qiu et al., 2012).

These smFRET results were combined to create a model for the conformational changes experienced by *Taq* MutS during sliding clamp formation (Figure 3.1). In this model, free MutS exists in a conformational equilibrium between an "open" and "closed" state when not bound to DNA (Figure 3.1, state i). Upon mismatch recognition (Figure 3.1, state ii), MutS binds to the DNA in a high FRET state in a "closed" conformation. A subsequent conformation change with no associated change in FRET was identified by the kinetic analysis and was proposed to be associated with exchange of ADP for ATP (Figure 3.1, state iii). This state then transitions to a state with intermediate FRET (Figure 3.1, state iv) associated with movement of the DNA binding domains. Finally, the sliding clamp (Figure 3.1, state v) with zero FRET is formed, and the DNA binding domains fully open. Importantly, this model depicts changes in

DNA bending throughout the process of sliding clamp formation without evidence for these changes (Qiu et al., 2012, 2015).

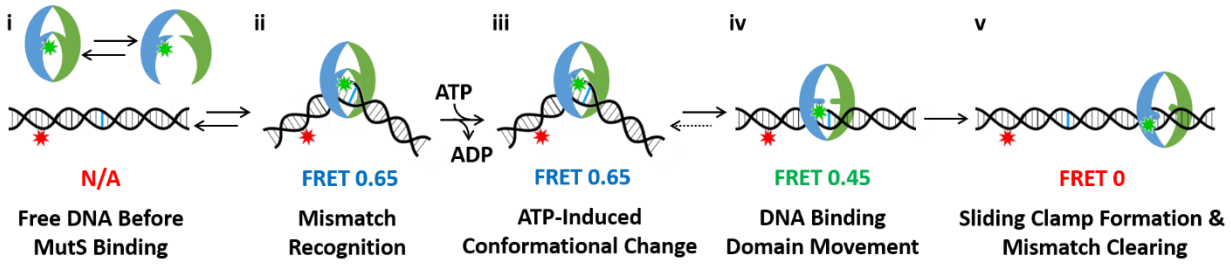

Figure 3.1: The existing model of sliding clamp formation by *Taq* MutS.

(i) Prior to binding DNA, MutS exists in a conformational equilibrium. **(ii)** MutS binds and bends the DNA at the error. MutS adopts a conformation where the two DNA binding domains are near each other. **(iii)** MutS exchanges ADP for ATP, and the conformation changes little. **(iv)** MutS undergoes a conformational change that slightly opens the DNA binding domains. **(v)** The DNA binding domains fully open as MutS becomes a sliding clamp and moves away from the mismatch. Adapted from (Qiu et al., 2015).

DNA bending by *Taq* MutS has been well characterized by X-ray crystallography and atomic force microscopy (Alani et al., 2003; Obmolova et al., 2000; Wang et al., 2003). This MutS-induced DNA bending is not static, and several single molecule studies have dissected the DNA bending dynamics of MutS:DNA complexes in the absence of adenosine nucleotides (Derocco et al., 2014; Sass et al., 2010). In these studies, a key "bent-to-unbent" transition has been identified, and preservation of this transition correlates with efficient repair (Derocco et al., 2014; Wang et al., 2003). However, little is known about the nucleotide dependence of these DNA bending conformations, nor is the contribution of DNA bending throughout process of forming the sliding clamp well understood. To fully understand the actions of MutS in mismatch repair initiation, we must elucidate the molecular changes occurring in both MutS and the DNA.

Here, we use smFRET to monitor changes in DNA bending induced by *Thermus aquaticus* (*Taq*) MutS in the presence of saturating concentration of ADP or ATP. The results of these experiments are then

correlated to the previous smFRET studies via their characteristic kinetics. Notably, a pathway of DNA bending states was identified, and the kinetics of this pathway correlate well with the sliding clamp formation kinetics from the previous studies. Moreover, a conformational change previously identified only by kinetic analysis was characterized by changes in the DNA bending conformation. We then refine the model for sliding clamp formation to account for conformational changes in both MutS and the DNA.

Results

To study the nucleotide dependence of MutS-induced DNA bending during sliding clamp formation using smFRET, we designed a 68 bp oligonucleotide that is doubly-labeled with a FRET pair of dyes, TAMRA (donor) and Cy5 (acceptor). These dyes are separated by 19 bp flanking a single thymine insertion error (referred to herein as T-bulge) near the midpoint of the oligonucleotide (Figure 3.2A). This oligonucleotide is also biotinylated on one end so that it could be immobilized via an interaction with streptavidin-biotinylated BSA on the surface of a quartz slide. The cysteine mutant C42A/M88C *Thermus aquaticus* MutS with wild type ATPase and DNA binding activities is used in this study to allow for direct comparison to existing smFRET data (Qiu et al., 2012, 2015).

MutS:ADP bends DNA to a single bent conformation.

We first determined the DNA bending properties of *Taq* MutS in the presence of saturating concentrations of ADP. We measured the smFRET efficiency from the donor to the acceptor dyes on either side of a T-bulge (Figure 3.2A). On linear DNA in the absence of MutS, the FRET efficiency between the dyes remained constant over time with a peak centered around 0.25 (Figure 3.2C, black dotted cityscape). Upon introduction of 10 nM MutS and 2 mM ADP, anti-correlated changes in the donor and acceptor fluorescence intensity time traces were observed, resulting in brief periods of increased FRET efficiency followed by a return to the original FRET efficiency (Figure 3.2B). In each time trace, the FRET efficiency

was observed to transition between only two distinct FRET states, a low FRET state (Figure 3.2B, red arrows) and a high FRET state (Figure 3.2B, cyan arrows). A histogram of the average FRET efficiency of the low FRET states (Figure 3.2C, red bars) shows a peak centered around 0.25, similar to the free DNA distribution. A histogram of the average FRET efficiency of the high FRET states in each trace (Figure 3.2D) is shifted to higher FRET with a peak centered around 0.35. This result is consistent with MutS-induced DNA bending upon binding to DNA containing a T-bulge, as was observed previously in the absence of ADP (Derocco et al., 2014; Sass et al., 2010; Wang et al., 2003). The FRET transitions detected in these transitions are depicted in a transition density plot (TDP), where warmer colors represent more frequent transitions (Figure 3.2E). Two dominant types of transitions are observed: (1) a low FRET state transitioning to a high FRET state, and (2) a high FRET state transitioning to a low FRET state.

Using the observed change in FRET efficiency and the law of cosines, the DNA bend angle can be calculated; however this calculation is complicated by variations in the Förster radius caused by linking the fluorophores to the DNA and fluorophore-DNA interactions. If these complications are ignored, the observed change in the average FRET efficiency from ~0.25 to ~0.35 corresponds to a DNA bend angle of ~45°, which agrees well with the DNA bend angle observed in the crystal structure and in the AFM studies (Alani et al., 2003; Obmolova et al., 2000; Wang et al., 2003).

For most of the observed transitions to the high FRET state, the time spent in the high FRET state before returning to the low FRET state (i.e. the dwell time) was measured. The distribution of dwell times (Figure 3.2F, cyan bars) fit well to a single exponential decay (Figure 3.2F, black line), which is consistent with a stochastic transition out of a single discrete molecular state. The characteristic lifetime of these states (i.e. the reciprocal of the exponential decay term, k) is 3.2 ± 0.6 sec. This value is in good agreement with the characteristic lifetime found for donor-labeled MutS binding to acceptor-labeled T-bulge DNA in the presence of ADP observed in smFRET experiments at the same conditions (2.7 sec) (Qiu et al., 2012, 2015).

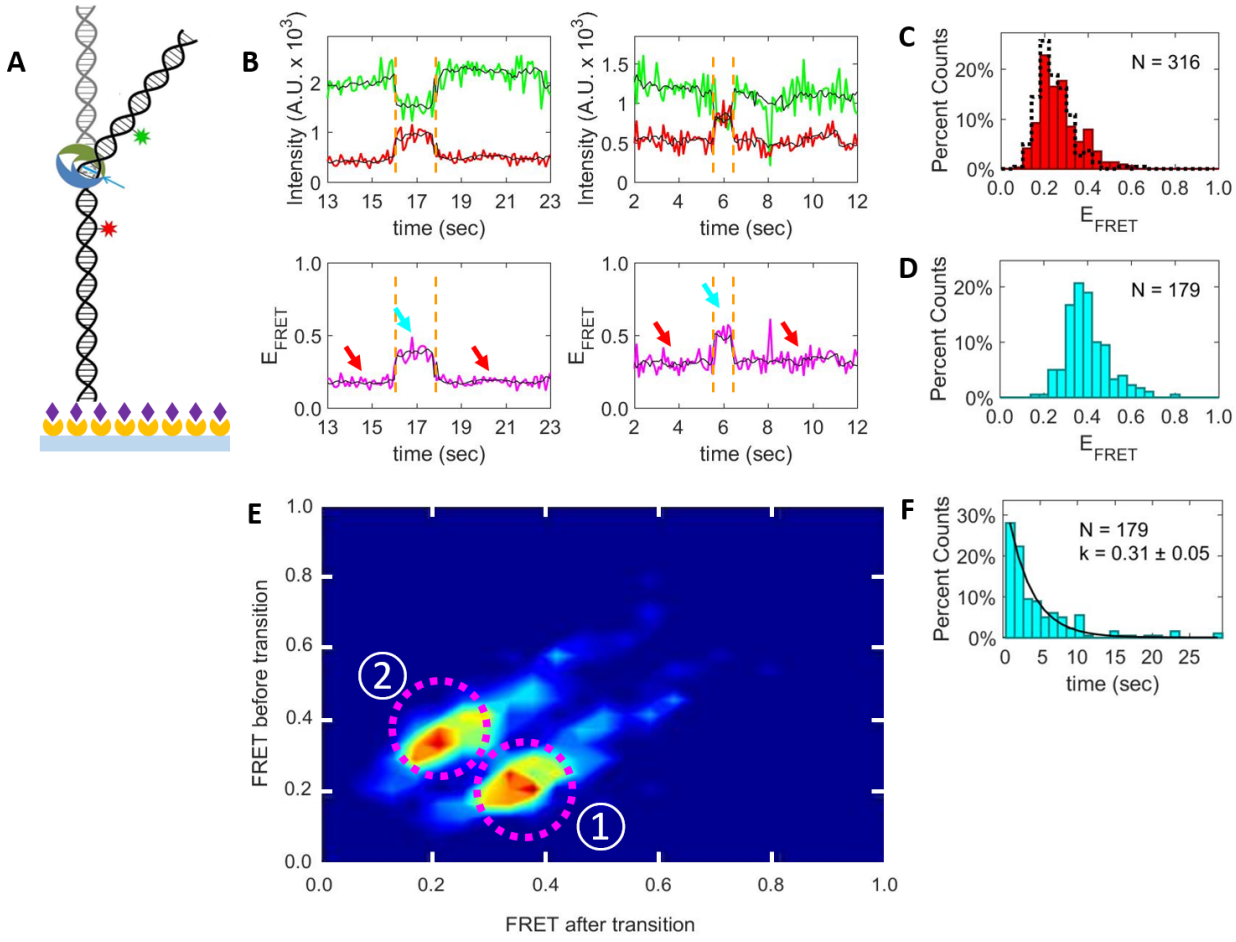

Figure 3.2: In the presence of ADP, *Taq* MutS bends T-Bulge DNA to a single bent state.

A) Schematic of a surface immobilized 70mer DNA molecule bound by *Taq* MutS at the location of the single thymine insertion (blue arrow). **B)** Example donor (green) and acceptor (red) intensity time traces and their corresponding FRET time traces (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ADP. The black line represents the smoothed signal. The red and cyan arrows denote low FRET and high FRET states, respectively. **C)** The distributions of FRET values for the low FRET states (red bars) and for DNA in the absence of *Taq* MutS (black dotted cityscape). **D)** The distribution of FRET values for the high FRET states (cyan bars). **E)** Transition density plot depicting the frequency of transitions converting between the low and high FRET states. **F)** The distribution of dwell times for the high FRET states (cyan bars) fit to a single exponential decay (black line).

In the presence of ATP, most MutS:DNA complexes adopt a single bent conformation.

To ascertain how the conformation of the DNA changes upon introduction of ATP, we monitored the smFRET efficiency of our doubly labeled oligonucleotide (Figure 3.2A) in the presence of *Taq* MutS and saturating concentrations of ATP. Again, anti-correlated changes in the donor and acceptor fluorescence intensity time traces are observed when MutS and ATP are added. Dynamic changes in the FRET efficiency were observed (Figure 3.3A). In the majority of these events (70%), only two FRET states were observed: a low FRET state (Figure 3.3A, red arrows) and a high FRET (Figure 3.3A, cyan arrows), similar to the ADP FRET time traces. Histograms of the average FRET efficiency of each state show that the low FRET state (Figure 3.3B, red bars) exhibits a peak centered around 0.25. This peak is slightly broader than that observed in the presence of ADP but is still quite similar to the free DNA FRET distribution (Figure 3.3B, black dotted cityscape). The high FRET distribution observed in the presence of ATP (Figure 3.3C) was shifted to higher FRET with a peak centered around 0.4. This distribution is also broader than the high FRET distribution observed in the presence of ADP, perhaps due to MutS being inherently more dynamic in the presence of ATP.

For these transitions between low FRET and high FRET states in the presence of ATP, the TDP (Figure 3.3D) again reveals two dominant transitions: (1) a low FRET state transitioning to a high FRET state, and (2) a high FRET state transitioning to a low FRET state The breadth observed in the FRET efficiency histograms (Figures 3.3B and 3.3C) is also evident in the TDP, as the peaks representing the major transitions span a wider range of FRETs. Interestingly, the TDP reveals a systematic error in these data. Traces starting in a below average low FRET state (i.e. FRET < 0.25) transition to a state with a correspondingly below average high FRET state (i.e FRET < 0.4); likewise, traces starting in an above average low FRET state (i.e. FRET > 0.25) transition to a state with a correspondingly above average high FRET state (i.e. FRET > 0.4). This systematic error may explain the breadth of the distributions for the events involving a single bent conformation in the presence of ATP.

For this subset of transitions, the distribution of dwell time in the high FRET state (Figure 3.3E, cyan bars) fit well to a single exponential decay (Figure 3.3E, black line), yielding a characteristic lifetime of 4.2 ± 0.6 sec, within error of the lifetime observed in the presence of ADP and very close to that observed previously (2.7 sec) (Qiu et al., 2012, 2015).

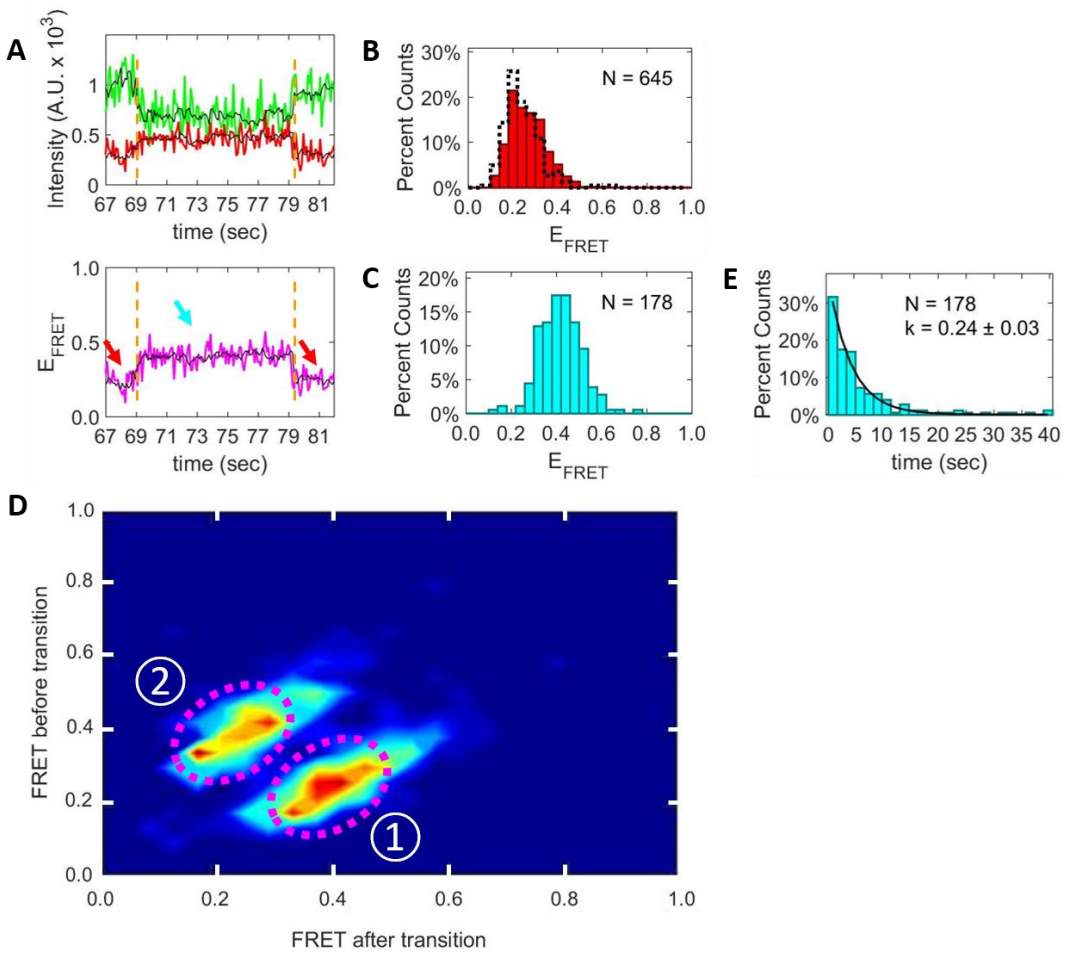

Figure 3.3: In the majority of *Taq* MutS:DNA complexes formed in the presence of ATP, the DNA adopts a single bent state.

A) Example donor (green) and acceptor (red) intensity time trace and its corresponding FRET time trace (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ATP. The black line represents the smoothed signal. The red and cyan arrows denote low FRET and high FRET states, respectively. **B)** The distributions of FRET values for the low FRET states (red bars) and for DNA in the absence of *Taq* MutS (black dotted cityscape). **C)** The distribution of FRET values for the high FRET states (cyan bars). **D)** Transition density plot depicting the frequency of transitions converting between the low and high FRET states. **E)** The distribution of dwell times for the high FRET states (cyan bars) fit to a single exponential decay (black line).

ATP induces a subset of MutS:DNA complexes to adopt multiple conformations.

In the remaining subset (30%) of the DNA bending events observed in the presence of *Taq* MutS and ATP, more than two FRET states are observed: a low FRET state, an intermediate FRET state, and a high FRET state (Figure 3.4A, red, blue, and yellow arrows, respectively). The distribution of FRET values for the low FRET state (Figure 3.4B, red bars) is again similar to the free DNA FRET (Figure 3.4B, black dotted cityscape), as both exhibit a peak centered around 0.25. The average FRET efficiency peaks for the intermediate and high FRET states (Figures 3.4C and 3.4D, respectively) are centered around 0.35 and 0.45, respectively. Such shifts in the FRET values predict DNA bend angles of ~45° and ~60°, respectively. These results are consistent with the previously observed range of *Taq* MutS:DNA complex conformations with distinct extents of DNA bending in the absence of nucleotides using T-bulge DNA (Alani et al., 2003; Obmolova et al., 2000; Derocco et al., 2014; Wang et al., 2003).

Examination of the TDP (Figure 3.4E) reveals four major transitions: (1) a low FRET state transitioning to an intermediate FRET state, (2) an intermediate FRET state transitioning to a high FRET state, (3) a high FRET state transitioning back to an intermediate FRET state, and (4) an intermediate FRET state returning to a low FRET state. Notably, the systematic error responsible for broadening the peaks in the TDP of the events with a single bent (Figure 3.3D) is again apparent in the TDP for the events transitioning through multiple bent states.

For this subset of DNA bending events in the presence of ATP, the total time spent in a FRET state other than the low FRET state (Figure 3.4A, asterisked braces) was recorded. The distribution of these overall dwell times (Figure 3.4G, gray bars) fit to a single exponential decay yielding a characteristic lifetime of 13 ± 2 sec. This lifetime is similar to the previously determined characteristic time (11.7 sec) that MutS spent at the T-bulge before forming the sliding clamp state. (Qiu et al., 2012, 2015).

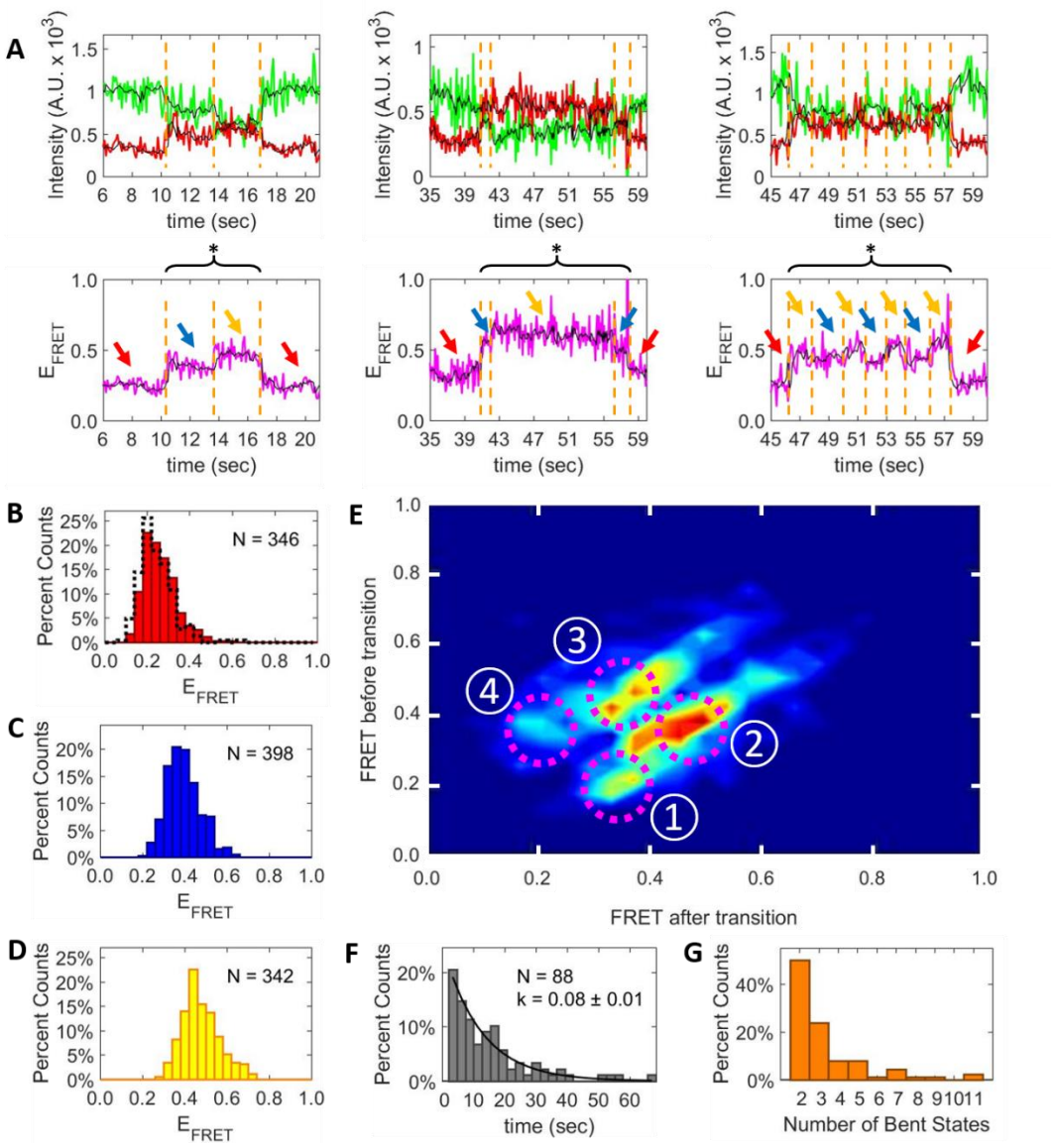

Figure 3.4: In a subset of *Taq* MutS:DNA complexes formed in the presence of ATP, switching between multiple bent states is observed.

A) Example donor (green) and acceptor (red) intensity time trace and its corresponding FRET time trace (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ATP. The black line represents the smoothed signal. The red, blue and yellow arrows denote low, intermediate, and high FRET states, respectively, between the detected FRET transitions (dotted gold lines). **B)** The distribution of the number of FRET states per bending event. **C)** The distributions of FRET values for the low FRET states (red bars) and for DNA in the absence of *Taq* MutS (black dotted cityscape). **D)** The distribution of FRET values for the intermediate FRET states (blue bars). **E)** The distribution of FRET values for the high FRET states (yellow bars). **F)** Transition density plot depicting the frequency of transitions between the low, intermediate and high FRET states (dotted magenta circles). (1) low FRET to intermediate FRET; (2) intermediate FRET to high FRET; (3) high FRET to intermediate FRET; and (4) intermediate FRET to low FRET. **G)** The distribution of dwell times (gray bars) for the total time in a bent state (i.e. the time marked by the asterisked braces) fit to a single exponential decay (black line).

DNA bending by MutS:ATP follows a preferred pathway of transitions.

In the events with multiple bent states, the number of intermediate and high FRET states entered before returning to the low FRET state (Figure 3.4F) shows a preference for fewer transitions. Approximately 75% of the events with multiple bent states adopt only two or three bent FRET states before returning to the low FRET state (Figure 3.5A). Rarely, the FRET oscillates between the high and intermediate FRET states 4 or more times before returning to the low FRET state (Figure 3.4A, right column).

The events adopting only two or three bent FRET exhibit a three preferred transition paths: Path 1) low FRET → intermediate FRET → high FRET → low FRET (Figure 3.5A, left), Path 2) low FRET → high FRET → intermediate FRET → low FRET (Figure 3.5A, middle), and Path 3) low FRET → intermediate FRET → high FRET → intermediate FRET → low FRET (Figure 3.5A, right). As depicted in the simplified schematics (Figure 3.5B), Paths 1 and 2 may represent subsets of Path 3 where one conformation was not observed. Using this model, each state is assigned an extent of MutS-induced DNA bending based on the FRET value as follows: Low FRET states (Figure 3.5A, red arrows) are assumed to be linear DNA (Figure 3.5C, **D**). This assumption is supported by the FRET distribution of these states (Figure 3.5D, red bars **D**), which is centered around 0.25 similar to free DNA. The intermediate FRET states (Figure 3.5A, blue arrows) are attributed to slightly bent DNA (Figure 3.5C, $\mathbf{U_1}$ and $\mathbf{U_2}$) because their FRET distributions (Figure 3.5D, blue bars and green bars for $\mathbf{U_1}$ and $\mathbf{U_2}$ respectively) have shifted to higher FRETs and are centered around 0.35. Interestingly, the FRET associated with $\mathbf{U_2}$ appears to be significantly broader than that of $\mathbf{U_1}$. Finally, the high FRET state (Figure 3.5B, yellow arrows) represents more sharply bent DNA (Figure 3.5C, **B**), as suggested by the FRET distribution (Figure 3.5D, yellow bars **B**) shifting to even higher FRET values centered around 0.45. For the three internal states, $\mathbf{U_1}$, **B**, and $\mathbf{U_2}$, the distribution of dwell times fit well to a single exponential decay yielding characteristic time constants of 3.3 ± 1.2, 2.3 ± 0.3, and 1.3 ± 0.4 sec, respectively. Importantly, both FRET distributions (Figure 3.6).and the kinetic analyses (Figure 3.7) for each of these states appear to be independent of the transition path.

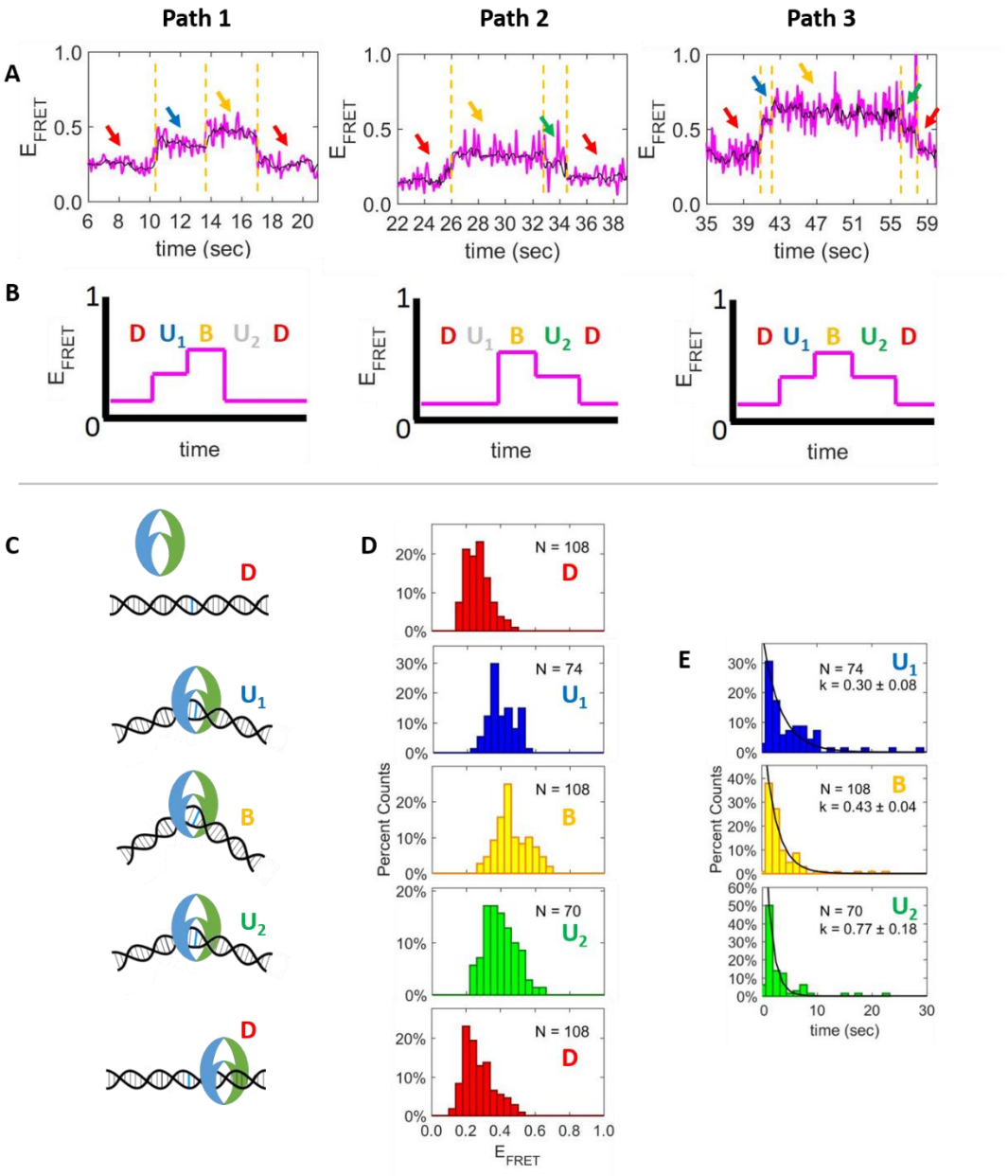

Figure 3.5: The DNA bending transitions for the multi-state bending events follows a D-U_1-B-U_2-D pattern.

A) Example FRET time traces (magenta) representing the three most common paths for multi-state traces. The black line represents the smoothed signal, and the arrows point to FRET states between the detected FRET transitions (dotted gold lines). Throughout the figure, the D, U_1, B, and U_2 states are color coded in red, blue, yellow, and green, respectively. **B)** Schematic representations of the most common transition paths. **C**) Models depicting the DNA bending conformations through the D-U_1-B-U_2-D pathway. **D)** The combined distributions of FRET values for each state. **E)** The combined distribution of dwell times for each state fit to a single exponential decay (black line).

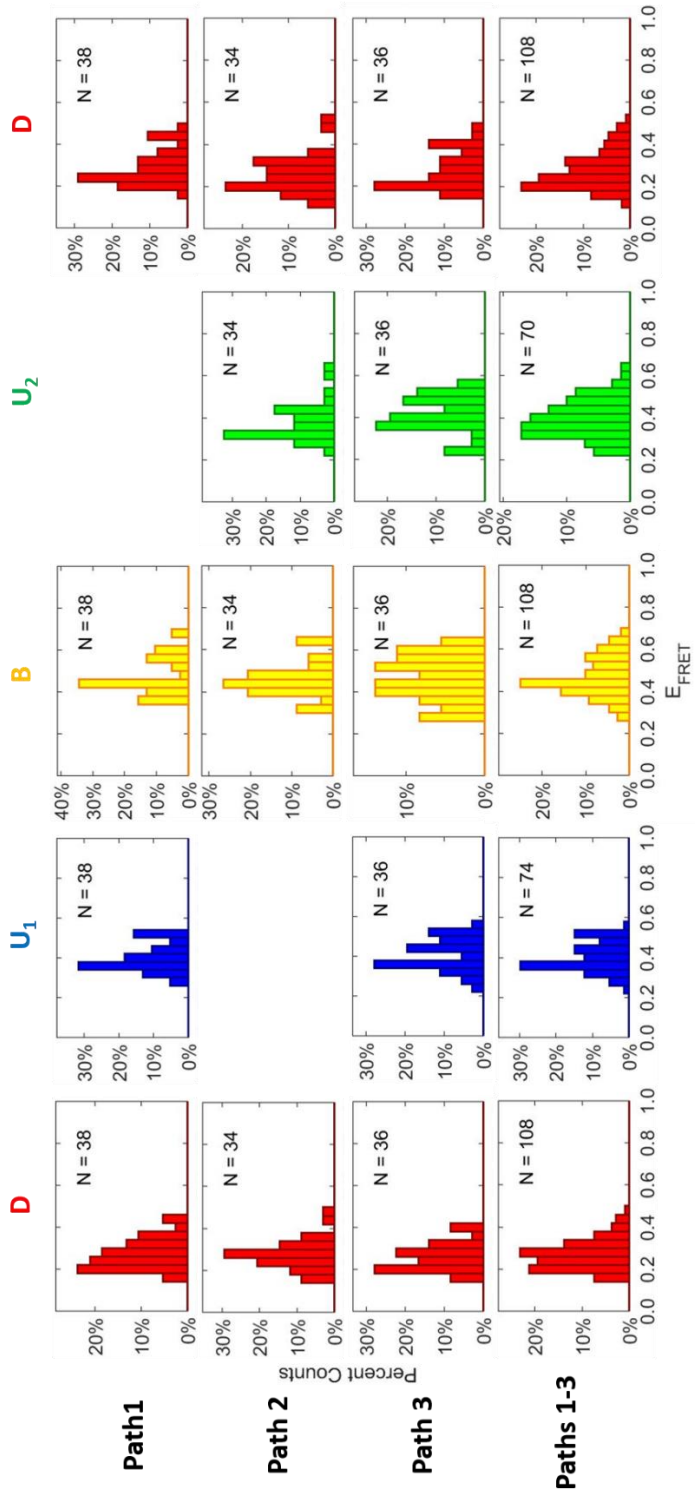

Figure 3.6: The distributions of FRET values for each state in the D-U_1-B-U_2-D pathway.

FRET distributions for each DNA bending state are shown separately for the Path 1 (*row 1*), Path 2 (*row 2*), and Path 3 (*row 3*). The combined distributions are also shown (*row 4*). Throughout the figure, the D, U_1, B, and U_2 states are color coded in red, blue, yellow, and green, respectively.

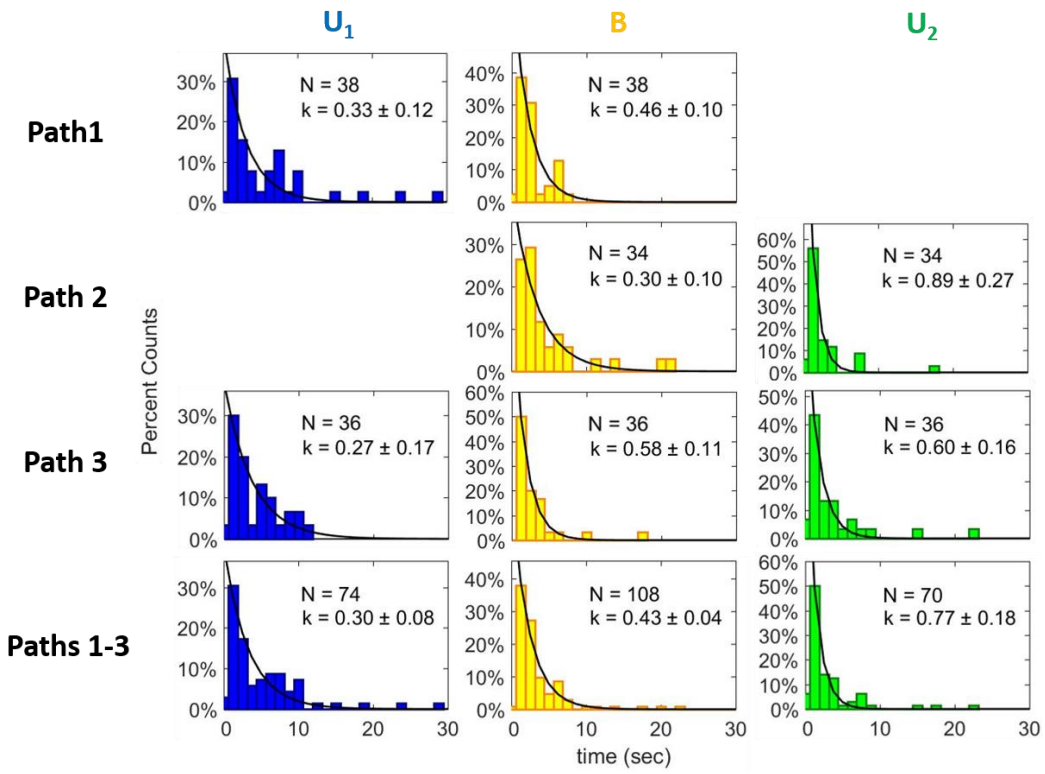

Figure 3.7: The dwell time distributions for the U_1, B, and U_2 states.

Dwell time distributions are shown separately for the Path 1 (*row 1*), Path 2 (*row 2*), and Path 3 (*row 3*). The combined distributions are also shown (*row 4*). Throughout the figure, the U_1, B, and U_2 states are color coded in blue, yellow, and green, respectively.

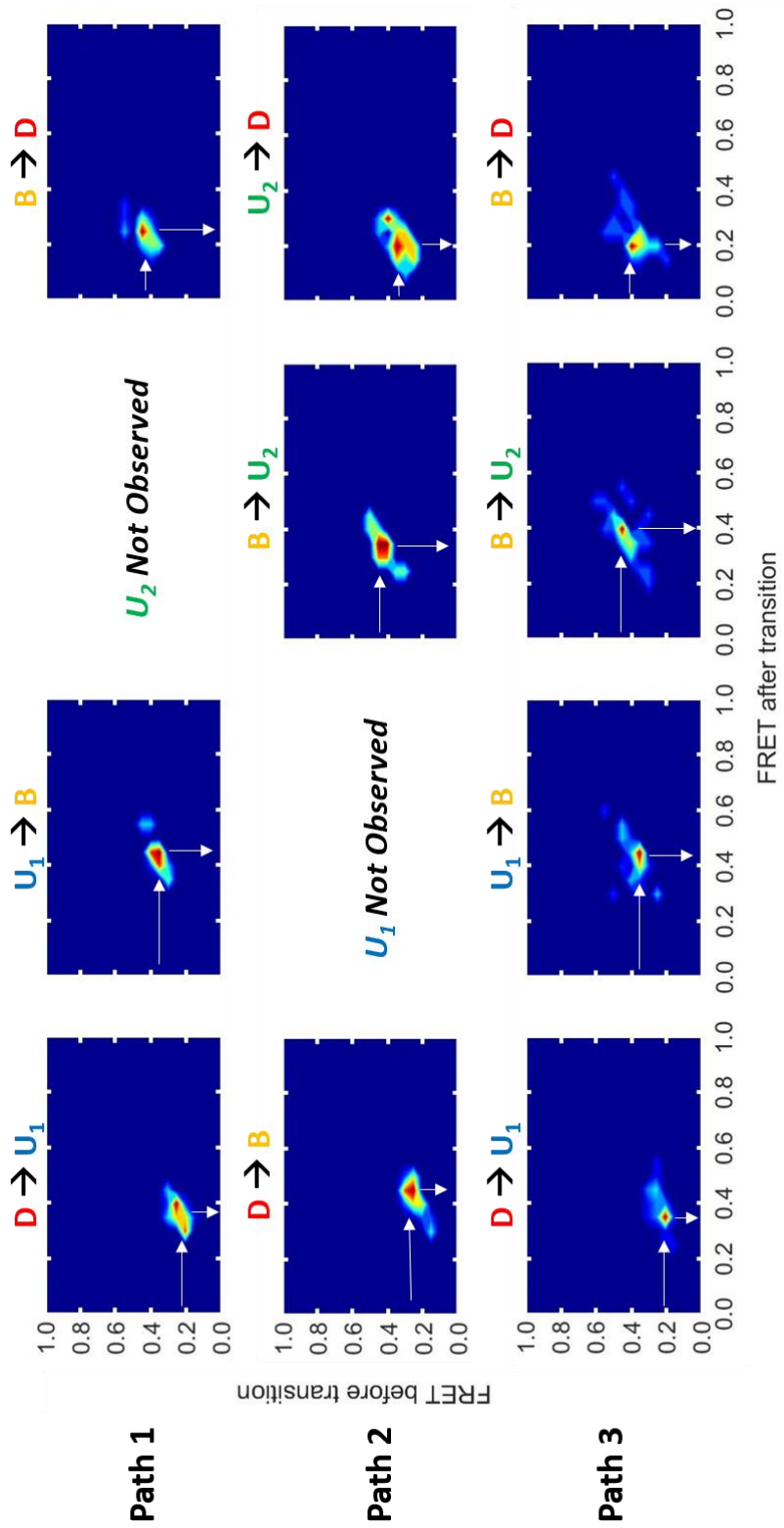

Figure 3.8: Transition density plots depicting each step of the D-U_1-B-U_2-D pathway.

Transition density plots depicting the preferred pathway (white arrows) of conversion between the DNA bending states. Each step is shown separately for Path 1 (*row 1*), Path 2 (*row 2*), and Path 3 (*row 3*) traces.

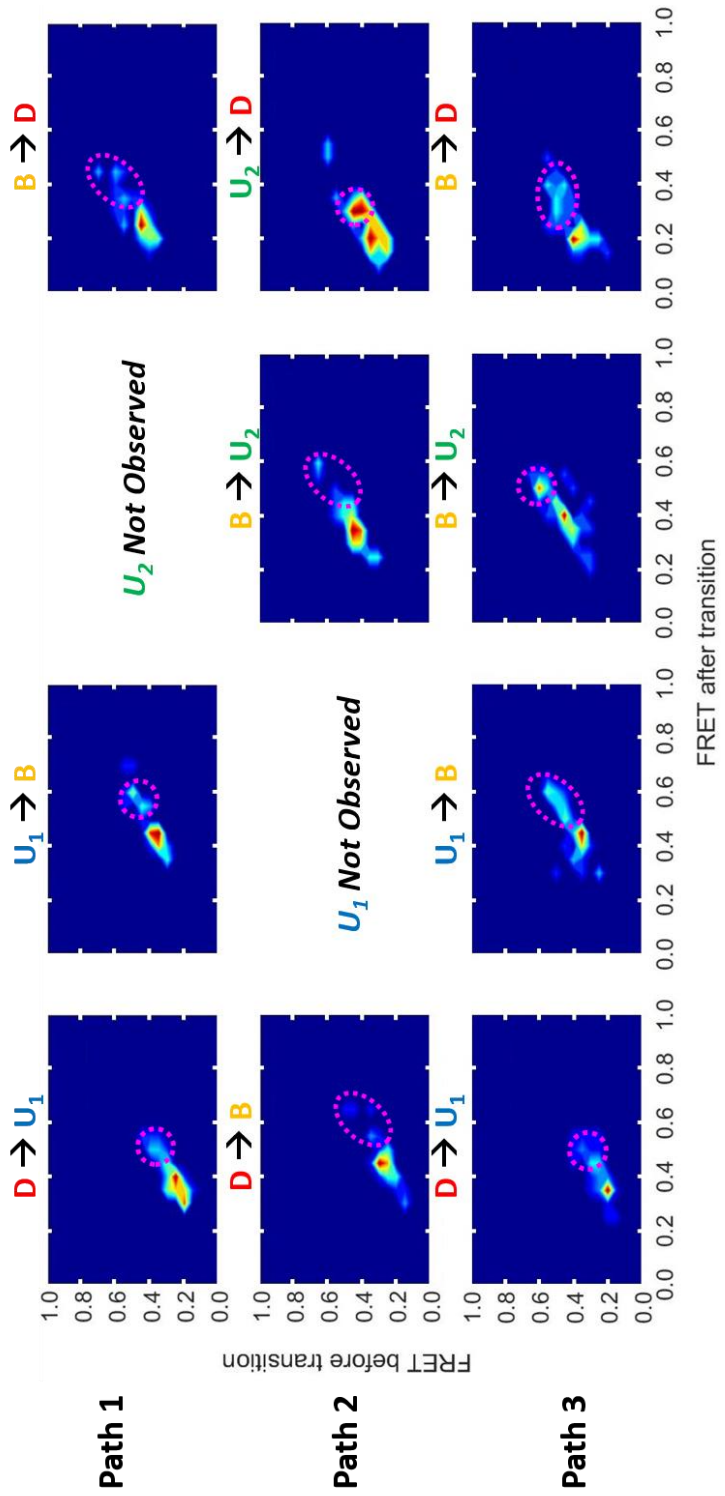

Figure 3.9: Transition density plots depicting the pathway of conversion between the DNA bending states for all molecules studied.

Each step is shown separately for the Path 1 (*row 1*), Path 2 (*row 2*), and Path 3 (*row 3*). Transitions occurring in molecules with a systematically high **D** FRET (omitted in Figure 3.4F) are circled in magenta.

To further analyze the transitions involved in these paths, the FRET efficiency before and after each transition are depicted in TDPs (Figure 3.8). Moving from left to right across each row, the transitions through each pathway become apparent. For Path 1, a clear transition from **D** (FRET = 0.2-0.3) to $\mathbf{U_1}$ (FRET = 0.35-0.4) is observed. Next, $\mathbf{U_1}$ transitions to **B** (FRET = 0.4-0.45). As $\mathbf{U_2}$ is not observed in this pathway, **B** then transitions directly to **D** (FRET = 0.2-0.3). For Path 2 (Figure 3.4F, middle row), **D** (FRET = 0.2-0.3) transitions directly to **B** (FRET = 0.4-0.45), skipping $\mathbf{U_1}$. **B** then transitions to $\mathbf{U_2}$ (FRET = 0.3-0.4) before finally arriving back at **D** (FRET 0.15-0.25). Finally, in Path 3 (Figure 3.4F, bottom row), all of the states are observed. **D** (FRET = 0.2) transitions to $\mathbf{U_1}$ (FRET = 0.35), which switches to **B** (FRET = 0.45). **B** then transitions to $\mathbf{U_2}$ (FRET = 0.35-0.4) before returning to **D** (FRET = 0.2). While slight variations in the FRET values for each state are apparent, the sequence of transitions is conserved. Notably, the systematic error that has been observed in all of the experiments in the presence of ATP complicates this analysis. Thus, for clarity, only transitions from FRET time traces whose **D** states had an average FRET efficiency below 0.3 were included in these TDPs. Importantly, though, the remaining transitions (i.e. those from traces whose state **D** had a FRET efficiency above 0.3) follow the same **D-**$\mathbf{U_1}$**-B-**$\mathbf{U_2}$**-D** transition pattern with every state exhibiting a systematically shifted higher FRET (Figure 3.9, magenta circles).

Discussion

During mismatch repair initiation, MutS is tasked with recognizing the error in the DNA. Upon error recognition, MutS undergoes a series of ATP-dependent conformational changes resulting in formation of a sliding clamp state and/or recruitment of MutL. These conformational changes are necessary for signaling downstream mismatch repair events. Using smFRET, much has been learned about nucleotide-dependent conformational changes involved in forming a MutS sliding clamp; however, little is known about the conformation of the DNA during this process. In this study, we used smFRET to monitor changes in DNA bending induced by MutS in the presence of saturating concentrations of ADP and ATP.

MutS and DNA conformational changes during error recognition and sliding clamp formation.

In two recent smFRET studies (Qiu et al., 2012, 2015), MutS-DNA interactions during sliding clamp formation and intramolecular MutS conformational dynamics were monitored. By placing the donor and acceptor fluorophore in different positions, we gain insight into the conformational properties of various components of the *Taq* MutS:DNA complex. Placing the donor fluorophore on the DNA binding domain of MutS and the acceptor fluorophore on the DNA near the mismatch (Figure 3.10A), the dynamics of MutS binding to the mismatch and the subsequent conformational changes leading to sliding clamp formation were tracked (Qiu et al., 2012, 2015). In these experiments, sliding clamp formation was characterized by a consistent pattern of transitions: MutS bound to the T-bulge in a high FRET state (FRET = 0.65) before transitioning to an intermediate state (FRET = 0.45). This intermediate then transitioned to the sliding clamp (FRET = 0) (Qiu et al., 2012, 2015).

With both the donor and acceptor fluorophores on the DNA binding domains of MutS (Figure 3.10B) intramolecular conformational changes were monitored (Qiu et al., 2012). These experiments revealed that, in the absence of DNA, *Taq* MutS exists mainly in two conformations: a low FRET "open" structure (FRET = 0.2) a high FRET "closed" structure (FRET = 0.9). Upon binding to DNA containing an error, MutS favors the high FRET "closed" conformation (FRET = 0.9). In the presence of ATP, a subset of MutS dimers transitioned to an "open" conformation following a consistent pathway: MutS bound to the error in the high FRET "closed" conformation (FRET = 0.9) transitioned to a "partially open" state (FRET = 0.65) before finally adopting the low FRET "open" state (FRET = 0.2). Not only were these transitions ATP- and mismatch-dependent, but they also correlated kinetically with the protein-to-DNA FRET experiments. Thus, these intramolecular conformational changes were attributed to MutS sliding clamp formation (Qiu et al., 2012).

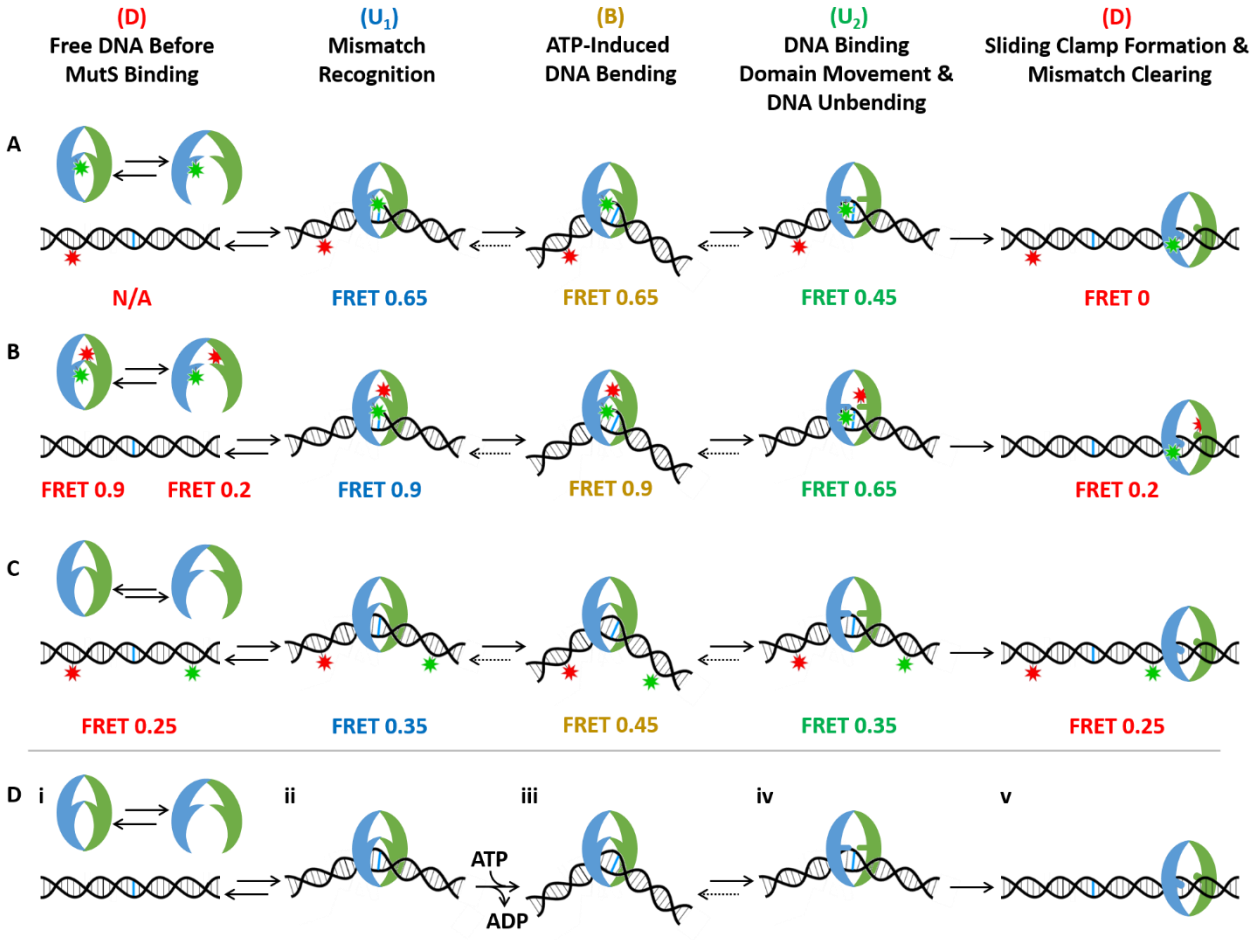

Figure 3.10: A model of sliding clamp formation using the results of three smFRET experimental designs.

Molecular states identified by smFRET experiments where **A)** The donor fluorophore is on the DNA binding domain of one MutS monomer, and the acceptor fluorophore is on the DNA near the mismatch, **B)** the donor and acceptor fluorophores are on the DNA binding domains of each MutS monomer, or **C)** the donor and acceptor fluorophores are on the DNA flanking the error. Average FRET values for each molecular state identified by each experiment are given below each cartoon. **D)** A model for sliding clamp formation for the *Taq* MutS:DNA complex: **(i)** Prior to binding DNA, MutS exists in a conformational equilibrium. DNA not bound by MutS remains unbent. **(ii)** MutS binds and bends the DNA at the error. MutS adopts a conformation where the two domains I are near each other. **(iii)** The DNA is briefly bent at a sharper angle while the MutS conformation changes little. **(iv)** The DNA unbends slightly as MutS undergoes a conformational change that increases the distance between the domains I. **(v)** The separation between the domains I increases again, forming the sliding clamp. The DNA near the error returns to its original unbent form.

While these approaches have provided much insight into the conformation changes experience by the MutS:DNA complexes, neither technique has reported on the DNA bending conformation through the sliding clamp formation process. Here, we track the extent of DNA bending by MutS in the presence of ADP and ATP using DNA labeled with both the donor and acceptor fluorophores (Figure 3.10C). Our results reveal at least three distinct FRET states: 1) a low FRET state (FRET = 0.25, **D**) that corresponds to free DNA that is not bound by MutS and is therefore straight, 2) an intermediate FRET state (FRET = 0.35, $\mathbf{U_1}$ or $\mathbf{U_2}$) representing DNA that has been slightly bent or unbent by MutS, and 3) a high FRET state (FRET = 0.45, **B**) attributed to a state with a greater extent of MutS-induced DNA bending. In a subset of MutS-induced DNA bending events (30%) in the presence of ATP, a dominant transition sequence through these DNA bending states is apparent: **D** (low FRET = 0.25) → $\mathbf{U_1}$ (intermediate FRET = 0.35) → **B** (high FRET = 0.45) → $\mathbf{U_2}$ (intermediate FRET = 0.35) → **D** (low FRET 0.45). In some of these events, the $\mathbf{U_1}$ or $\mathbf{U_2}$ states are not detected, which may be due to the relatively short lifetime of these states.

In the absence of nucleotides, MutS-induced DNA bending states similar to those observed in the presence of ADP and ATP have been observed, both in control experiments (Figures 3.11A and 3.11B) and in previous studies using a different T-bulge DNA substrate (Derocco et al., 2014); however, these bent states are significantly longer lived rarely transition to other states during the observation period (100 sec). Most likely this lack of dynamics can be explained by the stability of the *Taq* MutS on T-bulge DNA in the absence of ADP or ATP. As expected, high FRET states as a result of DNA bending are not observed on homoduplex DNA labeled with a FRET pair of fluorophores (Figure 3.11C).

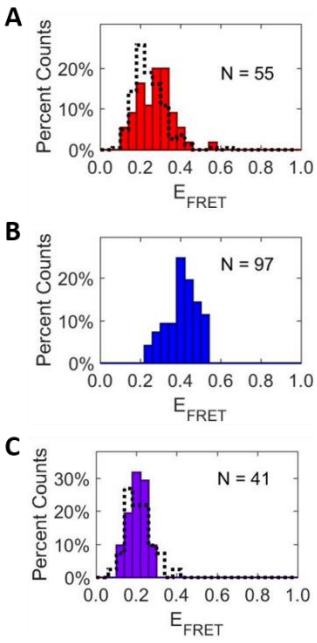

Figure 3.11: Control Experiments.

A) The distribution of FRET values for the low FRET states for experiments conducted with T-bulge DNA and 10 nM *Taq* MutS (red bars) and the distribution of FRET values for T-bulge DNA in the absence of *Taq* MutS (black dotted cityscape). **B)** The distribution of FRET values for the high FRET states for experiments conducted with T-bulge DNA and 10 nM *Taq* MutS (blue bars). **C)** The distribution of FRET values for the high FRET states for experiments conducted with homoduplex DNA in the presence (purple bars) and absence (dotted black cityscape) of 10 nM *Taq* MutS (purple bars).

The molecular states identified by smFRET can be unified by their kinetics and transition sequence.

In each of the experiments discussed above, changing the position of the donor and acceptor fluorophores provides unique insight into the nucleotide-dependent conformational changes within the MutS-DNA complexes. The distinct molecular states identified by each of these independent experimental designs can be correlated using their kinetics and by the transition sequence. Importantly, the characteristic lifetime a given MutS:DNA complex conformation should be approximately equal regardless of the fluorophore locations, and both the number of events as well as the preferred pathway of transitions should be preserved.

The protein-to-DNA FRET experiments (Figure 3.10A) directly visualized formation of the sliding clamp and observed three FRET states (high FRET → intermediate FRET → zero FRET). The dwell time distribution of the high FRET state required a two-step kinetic model to fit the data. Thus, four characteristic lifetimes were identified in these experiments: 2.2 and 0.9 sec (FRET = 0.65), 1.8 sec (FRET = 0.45), and 2.2 sec (FRET = 0) (Qiu et al., 2015). In the intramolecular FRET experiments (Figure 3.10B), sliding clamp formation was characterized by three FRET states (high FRET = 0.9 → intermediate FRET = 0.65 → low FRET = 0.2). The characteristic lifetimes of these states were found to be 4.7, 1.3, and 2.4 sec, respectively (Qiu et al., 2012).

With no label on MutS in the DNA-to-DNA FRET experiments, there is no way to directly visualize whether or not MutS is bound to the DNA during the experiment; however, several pieces of evidence allow us to conclude that the dynamic changes in FRET are due to MutS-induced changes in DNA bending. First, these events are not observed in the absence of MutS. Second, the characteristic lifetimes identified of the high FRET states in the presence of ADP and for the single-state events in the presence of ATP are very similar to those observed in the protein-to-DNA experiments at the same conditions. Third, the population of ADP-like events in the presence of ATP are similar in both the protein-to-DNA and DNA-to-DNA experiments (80% and 70%, respectively). Finally, the events known to form the sliding clamp in the protein-to-DNA experiments (20% of events in the presence of ATP) had a total characteristic lifetime at the mismatch of around 11.7 sec, which is strikingly similar to that of the subset of DNA bending events exhibiting multiple bent states (30% of events in the presence of ATP), 12.5 sec. Taken together, these data suggest that this subset of DNA bending events represents the population of MutS:DNA complexes that are competent to form sliding clamps in the presence of ATP.

The dominant pathway of DNA bending transitions identified in the events leading to sliding clamp formation was: **D** (low FRET, free DNA) → $\mathbf{U_1}$ (intermediate FRET, slight bending) → **B** (high FRET, greater extent of bending) → $\mathbf{U_2}$ (intermediate FRET, slight bending) → **D** (low FRET, free DNA). Not only does the number of states in this pathway correlate with the number of states in the existing model for

MutS sliding clamp formation (Figure 3.1), but the kinetics of these states also correlate (Table 3.1). Thus, the states identified in each of the smFRET experiments are combined in a unified model for sliding clamp formation by *Taq* MutS (Figure 3.10D).

Table 3.1: Kinetic parameters from the smFRET experiments.

Donor Location	Acceptor Location	Free DNA (D)	State 1 (U_1)	State 2 (B)	State 3 (U_2)	Sliding Clamp (D)
MutS	MutS	-	4.7 sec (total)		1.3 sec	2.4 sec
MutS	DNA	-	2.2 sec & 0.9 sec		1.8 sec	2.2 sec
DNA	DNA	-	3.3 sec	2.3 sec	1.3 sec	-

Model of sliding clamp formation

Single-molecule FRET is a powerful tool to resolve asynchronous dynamic changes in complicated samples at steady-state conditions. The sequence of FRET states observed during error recognition and sliding clamp formation by *Taq* MutS, both in this work and previous single-molecule FRET studies, allowed us to develop the following model for the MutS:DNA complexes during mismatch repair initiation (Figure 3.10D):

Based on the intrinsic ATPase activity MutS and the physiologically high concentrations of ADP and ATP, free MutS that is not bound to a mismatch most likely exists in an ADP:ADP or ATP:ADP liganded state. In either case, the MutS-to-MutS FRET experiments suggest that these MutS molecules will exist in a conformational equilibrium between an “open” and “closed” state (Figure 3.11D, i). The “open” state may facilitate loading onto the DNA, while the “closed” state may stabilize the MutS:DNA interaction. ATP:ATP liganded MutS molecules, if they exist, are expected to exist exclusively in a “closed” conformation, and are therefore not expected to bind DNA.

Upon encountering the error (Figure 3.8E, ii), MutS adopts a “closed” conformation and the DNA is bent slightly, in an initial recognition complex. The fate of this state now depends on the nucleotide ligation status of the involved MutS. The ADP:ADP liganded molecules would directly dissociate from the mismatch, as would the majority of the ATP:ADP molecules. A subset of the ATP:ADP molecules,

however, will successfully transition to the penultimate recognition state (Figure 3.8E, iii). In this state, the DNA is significantly bent, but conformations of the DNA binding domains of MutS change very little. Based on the kinetics of the transition from the initial recognition state to the penultimate recognition state could be the result of ADP release followed by rapid ATP binding. Thus, the penultimate recognition state is likely ATP:ATP liganded, which could serve to stabilize the MutS-DNA interactions by locking MutS into a "closed" conformation.

Next MutS transitions to the ultimate recognition complex (Figure 3.8E, iv) via two major conformational changes: 1) the DNA binding domains of MutS "partially open", and 2) the DNA unbends slightly. This transition to the ultimate recognition state includes the hallmark "bent-to-unbent" change in the DNA conformation that has been observed in several studies (Derocco et al., 2014; Sass et al., 2010; Wang et al., 2003). Notably, this change in DNA bending is predictive for repair efficiency (Derocco et al., 2014; Tessmer et al., 2008). The nucleotide ligation status of MutS in the ultimate recognition complex is unclear, as ATP hydrolysis may occur in one (or both) ATPase sites. In the final transition, MutS becomes a sliding clamp (Figure 3.8E, v). In the sliding clamp state, the DNA binding domains are fully "open", and MutS is free to move away from the mismatch. Though it is unclear whether MutS sliding clamps bend DNA, the DNA near the mismatch that is left behind by MutS returns to its original straight conformation.

Biological Significance

Using smFRET, we identified a preferred pathway of *Taq* MutS-induced DNA bending state in the presence of ATP. These results have refined our understanding of MMR initiation by *Taq* MutS by providing evidence of dynamic changes in DNA bending during sliding clamp formation. Interestingly, the "bent-to-unbent" transition that is known to correlate with repair efficiency (observed in previous studies in the absence of nucleotide) was identified here (Figure 3.10D, **B** to $\mathbf{U_2}$) as part of the conformational changes expected during sliding clamp formation. Thus, the conformational dynamics observed in the absence of ATP seems to be predictive of those observed with ATP.

Materials and Methods

Protein and DNA substrates

C42A/M88C *Taq* MutS was provided graciously provided by Dr. Keith Weninger. The following fluorescently labeled DNA oligonucleotides (IDT) were used to prepare the 68 bp T-bulge DNA substrate: b-CTC TAG AGG ATC CGC TGA GGC CAG CTG AGG CCT GGC TGA GGA TTG CTG A(T*)G AAT TCA CTG GCC GTC G; dig-CGA CGG CCA GTG AAT TCA TCA GCA ATC **T**CT CAG CCA G/iCy5/GC CTC AGC TGG CCT CAG CGG ATC CTC TAG AG. The "b" represents a 5′ biotinylation, and "dig" represent a 5′ digoxigenin modification added to facilitate blocking of the free end. The bold and underlined thymine is the single base insertion. TAMRA was linked to the thymine base denoted (T*), and an internal Cy5 was incorporated at the position marked /iCy5/. The sequence of this oligonucleotide was based on the DNA substrates used in previous studies, with the only change being substituting a thymine (T*) into the sequence to allow labeling (Qiu et al., 2012, 2015). The dye separation (19 bp) was selected based on previous studies (Derocco et al., 2014; Sass et al., 2010).

Single-molecule FRET experiments

Quartz slides were plasma cleaned and used to construct home-built flow cells. The slides were then pre-functionalized by sequentially incubating the flow cells with 1 mg/ml biotinylated BSA and 0.1 mg/ml streptavidin. Biotinylated fluorescently-DNA oligonucleotides were then immobilized onto the surface of the quartz slide (as depicted in Figure 3.2A). The samples were then imaged using a through-prism total internal reflection single molecule fluorescence microscope. Donor and acceptor excitation was achieved using 532 nm and 638 nm lasers respectively. The fluorophore emission is collected through a 60X water immersion, 1.2 N.A. objective, and the image is split by a DualView optical splitter with a 645 nm dichroic mirror. The donor and acceptor signals then pass through optical filters (i.e. a 585/70 bandpass filter for TAMRA, and a 655 longpass filter for Cy5) before detection by an emCCD camera. To observe

changes in DNA bending over time, movies of approximately 1000 frames are collected using the following excitation sequence: 1) Brief excitation of the acceptor dye (~ 1 sec) to locate DNA molecules; 2) Excitation of the donor dye (~2 min) to monitor changes in FRET; 3) Brief excitation of the acceptor dye (~5 sec) to reveal whether the acceptor has photobleached. All experiments were performed at room temperature in 50 mM Tris, 100 mM sodium acetate, 5 mM magnesium chloride, 2% glucose (w/v), pH 7.8, in the presence of an oxygen scavenging and triplet state quenching system of 100 U/ml glucose oxidase, 1000 U/ml catalase, 0.05 mg/ml cyclooctatetraene and 143 mM 2-mercaptoethanol.

Data analysis

Data analysis was carried out using the approach described in Chapter 2. Only those FRET time traces exhibiting changes in FRET were analyzed; traces whose FRET was constant over the entire observation window were discarded. Only traces with intensities consistent with one MutS were analyzed. FRET traces showing evidence of protein:dye interactions (i.e. changes in fluorescence intensity of either the donor or acceptor dye that are not anti-correlated) were also discarded from analysis.

CHAPTER 4: PRELIMINARY STUDIES MONITORING DNA BENDING BY MUTS

Introduction

Mismatch repair is initiated by MutS, which recognizes errors left behind by DNA replication. These errors can be base-base mismatches and/or insertion/deletion loop errors (Iyer et al., 2006; Kunkel & Erie, 2005). Error recognition by MutS occurs upon formation of specific interactions with the mispaired base(s) (Obmolova et al., 2000). Subsequently, MutS undergoes ATP-dependent conformational changes to form a mobile "sliding clamp" state and/or to recruit MutL, a downstream mismatch repair protein (Qiu et al., 2012, 2015). Recent studies using single-molecule fluorescence resonance energy transfer (smFRET) successfully characterized the conformational and kinetic properties of MutS during sliding clamp formation and MutL recruitment on a DNA substrate containing a single thymine insertion (T-bulge) (Qiu et al., 2015). To assess the stoichiometries of these complexes, fluorescently labelled MutS and MutL were incubated with DNA containing an error and then cross-linked. The resulting complexes were then imaged, and the number of each protein was counted using stochastic photobleaching. Ternary MutS:MutL:DNA complexes studied this way were found to be widely varied, but they typically consisted of one MutS homodimer and two MutL homodimers (Qiu et al., 2015). While the mechanism of error recognition and mismatch repair initiation appears to be conserved, repair efficiency varies among the different DNA errors.

Dynamic DNA bending by MutS is a well-established and well-characterized phenomenon (Derocco et al., 2014; Sass et al., 2010). These DNA bending conformational dynamics, particularly a transition from "bent" to "unbent" DNA, appear to correlate with repair efficiency for the different types

of errors. While the evidence suggests that DNA bending plays a key role in repair, DNA bending states adopted by the mismatch repair initiation complex have not been studied in the presence of MutL. In addition, differences between the DNA bending dynamics on DNA substrates containing different types of DNA errors have not been studied in the presence of ATP. Finally, changes in the DNA bending states caused by blocking the free end of the DNA, thereby preventing dissociation of the MutS sliding clamp from the end of the DNA, have also gone uncharacterized.

Here, preliminary studies to address these gaps in our understanding mismatch repair initiation are presented. We use smFRET to monitor changes in DNA bending induced by *Thermus aquaticus* (*Taq*) MutS in the presence of saturating concentration of ADP or ATP on DNA substrates containing a GT mismatch. Subsequently, we study DNA bending on end-blocked substrates containing a T-bulge error. Finally, we monitored bending of T-bulge DNA in the presence of MutL. The results of these experiments are then compared to the results presented in Chapter 3 (T-bulge DNA with free ends and without MutL). Key findings of these studies include: 1) MutS bends GT DNA to a wider range of conformations, complicating the FRET-TACKLE analysis, 2) the kinetics of the DNA bending events observed on GT DNA do not correlate with previous FRET studies, complicating interpretation of these results, 3) MutS sliding clamps on end-blocked DNA substrates may return to the site of the mismatch, though these events are very rare, and 4) preliminary evidence of an equilibrium between two DNA bending states in the presence of MutL exists, though protein:dye interactions complicate their interpretation.

Results and Discussion

MutS bends GT DNA to a broad range of conformations in the presence of ADP and ATP.

To study the nucleotide dependence of MutS-induced DNA bending on oligonucleotides containing a GT mismatch using smFRET, we designed a 68 bp oligonucleotide that was doubly-labeled with a FRET pair of dyes, TAMRA (donor) and Cy5 (acceptor). These dyes were separated by 19 bp flanking GT

mismatch error near the midpoint of the oligonucleotide (Figure 4.1A). This oligonucleotide was also biotinylated on one end so that it could be immobilized via an interaction with streptavidin-biotinylated BSA on the surface of a quartz slide. The cysteine mutant C42A/M88C *Thermus aquaticus* MutS with wild type ATPase and DNA binding activities was used in this study for direct comparison to existing smFRET data (Qiu et al., 2012, 2015).

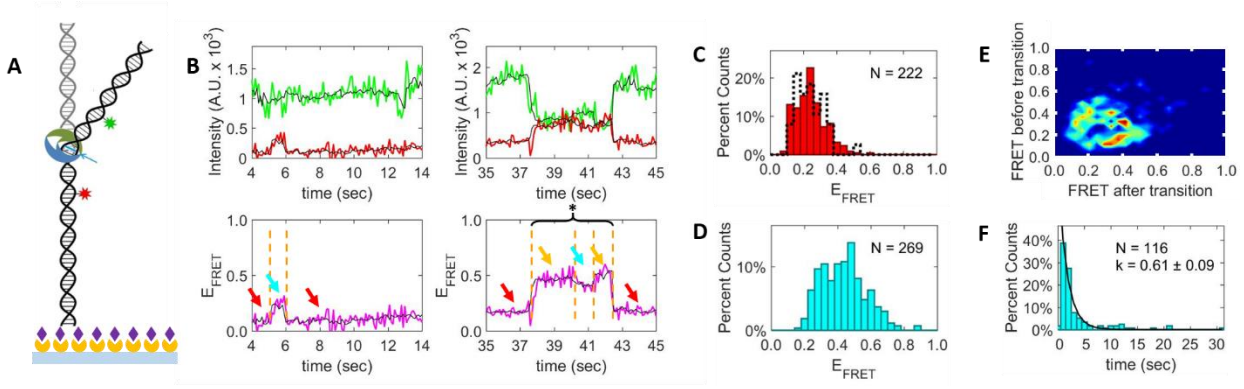

Figure 4.1: In the presence of ADP, GT DNA bound by *Taq* MutS adopts multiple bent states.

A) Schematic of a surface immobilized 70mer DNA molecule bound by *Taq* MutS at the location of the GT mismatch (blue arrow). **B)** Example donor (green) and acceptor (red) intensity time traces and their corresponding FRET time traces (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ADP. The black line represents the smoothed signal. The red, cyan, and yellow arrows denote low, intermediate, and high FRET states, respectively. **C)** The distributions of FRET values for the low FRET states (red bars) and for DNA in the absence of *Taq* MutS (black dotted cityscape). **D)** The distribution of FRET values for the high FRET states (cyan bars). **E)** Transition density plot depicting the frequency of transitions converting between the low and high FRET states. **F)** The distribution of dwell times for the high FRET states (cyan bars) fit to a single exponential decay (black line). For multi-step events, the dwell time of the full event (asterisked brace) was used.

We first determined the DNA bending properties of *Taq* MutS in the presence of saturating concentrations of ADP. We measured the smFRET efficiency from the donor to the acceptor dyes on either side of a T-bulge (Figure 4.1A). The FRET efficiency associated with free DNA was broadly distributed near 0.3 (Figure 4.1C, black dotted cityscape). Introduction of 10 nM MutS and 2mM ADP produced donor and acceptor time traces (Figure 4.1B, green and red lines, respectively) with anti-correlated dynamics. These fluorescence intensity changes resulted in changes in FRET efficiency (Figure 4.1B, magenta lines).

Most of these FRET time traces showed transitions between only two states: a low FRET state consistent with the free DNA (Figure 4.1B, red arrows) and a high FRET state (Figure 4.1B, cyan arrows). In contrast to the T-bulge studies, intermediate FRET states were also observed in the presence of ADP (Figure 4.1B, yellow arrow). The distribution of FRET values for the low FRET states (Figure 4.1C, red bars) was near 0.3, which is loosely similar to that of the free DNA. However, the distribution of FRET values for the intermediate and high FRET states (Figure 4.1D) is very broad and appears to include several distinct FRET states. This wide range of FRET states is also visible in the transition density plot (TDP) which depicts the distribution of transitions observed by the MutS:GT DNA complexes in the presence of ADP (Figure 4.1E, warmer colors represent more frequent transitions). Many states with distinct FRET are observed, and the transitions between these states are widely varied, which is similar to previous DNA bending studies in the absence of nucleotide (Sass et al., 2010).

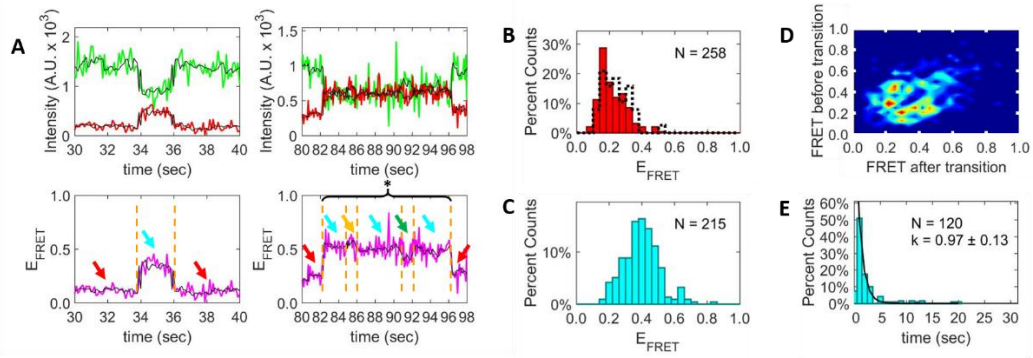

Figure 4.2: In the presence of ATP, GT DNA bound by *Taq* MutS also adopts multiple bent states.

A) Example donor (green) and acceptor (red) intensity time traces and their corresponding FRET time traces (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ATP. The black line represents the smoothed signal. The red arrows denote the low FRET states, and the cyan, yellow, and green arrows denote different high FRET states. **B)** The distributions of FRET values for the low FRET states (red bars) and for DNA in the absence of *Taq* MutS (black dotted cityscape). **C)** The distribution of FRET values for the high FRET states (cyan bars). **D)** Transition density plot depicting the frequency of transitions converting between the low and high FRET states. **E)** The distribution of dwell times for the high FRET states (cyan bars) fit to a single exponential decay (black line). For multi-step events, the dwell time of the full event (asterisked brace) was used.

In the presence of 10 nM MutS and 2 mM ATP, donor and acceptor time traces with anti-correlated transitions were again observed (Figure 4.2A, green and red lines, respectively). The resulting FRET time traces (Figure 4.2A, magenta lines) revealed two classes of transitions. Most traces showed traces from a low FRET (Figure 4.2A, red arrows) state to a higher FRET state (Figure 4.2A, cyan arrows), while a few rare bending events transitioned through multiple states of relatively high FRET (Figure 4.2A, cyan, yellow, and green arrows). The distribution of low FRET states (Figure 4.2B, red bars) was consistent with the free DNA distribution (Figure 4.2B, black dotted cityscape). The high FRET state states were broadly distributed across higher FRET values (Figure 4.2C), though a bending state with a FRET of ~0.4 appears to dominate. Analysis of the TDP in the presence of ATP reveals an increase in transitions between many distinct FRET states relative to those observed with ADP. This increase in dynamics upon introduction of ATP could be caused by the ATP-dependent conformational changes previously observed in MutS on DNA containing a mismatch (Qiu et al., 2012, 2015).

Kinetic analysis of the observed increases in FRET yields lifetimes that are significantly shorter than those seen in MutS DNA studies monitoring sliding clamp formation. Here, overall time the DNA is bent in the presence of ADP (Figure 4.1F) or ATP (Figure 4.2E) was fit to a single exponential decay, resulting in characteristic lifetimes of 1.6 ± 0.3 sec and 1.0 ± 0.2 sec, respectively. These results are in stark contrast to those observed in previous single-molecule FRET studies, which found that MutS was bound to GT mismatch DNA with lifetimes on the order of 30-60 seconds (Qiu et al., 2012, 2015; Sass et al., 2010). This discrepancy makes interpretation of the observed changes in FRET on GT DNA difficult.

Three models may be able to unify these seemingly conflicting results: 1) Upon recognizing the mismatch, MutS bends the DNA for approximately one second before transitioning to an unbent state. This unbent state then persists for tens of seconds; 2) MutS initially binds a GT mismatch in an unbent state and then transiently bends the DNA for about one second before returning to an unbent state. In total, these transitions would take around 30-60 seconds. This model is unlikely given that MutS is expected to bend DNA substrates upon mismatch recognition; or 3) MutS binds GT mismatch DNA in an unbent state and

persists this way for 30-60 seconds. Then, just before forming the sliding clamp, MutS bends the DNA for about one second before sliding away. To clarify which of these models is most likely, three color FRET experiments using tagged MutS could be done, just as has been done previously with yeast MutSα (DeRocco, Anderson, Piehler, Erie, & Weninger, 2010). Using this approach, the DNA bending event could be placed within the MutS binding event.

MutS sliding clamps may return to the T-bulge on end-blocked DNA.

To study MutS-induced DNA bending on end-blocked oligonucleotides containing a T-bulge using smFRET, we used the 68 bp oligonucleotide described in Chapter 2. Anti-digoxigenin antibodies (20 µg/ml, incubated for 10 min) were used to block the ends, thereby prevent MutS sliding clamps from dissociated directly from the free end of the DNA (Figure 4.3A). This oligonucleotide was also biotinylated on one end so that it could be immobilized via an interaction with streptavidin-biotinylated BSA on the surface of a quartz slide. The cysteine mutant C42A/M88C *Thermus aquaticus* MutS with wild type ATPase and DNA binding activities was used in this study for direct comparison to existing smFRET data (Qiu et al., 2012, 2015).

MutS-induced DNA bending of end-blocked T-bulge oligonucleotides was monitored in the presence of 10 nM MutS and 2 mM ATP. At these concentrations of MutS and ATP, approximately 20-30% of the DNA bending events were observed to form sliding clamps in previous studies (Chapter 3, Qiu et al., 2012). Analysis of hundreds of traces revealed that, generally, the donor and acceptor intensity time traces exhibited non-correlated transitions, likely due to protein:dye interaction. These events could be due to loading of multiple MutS sliding clamps onto the DNA, as has been observed previously (Qiu et al., 2015). In a handful of traces (< 1%), the data were of sufficient quality to see the following pattern of events: a low FRET state (Figure 4.3B, red arrows) → a multi-state, higher FRET event (Figure 4.3B, cyan arrows) → a second low FRET state → a single-state, high FRET event (Figure 4.3B, gold arrows) → a

low FRET state. These states can be interpreted as follows: free DNA being bound by MutS (red arrow to cyan arrow) → the MutS:DNA complex undergoing conformational changes typical of sliding clamp formation (cyan to cyan arrows) → MutS sliding away from the T-bulge (cyan arrows to red arrow) → MutS sliding back to the T-bulge (red arrow to yellow arrow) → MutS dissociating or sliding away from the T-bulge site. Alternatively, the second event may be interpreted as loading of a second MutS at the T-bulge site. Extensive analysis of these traces would require collecting larger amounts of data in hopes of observing more of these rare events. In addition, the three-color experiment using tagged MutS could allow us to ascertain if multiple MutS loading events were taking place.

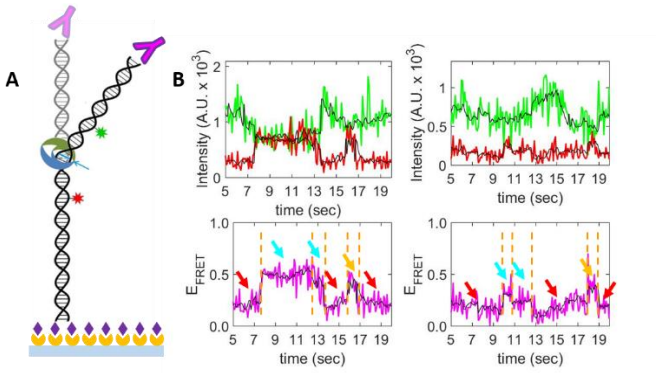

Figure 4.3: On end-blocked T-bulge DNA, *Taq* MutS may return to the error.

A) Schematic of a surface immobilized 70mer DNA molecule bound by *Taq* MutS at the location of the T-bulge (blue arrow). Note the free end of the DNA is blocked by an anti-digoxigenin antibody (magenta). **B)** Example donor (green) and acceptor (red) intensity time traces and their corresponding FRET time traces (magenta) for experiments conducted at 10 nM *Taq* MutS and 2 mM ATP. The black line represents the smoothed signal. The red, cyan, and yellow arrows denote the low FRET state, a multi-FRET state, and a subsequent bending event, respectively.

MutS:MutL complexes may adopt a rapid equilibrium between DNA bending states.

To study MutS:MutL-induced DNA bending on T-bulge using smFRET, we used the 68 bp oligonucleotide described in Chapter 2 (Figure 4.4A). This oligonucleotide was also biotinylated on one end so that it could be immobilized via an interaction with streptavidin-biotinylated BSA on the surface of a quartz slide. The cysteine mutant C42A/M88C *Thermus aquaticus* MutS with wild type ATPase and DNA

binding activities was used in this study for direct comparison to existing smFRET data (Qiu et al., 2012, 2015). Nickel-affinity chromatography was used to purify N-terminally his-tagged *Thermus aquaticus* MutL that has been used in previous studies (Qiu et al., 2015).

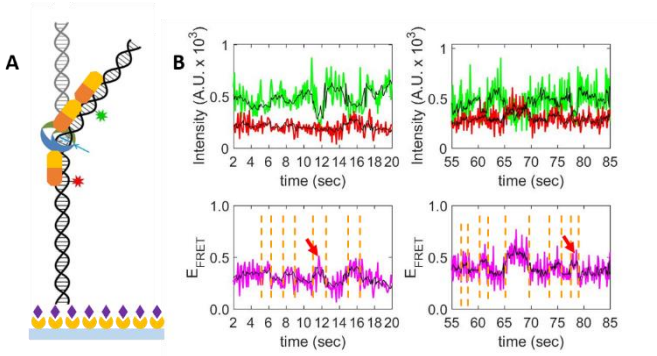

Figure 4.4: DNA bending by *Taq* MutS may be affected by MutL.

A) Schematic of a surface immobilized 70mer DNA molecule bound by *Taq* MutS and MutL at the location of the T-bulge (blue arrow). **B)** Example donor (green) and acceptor (red) intensity time traces and their corresponding FRET time traces (magenta) for experiments conducted at 10 nM *Taq* MutS, 100 nM *Taq* MutL and 2 mM ATP. The black line represents the smoothed signal. The red arrows point to "false-positive" apparent FRET changes likely due to protein:dye interactions.

DNA bending by the ternary complex of MutS:MutL:T-bulge DNA was studied at 10 nM MutS, 100 nM MutL, and 2 mM ATP. Previous studies at these concentrations revealed that MutS existed in a rapid equilibrium between two states while in a MutS:MutL complex that was trapped at the site of the error. Analysis of hundreds of traces revealed that, generally, the donor and acceptor intensity time traces exhibited non-correlated transitions, likely due to protein:dye interaction. This observation was not entirely unexpected, as the position of the fluorophores are within the expected footprint of a MutS:MutL complex formed at the T-bulge. Rarely (>1% of the traces), rapid oscillation between two (or more) FRET states was observed (Figure 4.4B). However, these observations suffer from several shortcomings: 1) The changes in donor and acceptor fluorescence intensity, and thereby, the changes in FRET efficiency, are very small; 2) Non-correlated changes in donor and acceptor intensity intermittently occur during these oscillations (Figure 4.4B, red arrows), and they often produce changes in FRET of approximately equal magnitude to

the identified oscillations; 3) DNA bending dynamics between similar states were observed in the absence of MutL (Chapter 3, Figure 3.3A); and 4) The dwell times of the observes states (~1-5 sec) are much longer than the MutS conformational states observed previously (typically <2 sec) (Qiu et al., 2015). To effectively monitor DNA bending in the presence of MutL, it may be necessary to redesign the DNA substrate to reposition the dyes outside of the expected footprint of the MutS:MutL complex. Unfortunately, the exact nature of this complex is not well understood.

Conclusion

The preliminary studies presented in this chapter sought to compare DNA bending by *Taq* MutS on T-bulge DNA with other contexts. While each of these experiments suffer from unique limitations, the preliminary results can be compared to those presented in Chapter 3.

GT mismatch DNA vs. T-bulge DNA

Studies on GT-mismatch DNA reveal that MutS is able to bend GT DNA into many more conformations than was observed with T-bulge DNA. These results are in agreement with previous studies in the absence of ADP or ATP. (Derocco et al., 2014; Sass et al., 2010) Not only are were DNA bending states observed in the GT DNA substrate, but there were also many more transitions between these states. Notably, the nucleotide-dependent kinetics of DNA bending observed here did not correlate well with previous studies (Qiu et al., 2012, 2015). As a result, development of a model for sliding clamp formation on GT DNA substrates that includes DNA bending, as was presented in Chapter 3 for T-bulge substrates, is impossible.

T-bulge DNA with blocked vs. free ends

On blocked DNA, FRET traces were identified that could be interpreted as MutS sliding clamps moving away from and then returning to the site of the T-bulge error. Similar FRET traces were observed on DNA with free ends (i.e. multiple DNA bending events were observed on the same molecule during the observation window), but these events were also rare. The rarity of these events makes direct comparison between the experiments with blocked and unblocked ends difficult. Further study is needed to see if these events are favored on blocked DNA substrates.

DNA bending by MutS vs. MutS:MutL

The DNA substrates used to study bending by *Taq* MutS are ill-suited to studying bending by larger complexes. This observation is likely due to the dyes being positioned within the protein complex's footprint on the DNA. Still, some FRET traces appeared to show signs of a conformational equilibrium between two DNA bending states, which may corroborate a previously identified MutL-induced MutS conformational equilibrium (Qiu et al., 2015). Further study, perhaps with a DNA substrate using different fluorophore positions, is necessary to further elucidate these events.

Materials and Methods

Protein and DNA substrates

C42A/M88C *Taq* MutS was provided graciously provided by Dr. Keith Weninger. His-tagged MutL was purified using nickel affinity chromatography as previously described (Qiu et al., 2015). The following fluorescently labeled DNA oligonucleotides (IDT) were used to prepare the 68 bp GT DNA substrate: b-CTC TAG AGG ATC CGC TGA GGC CAG CTG AGG CCT GGC TGA GGG ATT GCT GA(T*) GAA TTC ACT GGC CGT CG; dig-CGA CGG CCA GTG AAT TCA TCA GCA ATC **T**CT

CAG CCA G/iCy5/GC CTC AGC TGG CCT CAG CGG ATC CTC TAG AG. The following fluorescently labeled DNA oligonucleotides (IDT) were used to prepare the 68 bp T-bulge DNA substrate: b-CTC TAG AGG ATC CGC TGA GGC CAG CTG AGG CCT GGC TGA GGA TTG CTG A(T*)G AAT TCA CTG GCC GTC G; dig-CGA CGG CCA GTG AAT TCA TCA GCA ATC **T**CT CAG CCA G/iCy5/GC CTC AGC TGG CCT CAG CGG ATC CTC TAG AG. The "b" represents a 5′ biotinylation, and "dig" represent a 5′ digoxigenin modification added to facilitate blocking of the free end. The bold and underlined thymine is the single base insertion. TAMRA was linked to the thymine base denoted (T*), and an internal Cy5 was incorporated at the position marked /iCy5/. The sequence of this oligonucleotide was based on the DNA substrates used in previous studies, with the only change being substituting a thymine (T*) into the sequence to allow labeling (Qiu et al., 2012, 2015). The dye separation (19 bp) was selected based on previous studies (Derocco et al., 2014; Sass et al., 2010).

Single-molecule FRET experiments

Quartz slides were plasma cleaned and used to construct home-built flow cells. The slides were then pre-functionalized by sequentially incubating the flow cells with 1 mg/ml biotinylated BSA and 0.1 mg/ml streptavidin. Biotinylated fluorescently-DNA oligonucleotides were then immobilized onto the surface of the quartz slide (as depicted in Figure 4.1A). The samples were then imaged using a through-prism total internal reflection single molecule fluorescence microscope. Donor and acceptor excitation was achieved using 532 nm and 638 nm lasers respectively. The fluorophore emission is collected through a 60X water immersion, 1.2 N.A. objective, and the image is split by a DualView optical splitter with a 645 nm dichroic mirror. The donor and acceptor signals then pass through optical filters (i.e. a 585/70 bandpass filter for TAMRA, and a 655 longpass filter for Cy5) before detection by an emCCD camera. To observe changes in DNA bending over time, movies of approximately 1000 frames are collected using the following excitation sequence: 1) Brief excitation of the acceptor dye (~ 1 sec) to locate DNA molecules; 2) Excitation of the donor dye (~2 min) to monitor changes in FRET; 3) Brief excitation of the acceptor dye (~5 sec) to

reveal whether the acceptor has photobleached. All experiments were performed at room temperature in 50 mM Tris, 100 mM sodium acetate, 5 mM magnesium chloride, 2% glucose (w/v), pH 7.8, in the presence of an oxygen scavenging and triplet state quenching system of 100 U/ml glucose oxidase, 1000 U/ml catalase, 0.05 mg/ml cyclooctatetraene and 143 mM 2-mercaptoethanol.

Data analysis

Data analysis was carried out using the approach described in Chapter 2. In the GT mismatch experiments, only those FRET time traces exhibiting changes in FRET were analyzed; traces whose FRET was constant over the entire observation window were discarded. Only traces with intensities consistent with one MutS were analyzed. FRET traces showing evidence of protein:dye interactions (i.e. changes in fluorescence intensity of either the donor or acceptor dye that are not anti-correlated) were also discarded from analysis. For the end-blocked experiments and experiments with MutL, only traces interpretable traces were considered.

CHAPTER 5: LINKING DNA MISMATCH REPAIR AND PROTEOTOXIC STRESS: AN EXPLORATORY STUDY IN *SACCHAROMYCES CEREVISIAE*

Introduction

Proteasomes and molecular chaperones work together to maintain cellular proteostasis by either degrading incorrectly translated proteins or helping proteins to fold properly. Without these systems, erroneously translated proteins can misfold, aggregate, and cause cellular proteotoxic stress. However, the ability of the proteostasis machinery to minimize proteotoxic stress is limited to some maximum threshold of proteomic errors. Beyond this threshold, cells will no longer maintain proteostasis ultimately leading to proteotoxicity and cell death (Leak, 2014; Vermulst et al., 2015).

Some genes are intrinsically prone to causing proteotoxic stress, and several of these genes are linked to human disease. For example, the Huntingtin gene *Htt* contains a CAG trinucleotide repeat which can be erroneously expanded during DNA replication. The codon CAG encodes for the amino acid glutamine (Q), so expansion of the CAG repeat leads to incorporation of excess glutamines into the polypeptide chain. Depending on the number of additional glutamines, these poly-Q regions can then lead misfolding and aggregation of the HTT protein. Notably, 25 glutamines are insufficient to produce symptoms, while 105 glutamines will lead to aggregates of HTT. This aggregation will cause proteotoxic stress that ultimately will lead to Huntingtin disease (McMurray, 2010). Similarly, the protein TDP-43 forms protein aggregates associated with Lou Gehrig's disease (Latouche et al., 2006). Finally, the prion RNQ1 is associated with the formation of protein aggregates in yeast (Sondheimer & Lindquist, 2000).

Errors introduced during DNA replication lower the integrity of the genomic information. Fortunately, DNA repair processes such as DNA mismatch repair (MMR) to increase the fidelity of

replication such that errors are safely minimized. However, if left unrepaired, genomic errors will ultimately be transcribed in to RNA and translated into protein, leading an increase in sporadic proteomic errors (Kunkel & Erie, 2005).

Budding yeast (*Saccharomyces cerevisiae*) are a common eukaryotic model system to study DNA repair due their simplicity and relatively tractable genome. In yeast, DNA MMR is initiated when one of two isoforms of a heterodimeric MutS homolog recognizes an error in the DNA. MutSα (a heterodimer of Msh2 and Msh6) is responsible for recognizing base-base mismatches, while MutSβ (Msh2-Msh3) recognizes small insertion/deletion loop errors. Thus, by knocking out Msh2, MMR is completely removed, while knocking out either Msh3 or Msh6 will allow only a subset of DNA errors to go uncorrected. Any of these knockouts will have an increased pool in sporadic errors in their proteome relative to WT strains (Kunkel & Erie, 2005).

Using a galactose promoter, we induced expression of four known proteotoxic genes (*Htt25Q*, *Htt105Q*, *TDP-43*, and *RNQ1*) and assessed the survivability of four strains of yeast (wild type (WT), Msh2 deletion (Msh2Δ), Msh3 deletion (Msh3Δ), and Msh6 deletion (Msh6Δ)). We hypothesized that the MMR deficient strains would exhibit decreased viability due to an increase in proteotoxic stress caused by an increase in sporadic proteomic errors. Furthermore, we hypothesized that expression of genes that cause proteotoxic stress would further exacerbate the proteotoxicity by further increasing the pool or proteomic errors to surpass the maximum threshold for the proteostasis machinery, leading to decreased cell viability.

Materials and Methods

Yeast Strains and Plasmids

Wild type (WT), Msh2 deletion (Msh2Δ), Msh3 deletion (Msh3Δ), and Msh6 deletion (Msh6Δ) yeast strains missing the *URA3* locus were provided by Dr. Mara Duncan. Plasmids containing the *URA3* auxotrophic marker and the *Htt25Q, Htt105Q, TDP-43, and RNQ1* genes under an inducible galactose promoter were provided by Dr. Marc Vermulst.

Transformations

Four plasmids containing the genes and one control plasmid with no insert were transformed into each of the four yeast strains as follows: A saturated 3 ml culture of each yeast strain was grown overnight at 30°C in liquid YAPD media (1% yeast extract (w/v), 2% peptone (w/v), 2% glucose (w/v), 80 ng/ml adenine). This culture was used to inoculate a 100 ml culture of YAPD, which was incubated at 30°C until the OD_{600} reached 0.5. The cells were then spun down (3000 rpm for 3 min), and the supernatant was discarded. The cells were then washed with sterile water and repelleted, again discarding the supernatant. The cells were resuspended in fresh sterile filtered 10 mM Tris-HCl, 1 mM EDTA, and 100 mM lithium acetate at pH 7.5. Salmon sperm DNA was prepared at 2 mg/ml and boiled for 5 min. In a 1.5 ml tube, 0.2-2 µg of plasmid DNA and 30-50 µl of salmon sperm DNA were added and mixed thoroughly. Then, 100 µl of the yeast cell suspension and 600 µl of fresh sterile filtered 40% PEG 3350, 10 mM Tris-HCl, 1 mM EDTA, and 100 mM lithium acetate at pH 7.5 were added to the tube and mixed very vigorously. The transformation reaction was then heat shocked for 30 min at 42°C and then spun for ~15 sec to produce a cell pellet. The supernatant was discarded, and the cells were resuspended in 10 mM Tris-HCl and 1 mM EDTA at pH 7.5. 100 µl of the transformation reaction was then plated onto –URA Simple Complete media (6.7 mg/ml nitrogen base without amino acids, 2% glucose, 60 ppm adenine, 300 ppm leucine, 600 ppm histidine, 600 ppm methionine, 1200 ppm lysine, 600 ppm tryptophan, 6 ppm threonine, 3 ppm isoleucine,

3 ppm valine, and 3 ppm phenylalanine) to select for successful transformants. These plates were incubated at 30°C until transformants appeared, typically 3-4 days.

Culture Normalization and Dilution Plates

Successful transformants for each genotype:plasmid pair (N = 20) were cultured in triplicate overnight at 30°C in –URA Simple Complete minimal media containing 2% raffinose instead of glucose as the food source. All of these cultures were normalized by diluting an appropriate amount (usually 50-100 µl) of the culture in 1 ml of –URA Simple Complete media to the same OD_{600} +/- 0.01. Each of these normalized cultures (N = 60) were used as a stock solution (dubbed 1:1) for four serial dilutions (1:4, 1:16, 1:64, and 1:256). 7 µl of each of these dilutions (N = 180) was plated in duplicate on –URA Simple Complete media containing either 2% glucose (control) or 2% galactose (to induce expression of the gene on the plasmid). These plates were incubated at 30°C, and growth was monitored at 2, 3, and 4 days after the initial plating. The greatest differences in cell growth were observed on Day 3. This procedure was repeated three times with fresh transformants (Vermulst et al., 2015).

Results and Discussion

Representative images of the WT and Msh2Δ yeast strains grown on –URA plates are shown in Figure 5.1. As can be seen in the right column of Figure 5.1, *S. cerevisiae* cells are able to survive even in the absence of functional MMR (Msh2Δ). Upon galactose-induced expression of the proteotoxic genes *Htt25Q*, *TDP-43*, and *RNQ1*, notable differences in survivability between the WT and Msh2Δ yeast strains appear as shown in the left column of Figure 5.1. Galactose-induced expression of *Htt105Q* is lethal to both WT and Msh2Δ yeast, though WT yeast grew very slightly at the highest concentration of cells. While the vector only control does show differences between the WT and Msh2Δ strains in the presence of galactose, these differences are slight and not as significant as the differences observed in the experimental strains.

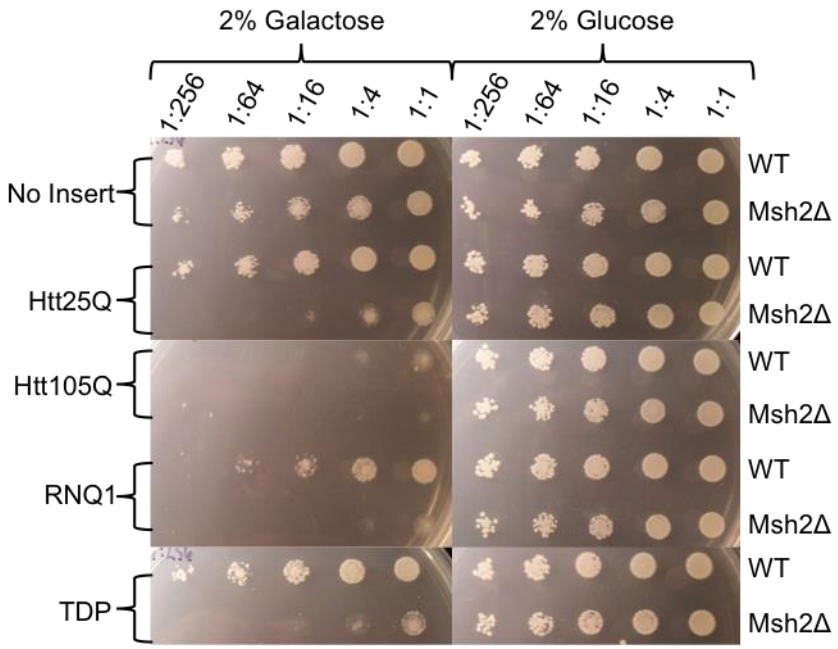

Figure 5.1: Msh2Δ strains are more sensitive than WT strains to proteotoxic stress.

Marked differences in survivability between WT and Msh2Δ yeast strains are observed upon galactose-induced expression of the *Htt25Q, RNQ1,* and *TDP-43* genes. The *Htt105Q* gene is lethal in both WT and Msh2Δ strains. No significant differences in survivability are observed in the presence of glucose (no gene expression).

Deletion of Msh2 completely knocks out mismatch repair, while deletion of either Msh3 or Msh6 knocks out repair of either insertion/deletion errors or base-base mismatches, respectively. To test if specific types of errors could be linked to proteotoxicity, the experiments were repeated in Msh3Δ and Msh6Δ strains. Representative images of these strains grown on –URA plates are shown in Figure 5.2. Again, no noticeable differences in survivability are apparent in between the different knock out strains on –URA media with glucose. Galactose-induced expression of either *Htt25Q* or *TDP-43* caused decreased survivability in the Msh2Δ, Msh3Δ, and Msh6Δ strains relative to the WT strain. Again, *Htt105Q* expression was lethal to all of the strains, most notably the Msh2Δ strain. Interestingly, upon expression of *RNQ1*, the survivability of Msh3Δ and Msh6Δ strains was decreased to relative to WT strains. However, this decrease was not as severe as the decrease noticed in the Msh2Δ strain. As expected, the vector only control show little to no effect on survivability.

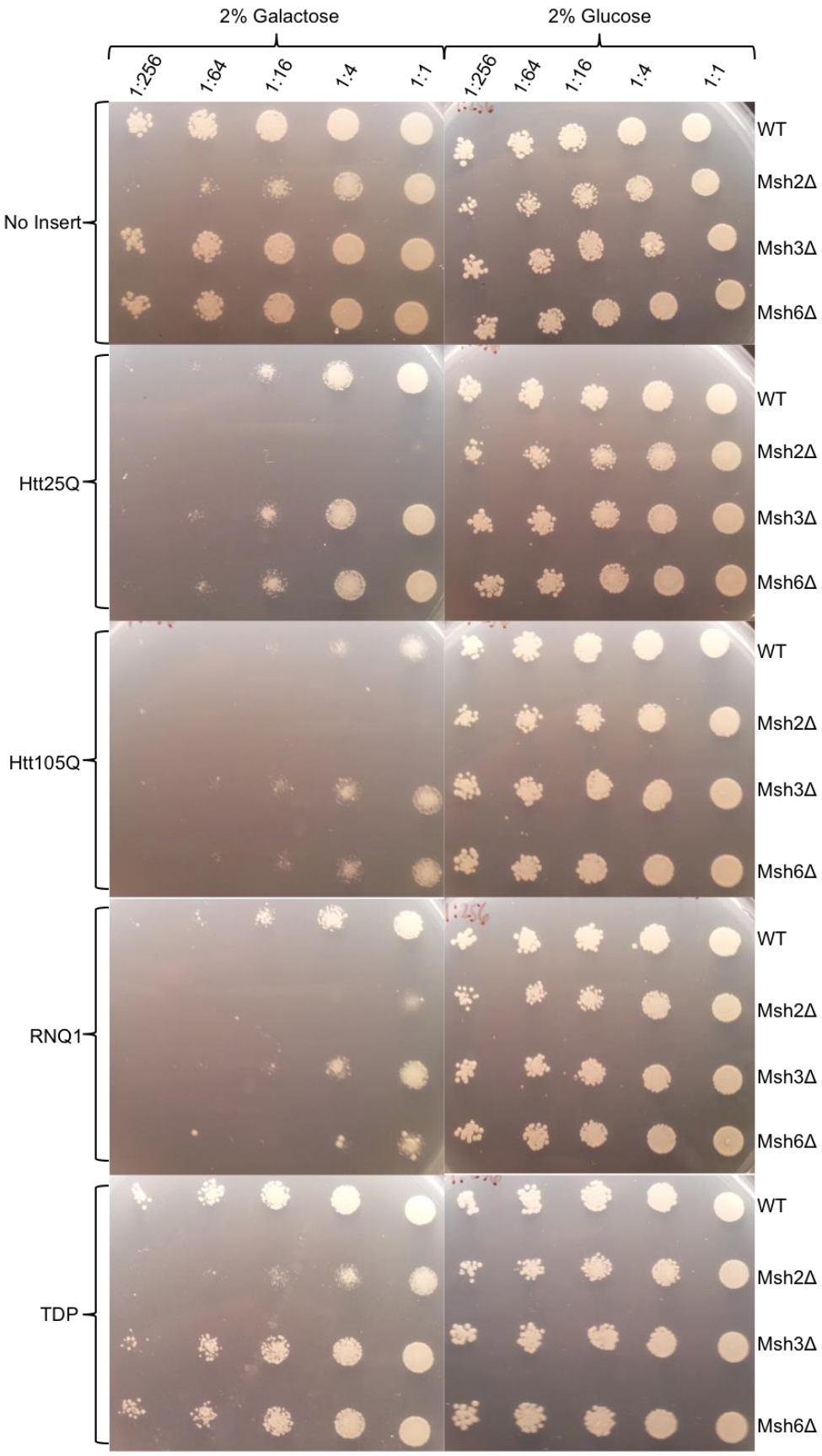
2% Galactose
2% Glucose
1:256
1:64
1:16
1:4
1:1
1:256
1:64
1:16
1:4
1:1
No Insert
Htt25Q
Htt105Q
RNQ1
TDP
WT
Msh2Δ
Msh3Δ
Msh6Δ

Figure 5.2: Msh3Δ and Msh6Δ strains display intermediate susceptibility to proteotoxic stress.

Marked differences in survivability between WT and Msh2Δ yeast strains are still observed upon galactose-induced expression of the *Htt25Q* and *TDP* genes; however, these differences are not apparent between the WT compared to the Msh3Δ and the Msh6Δ yeast strains. Upon expression of *RNQ1*, the survivability of the Msh3Δ and Msh6Δ strains are intermediate relative to the WT and Msh2Δ strains. The *Htt105Q* gene is lethal to all strains. No significant differences in survivability are observed in the presence of glucose (no gene expression).

Taken together, these data support the hypothesis that increased genomic errors resulting from deficiencies in mismatch repair can exacerbate proteotoxic stress leading to decreased survivability in yeast. These effects are clearly demonstrated in the strains expressing *Htt25Q* and *TDP-43* genes. Interestingly, in strains expressing *RNQ1*, there was a noticeable decrease in the survivability of the Msh2Δ strain relative to the Msh3Δ or Msh6Δ strains. This effect is expected, as deletion of Msh2 would knock out repair of all errors rather than one subset of errors, thereby producing a large pool of sporadic proteomic errors. This effect was not observed in any other experimental condition perhaps as a result of the relative crude sensitivity of this assay. Another expected observation was the more dramatic effect of *Htt105Q* relative to *Htt25Q*. The increased length of the poly-Q repeat is expected to lead to more misfolded and aggregated proteins, and therefore, decreased cell survival.

While these data do demonstrate that MMR deficient yeast expressing proteotoxic genes will have decreased survivability, they do not clearly demonstrate the mechanism behind this observation. We presume that the observed effects are the results of an increased pool of genomic errors that results in an increase pool of proteomic errors. We posit that the cells cannot survive this assault to their proteostasis and therefore die due to proteotoxicity. More experiments are necessary to demonstrate that proposed mechanism is true. For example, similar experiments should be done to test the survivability of yeast strains with both MMR proteins and chaperone proteins knocked out. These studies could provide more insight into the relationship between DNA repair and proteotoxicity.

APPENDIX A: USING THE TWO-COLOR TIRF MICROSCOPE

1) Turning the Scope On:

   a. Ensure that the appropriate emission filters are in place in the DualView splitter.

   b. Flip on the necessary lasers and the shutter control box using the switches (pink arrows below) and keys (blue arrows below). If you do not plan to use a specific laser, do not turn it on.

   c. **Be sure the shutter to the camera is closed**, and then turn on the camera (gold arrow below). Give the camera 20 minutes to warm up before collecting data.

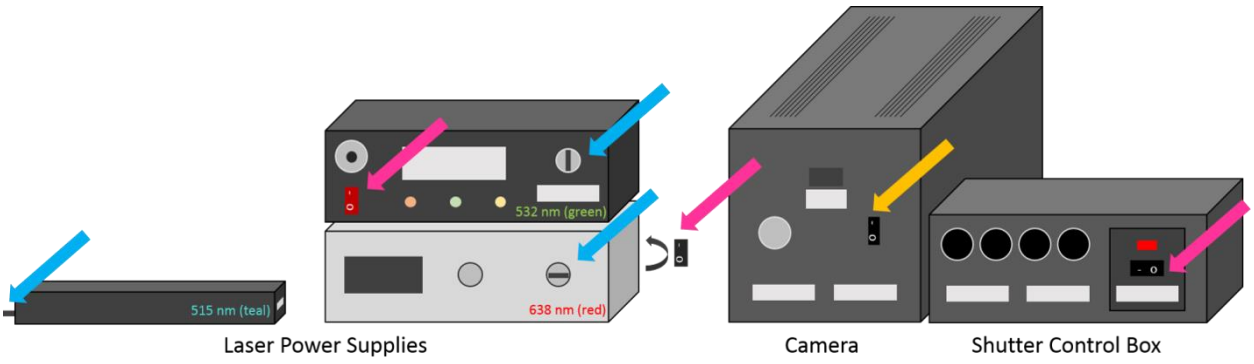

Figure A.1 Power-up Switches.

Lasers are turned on with switches (pink arrows) and keys (blue arrows). The camera is turned on with a switch (yellow arrow).

   d. Start the "DualView" computer, and open the program named WFI.exe and the "data" folder. (Note: There are shortcuts on the desktop for both.) Be sure the "data" folder is empty, as WFI.exe will write over any files already in the folder.

   e. Ensure that the computer has the capacity to store the amount of data you anticipate collecting. Movies around 1000 frames will take about half a gigabyte of storage space.

** This guide will describe aspects of WFI.exe useful in typical smFRET experiment. For more details on operation of WFI.exe, contact Dr. Keith Weninger, the program's author, at keith_weninger@ncsu.edu. **

2) Mounting a Slide (see the diagram on the next page):

** This description assumes that the objective you will use is the 60x UPlanApo water immersion objective (typical for a smFRET experiment). Be sure the desired objective is in the light path. If using a different objective (e.g., an oil immersion objective), be sure to use the appropriate immersive fluid. **

a. Lower the objective using the coarse adjustment knob to prevent the slide from hitting the objective lens. (Spin it clockwise.)
b. Pipet 30-70 µL of ddH_2O onto the objective. The resulting bead of water should be large enough such that it makes contact with the coverslip of the slide placed on top but sufficiently small so that it does not run off of the objective lens.
c. With a KimWipe, clean and dry the quartz slide loaded with your sample in the appropriate imaging buffer. Mount the slide **coverslip down** onto the stage. Secure it using tape or stage clips.
d. Move the stage so that the objective is directly below where you would like to image. Place a 30-70 µL drop of immersion oil (Type FF) on the slide above where you would like to image. Be sure none of the oil gets into the optics below the slide.
e. Pick up the prism mount **by the gold portion**. *Only one set screw secures the gold piece to the silver piece – holding the mount by the silver piece risks dropping the quartz prism!* Place the mount above the slide into the grooves. The prism should sit gently on the slide in the oil.

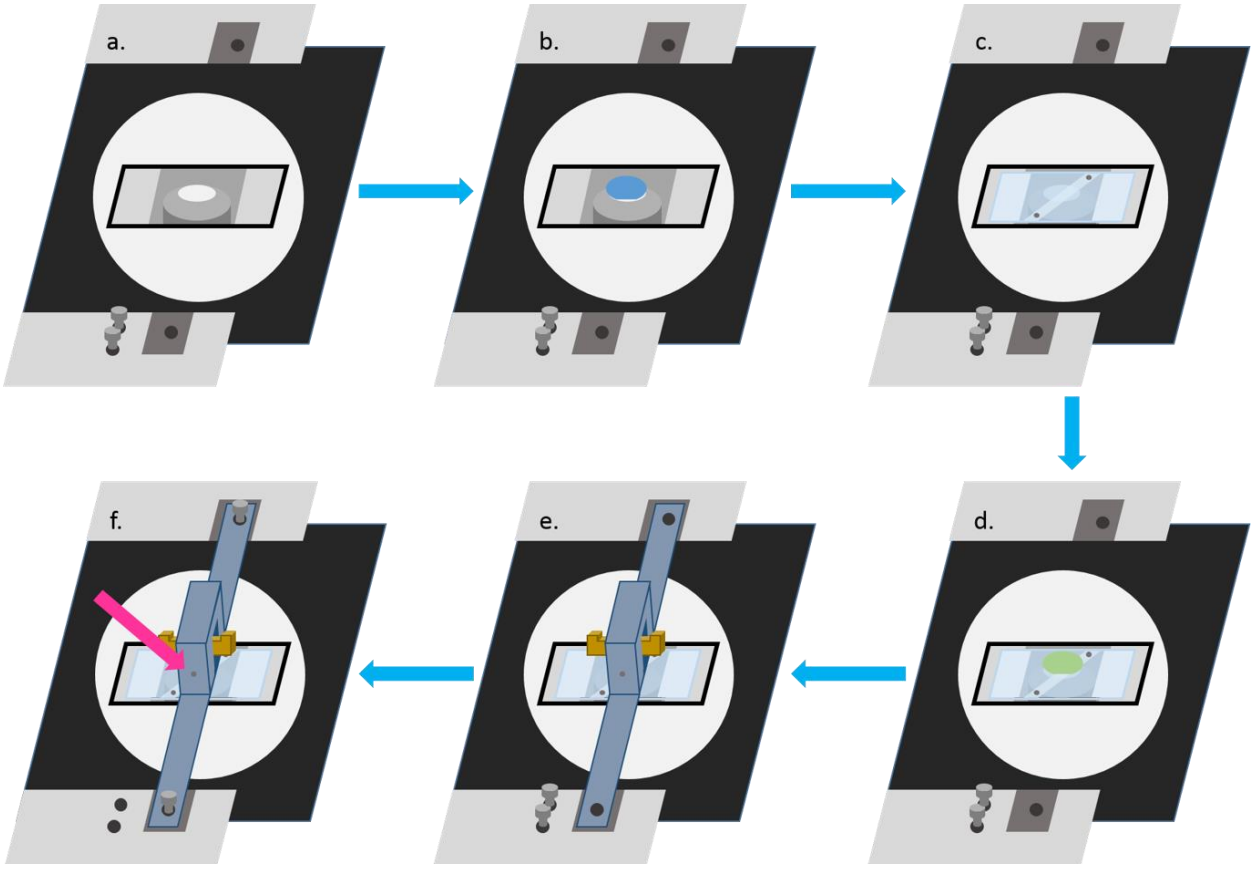

Figure A.2: Mounting a slide.

A) Microscope stage. **B)** Water added to objective. **C)** Slide mounted. **D)** Oil added to slide. **E)** Prism placed. **F)** Set screw adjusted.

** Note that the prism must be resettled each time a slide is mounted. This act is necessary because slight differences in slide thickness may cause the prism to sit too low (bowing the slide, which risks cracking the slide, prism, and/or objective) or too high (creating a gap between the prism and the slide, which disrupts the TIRF spot). Moving the prism changes the location of the TIRF spot such that it may not be directly above the objective (as shown to the right).* ***Resettling the prism is very challenging for new users and requires practice!*** *It is so challenging that when you cannot find a TIRF spot, you should assume that this step is what has gone wrong and resettle the prism. **

f. Partially screw the silver prism mount into place lining up the "X"s, but **do not fully tighten the screws**. Without allowing the gold prism mount to move left and right, loosen the set screw (pink arrow) with an Allen wrench. Allow the gold prism mount to settle, and retighten the set screw. Then hand-tighten the silver prism mount screws while ensuring the "X"s are aligned.

3) Focusing on a TIRF Spot:

a. Turn off the lights, and open a laser shutter using a switch on the shutter control box.

** Using the red laser to find the TIRF spot is usually best, as it has lower energy and will bleach less of the molecules while focusing. However, you can use the green lasers, if desired (e.g. for bead images). **

b. Look into the eyepieces. ***STOP LOOKING if the laser light is very bright!*** Don't ruin your eyes! Adjust the eyepieces as needed to fit the spacing between your eyes and to compensate for differences in your eye's vision. The light may be very bright because:

i. The slide may be positioned such that the laser is hitting tape or an air bubble, which causes a great deal of scatter. Move the stage slightly to look at a different spot.

ii. The prism may be positioned incorrectly such that the TIRF spot is not aligned properly within the field of view. Resettle the prism by adjusting the gold mount slightly to the left or right until there is no scatter throughout the prism.

** Note that when properly aligned, the laser should travel through the prism, reflect off of the slide, travel through the prism again, and then exit cleanly out the other side (Figure A.3, A). There should be little scattered light in the prism itself. If there is significant scatter (Figure A.3, B), consider slightly moving the prism to the left or right until the prism clears. **

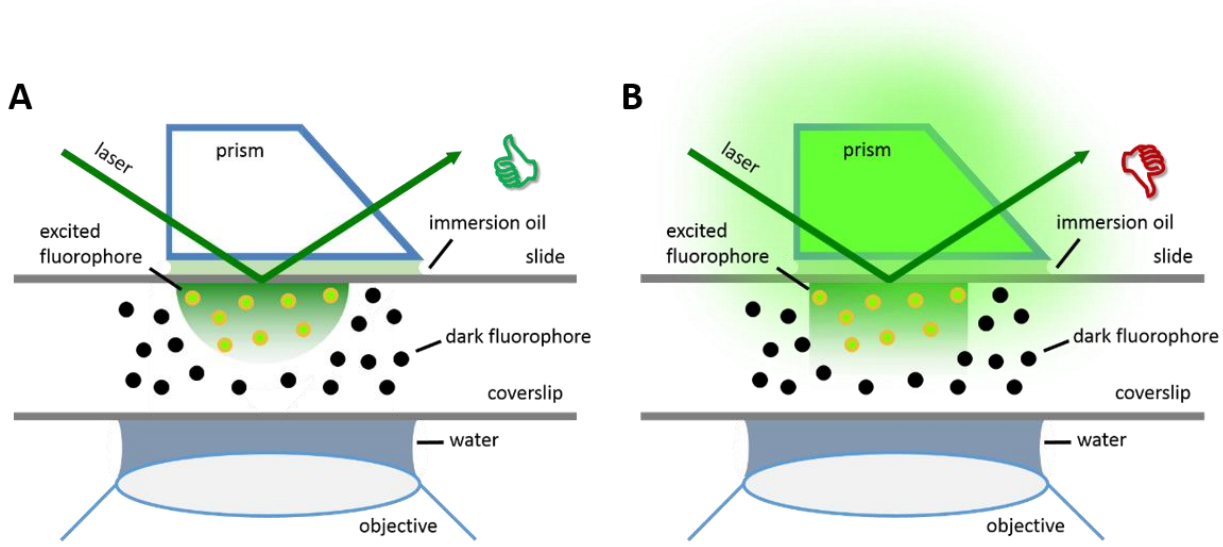

Figure A.3: Prism placement.

A) Proper prism placement will produce a relatively clear prism. **B)** Improper prism placement creates scatter.

c. Using the **fine adjustment knob**, **<u>slowly</u>** raise the objective closer to the slide. (Spin it counterclockwise.)

** **YOU MAY DAMAGE THE SCOPE** by ramming the objective into the slide. With the prism locked in place on the opposite side, this mistake can permanently crack the expensive objective lens! Be very careful! **

<u>NEVER</u> USE THE COARSE ADJUSTMENT KNOB TO FOCUS ON THE TIRF SPOT!

d. Continue focusing on the TIRF spot. ***STOP ADJUSTING*** the focus if:

iii. The spots of light visible through the eyepieces stop moving. **You have hit the slide with the objective!** Lower the objective with the fine adjustment knob (spin it clockwise).

iv. The entire field of view is dark. Either the laser is too dim or the TIRF spot may be misaligned. First, try turning up the laser power; if you can see laser light,

continue focusing. If not, try resettling the prism by adjusting the gold mount slightly to the left or right. Then, refocus on the TIRF spot.

IF YOU AREN'T SURE, <u>STOP</u> AND <u>ASK</u>!

** If you cannot find the TIRF spot after repeated tries, consider resettling the prism, moving it slightly to the left or right. If this still does not work, consider realigning the focusing lens. **

4) Aligning the TIRF Spots of Both Lasers:
    a. If properly aligned, the laser should travel through the center of the focusing lens to create a TIRF spot in the center of the view through the eyepieces.

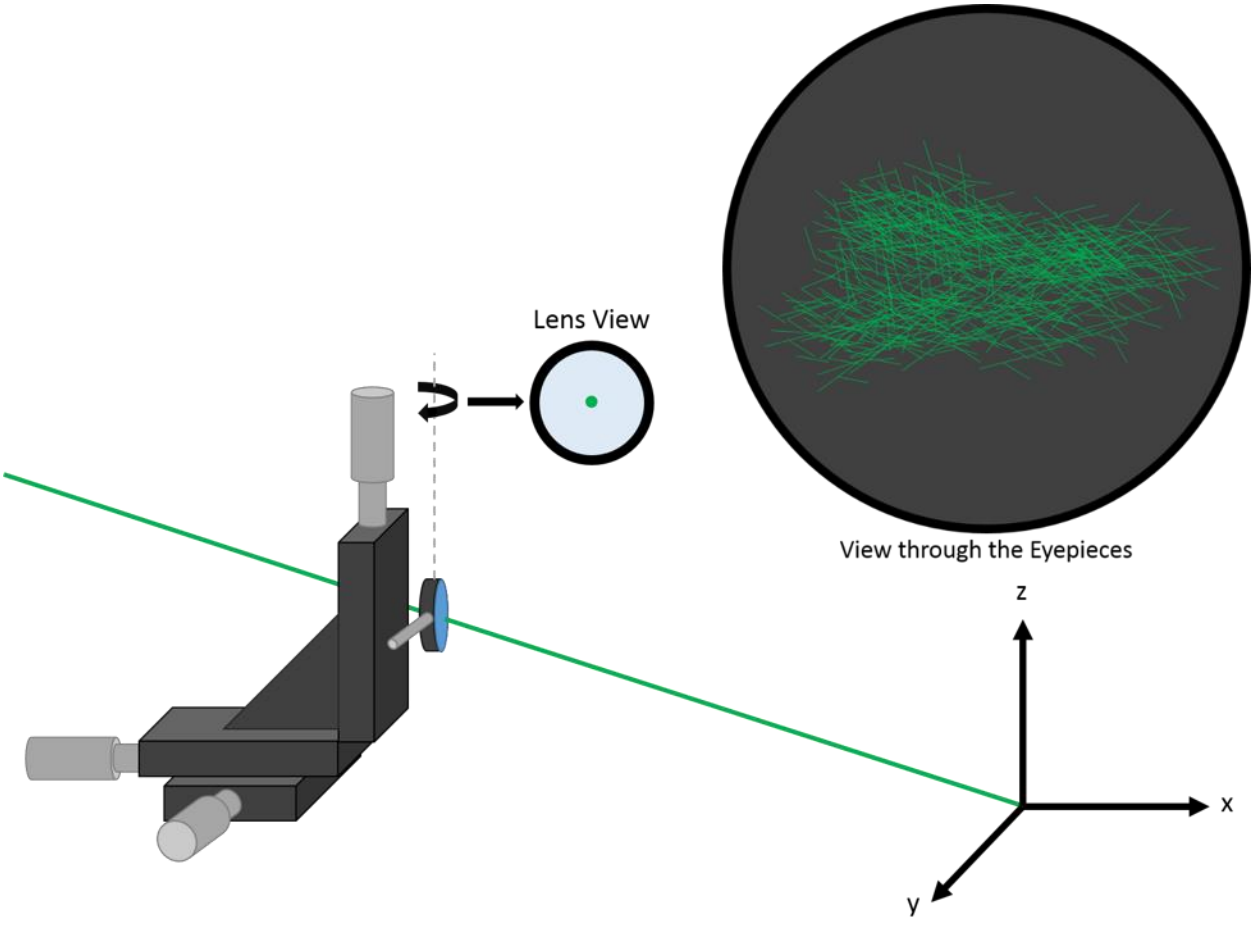

Figure A.4: TIRF spot.

With a good TIRF spot, scratches on the slide surface will be clearly visible and focus to a sharp image.

b. If the first laser's TIRF spot is off center, i.e. the most intense portion of the spot is not centered within the field of view, use the lens micrometers to move the spot (according to the instructions below and the diagrams on the next page).

** Note that the micrometers are calibrated. Before moving the micrometers, record their original position. That way, you can reset the scope to its previous settings if you need to start over. **

i. Use the front micrometer to move the TIRF spot in the eyepieces up and down.

ii. Use the upper micrometer and/or the left micrometer to move the TIRF spot in the eyepieces left and right.

c. Once the first TIRF spot is focused and centered, *look away from the eyepieces*.

** Note that having a clean TIRF spot from one laser does not mean that the TIRF spot from the other will be clean. There may be a lot of intense scattered light when turning on the second laser, so look away! **

d. Open the other laser shutter. Carefully look at the two spots to see if they are in the center of the field of view and overlap each other. ***STOP LOOKING if the laser light is very bright!*** Don't ruin your eyes! (Depending on the intensities of the lasers, you may need to turn the individual lasers off and on to see each spot.) If the two spots are:

i. Centered and overlapping – continue with your experiment.

** Note that differently sized spots are still usable so long as the spots are centered and overlapping as much as possible. Molecules in the area where the spots overlap will be excited by both lasers. **

ii. Overlapping but not centered – use the micrometers to center both spots.

iii. Not overlapping – start by centering the green/teal TIRF spot with the micrometers. Then, move the red spot to the green/teal spot using the alignment knobs on the "over" mirror nearest to the scope.

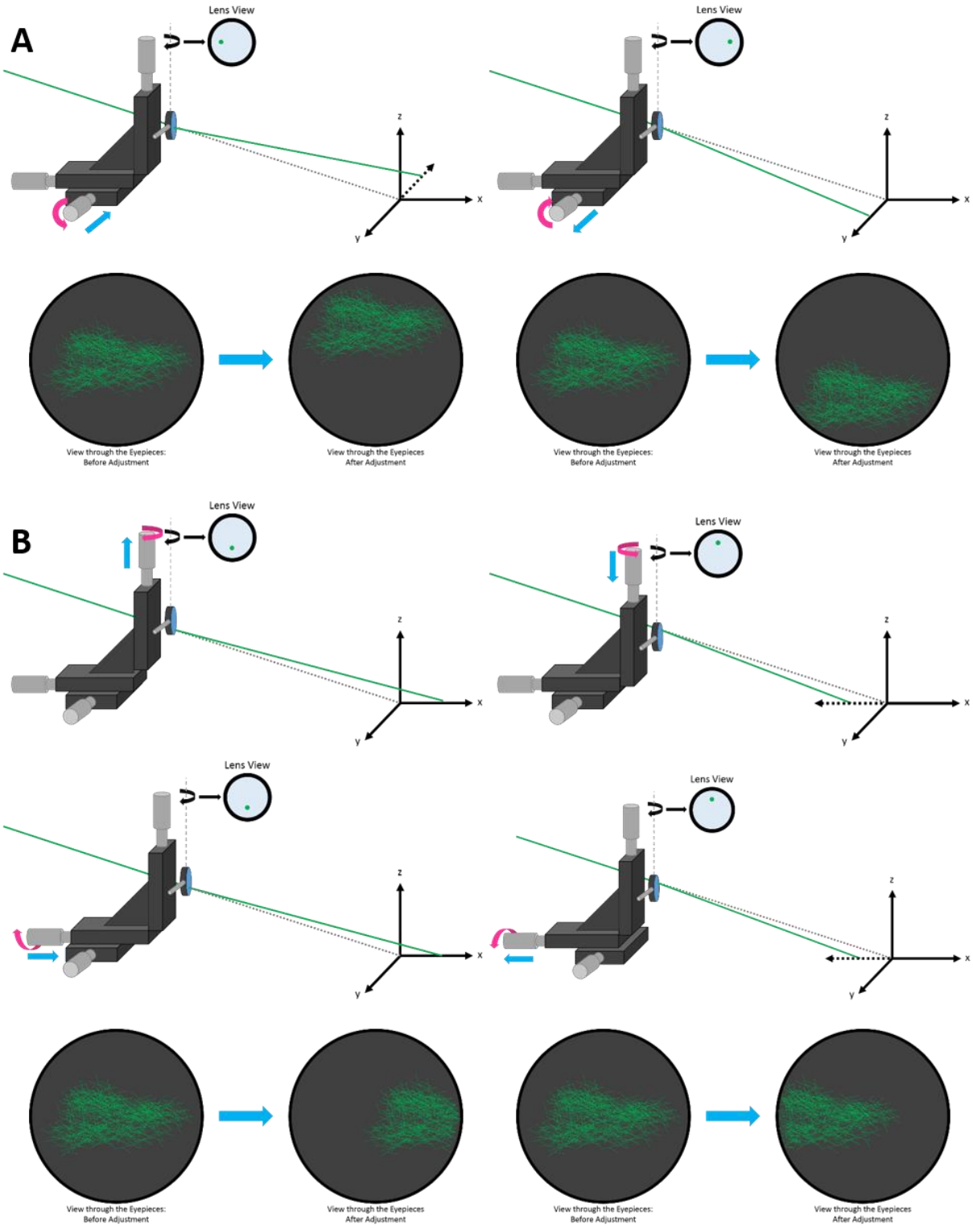

Figure A.5: Adjusting the TIRF spot.

A) Adjusting the TIRF spot up and down in the field of view. **B)** Adjusting the TIRF spot left and right in the field of view.

** It is possible that only one TIRF spot will be visible through the eyepieces, even with both laser shutters open, which most likely means that the lasers are misaligned. Using the micrometers, you can search for the second spot. If you find it, use the strategy described in iii. above to get the two spots to overlap. If you cannot find the second spot, consider resettling the prism OR realigning the whole table. **

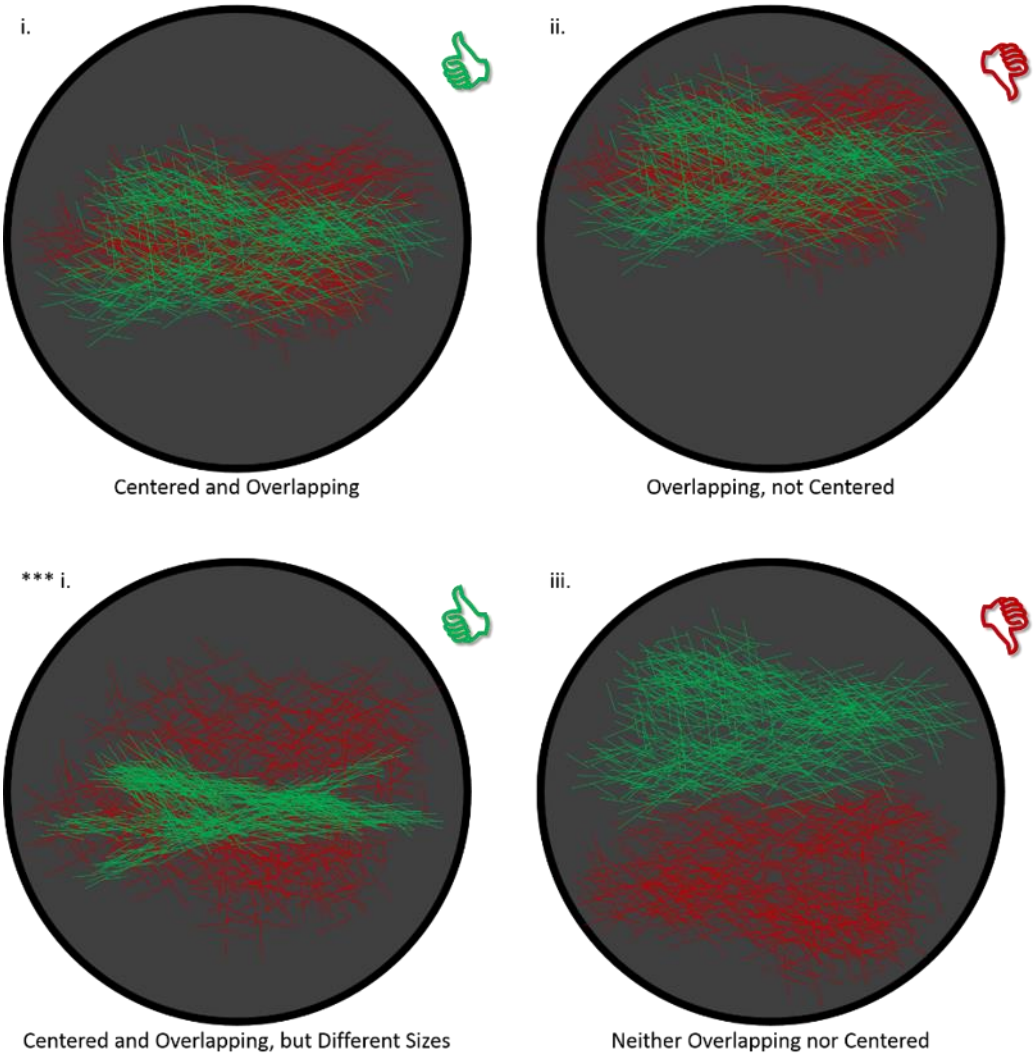

Figure A.6: Aligning the red and green TIRF spots.

Ideally, the spots will be centered and overlapping and the same size (i.) though different sizes are okay as well. (*** i.) Not centered (ii.) and not overlapping (iii.) is unacceptable.

5) Directing the Image to the Camera:

a. Turn off the lights. In WFI.exe, hit the *FOCUS* button to reveal the real-time image.

* ***YOU MAY DAMAGE THE SCOPE*** *by exposing the camera's detectors to the overhead lights, which are designed to detect a very small amount of photons, and exposure to high intensity light* ***can blow the detectors****. Be sure that the lights are turned off before opening the shutter to the camera.* *

NEVER EXPOSE THE CAMERA TO INTENSE LIGHT!

b. Open the shutter to the camera. If most of the field of view is white or orange, **close the shutter**! You are exposing the camera to too much light! Move to a new slide position.

WHEN IN DOUBT, CLOSE THE SHUTTER AND ASK!

c. With the appropriate laser(s) on, check to see if you can see fluorescence emissions from the individual molecules. If not, adjust the fine focus to sharpen the image.

* *Typically, the objective will need to be in a slightly different position to focus the image for your eyes versus the camera. Thus, a focused TIRF spot through the eyepieces may be appear out of focus when viewed by the camera. This difference is especially apparent if you have poor vision.* *

d. Ensure the TIRF spots are positioned properly as viewed by the camera. Be sure the spot(s) are:

i. Centered – if not, move them with the micrometers (as shown on the next page) until most of the spots in the camera's field of view are excited.

ii. Even – if not, use the top and left micrometers together to spread the spots (as described on the next page) until most of the spots in the camera's field of view are approximately the same intensity. Alternatively, consider introducing the diffuser if you'd like to spread only the green spot.

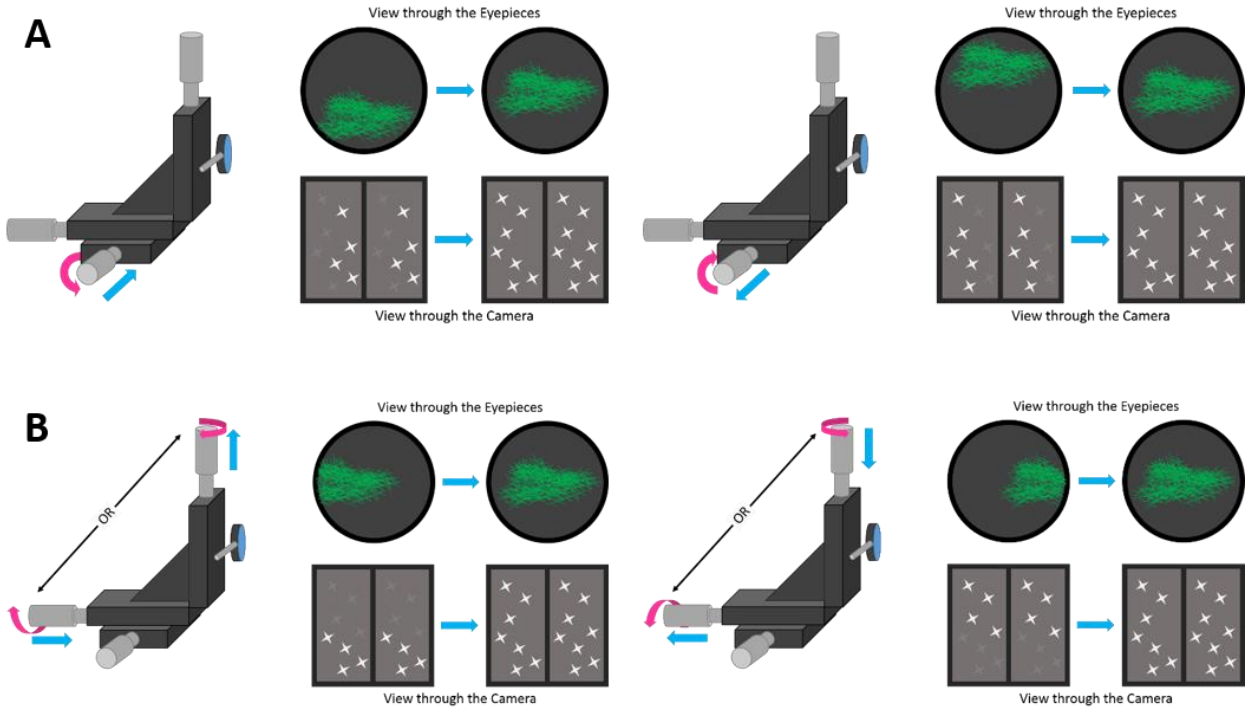

Figure A.7: Adjusting the TIRF spot for the camera.

A) Adjusting the TIRF spot up and down in the field of view moves the spot left and right for the camera.
B) Adjusting the TIRF spot left and right in the field of view moves the spot up and down for the camera.

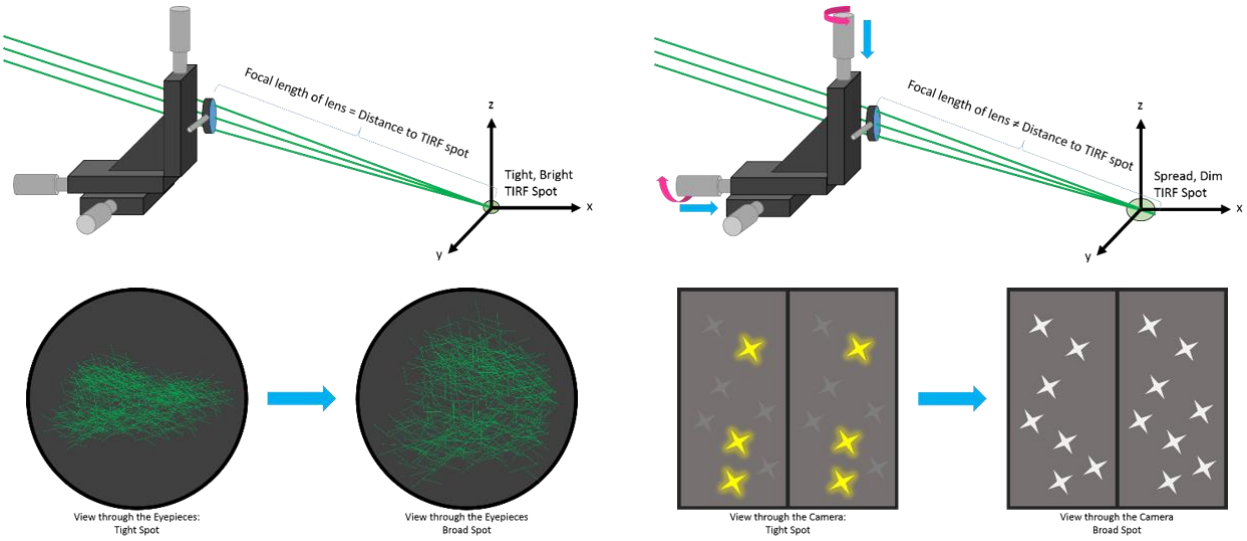

Figure A.8: Widening the TIRF spot for the camera.

Adjusting the TIRF spot up slightly out of the plane of focus will result a in a wider spot with more even excitation.

iii. Bright – if not, increase the laser power using the power supply or by adjusting the polarizer. Also, consider tightening the TIRF spot with the micrometers.

** After making all of these adjustments in one position on the slide, you may have bleached many or all of the molecules in that region. Consider moving to a new position before collecting data. **

e. In WFI.exe, if the background intensity makes identifying individual molecules difficult, try adjusting the Floor ADU (default value: 3900). Hit the *STOP* button, change the Floor ADU (in steps of 100 or so), and then hit the *FOCUS* button. Repeat until the background appears acceptably dark relative to the individual molecules.

6) Collecting Data:

a. Flip the switches on the shutter control box to the off position (down). The laser shutters can be controlled in WFI.exe using the *J*, *K*, and *L* keys for the red, green, and teal lasers, respectively.

b. To start and stop collecting data in WFI.exe, use the *spacebar*.

** A common mistake of first time WFI.exe users is to hit spacebar to stop the movie while WFI.exe is not the active window. This mistake will not stop the movie, and WFI.exe will continue collecting data until it is stops or crashes! If the program does not stop collecting data, be sure it is the active window! **

c. To collect a snapshot (e.g. for bead movies):

i. Open the shutter(s) for the appropriate laser(s).

ii. Ensure the molecules are in focus.

iii. Hit the *spacebar* twice in rapid succession (once to start, once to stop).

d. To collect a movie with constant excitation:

i. Open the shutter(s) for the appropriate laser(s).

ii. Ensure the molecules are in focus.

iii. Hit the *spacebar* to start the movie.

iv. After adequate time has elapsed, hit the *spacebar* to stop the movie.

e. To collect a FRET movie (0-15 frames, red; 15-1000 frames, green; 1000-X frames, red):

i. Be sure the *Red Green Alt Mode* box is checked. If it is not, hit the *STOP* button, check the box, and then hit the *FOCUS* button.

ii. Open the shutter(s) for the appropriate laser(s).

iii. Ensure the molecules are in focus.

iv. Leave only the red laser open (*J*).

v. Hit the *spacebar* to start the movie.

vi. The shutters will open and close automatically as programmed.

vii. After around 1020 frames, hit the *spacebar* to stop the movie.

** The extra 20 frames after 1000 (with the red laser on) are important for the background determination in the data analysis. Do not stop the movies after only 1000 frames! Also, do not let the movies run forever, as they will become very large data files (1000 frames is around half a gigabyte already!). **

f. If you would like to adjust other parameters in the data collection, such as frame rates or the time at which the shutters automatically switch open and closed, contact Dr. Keith Weninger, the program's author, at Keith_weninger@ncsu.edu.

7) Shutting Down the Microscope:

a. Close the shutter to the camera.

b. With all of the laser shutters closed, remove the prism mount and place it in its tray next to the microscope. Return the screws to their original place on the scope.

c. Remove your slide and cover the objective with a swatch of lens paper to avoid accumulation of dust on the objective.

d. Turn off the lasers, the camera, and the shutter control box using the switches and keys.

e. Close WFI.exe by hitting *STOP* and *QUIT*.

f. Cut and paste your data out of the data folder and into your own folder.

** It is important to 'cut and paste' (not 'copy and paste' or 'drag and drop') your data, as 'cut and paste' is a significantly faster way to transfer large amounts of data. **

g. As soon as possible, transfer your data from the "DualView" computer to a hard drive.

h. Check how much storage space remains on the "DualView" computer. If the computer is nearly at capacity, inform whoever is in charge of transferring the data to mass storage via killdevil. Transferring the data will open up space for the next user to collect data!

APPENDIX B: BUILDING AND ALIGNING A TIRF EXCITATION LASER PATH

1) Light Path Components:
    a. Lasers – Photodiode lasers at the desired wavelengths are used for excitation. Exciting at only one wavelength limits nonspecific background excitation. Other properties of the lasers include:
        i. Beam Diameter: 1 mm
        ii. Operation mode: Continuous wave (i.e. light is continuously emitted rather than pulsed)
        iii. Power: 100-300 mWatts
    b. Prism-Type TIRF – Total internal reflection occurs at the quartz:water interface of the slide because the refractive index of quartz is greater than that of water. This reflection creates an evanescent wave that penetrates only a few hundred nanometers into the sample. In this way, only molecules close to the slide are excited.

** The quartz prism, immersion oil, and quartz slide have the same refractive index; thus, minimal scatter, reflection, or refraction should occur as light travels through the interfaces between these materials. The prism is required to allow the beam to escape; without the prism, the beam would internally reflect through the slide. **

To excite the same region of a slide at multiple wavelengths, overlapping TIRF spots must be made from multiple lasers. To detect fluorescent emissions from this region, the overlapping TIRF spots must be positioned above the objective.

** Note that while this result can be achieved by directing the lasers to the spot via independent paths, best practice suggests aligning the lasers so that they travel the same light path to the path. This approach ensures that the TIRF spots will overlap and allows for simpler simultaneous realignment of all of the TIRF spots. **

c. Mirrors – Dielectric mirrors are be used to reflect the beam, as they scatter less light than regular mirrors, reducing losses in intensity.

** When directing the laser beam to the TIRF spot, use as few optics in the light path as possible. Each component is imperfect and will scatter/reflect some light; the incremental losses will add up! **

d. Lenses – When installing lenses, be sure the focal length is appropriate for the desired focusing and/or spreading. Notably, the lenses used to focus the TIRF spot should not be exchanged or mixed in with the other lenses, as they have been specifically chosen for their purpose.

e. Dichroic Mirrors – Dichroic mirrors have a critical wavelength (λ_{crit}). Light having a wavelength $\lambda < \lambda_{crit}$ will be reflected and light having a wavelength $\lambda > \lambda_{crit}$ will be transmitted. Using these properties, multiple laser beams can be brought into alignment on the same light path.

f. Half-Wave Plates and Polarizing Beamsplitting Cubes – Together, these components use polarization to control the beam intensity. A half-wave plate is made of a birefringent crystalline material, which allows it to change the polarization angle of an incident beam of polarized light (e.g. the laser beam). A polarizing beamsplitting cube splits an incident beam into the horizontal and vertical components. The horizontal component is transmitted through the cube, while the vertical component is reflected.

** Note that half-wave plates and polarizing beamslitting cubes are made for specific wavelengths! **

Used together, the intensity of an output beam can be controlled by rotating the half-wave plate, which changes the polarization angle of the light entering the polarizing beamsplitting cube. Since the magnitudes of the horizontal and vertical components of the polarized light depend on this angle, the intensity of the light transmitted through the cube is changed.

** Note that this approach produces a stray beam (the beam reflected by the cube). Be sure to install a beam blocker to stop the stray beam and prevent potential injuries. **

g. Diffuser - Since laser light is coherent (i.e. in phase), TIRF spots sometimes exhibit diffraction patterns, leading to uneven excitation of molecules within the TIRF spot. Introducing a spinning diffuser disc with pinholes into the beam path will render the beam incoherent (i.e. out of phase) due to the light periodically traveling through a medium with a different refractive index. The resulting TIRF spot will be more diffuse, even, and spread. Lenses are used to focus the beam through the pinholes.

2) Assembling and Aligning the Light Path

a. Secure the laser platforms in the desired positions on the table. Secure the lasers in place on the platforms pointing in the desired direction. Tape out the laser's outline onto the platform in case the laser is ever moved and needs to be replaced in the same position.

** It is best to orient the laser such that the beam will run parallel to the edge of the table. Use the screw holes in the breadboard as a guide to ensure that the laser is properly oriented. **

b. Attach the lasers to their power supplies and plug in the power supplies. If possible, put the power supplies on the table within easy reach of the microscope for ease of use.

** Cord management will quickly become an issue on the table top.* ***Keep the cords organized as you go!*** *Do not save untangling the cords until after the light path is assembled, as it will be very easy to knock the components out of alignment by struggling with the tangled cords. **

c. **Put on the laser safety goggles**. This safety gear is especially important during construction and alignment, as you do not know where the laser beams will be headed. ***Protect your eyes!***

* ***Always be aware of where the laser beam(s) are pointed.*** *If you are working as a team, be sure to point out the beam to each other and tell each other as you add new optical components. Be sure you do not direct the laser toward anyone's face.* ***Always know where the laser is and think about where it'll be!*** *

d. When it is safe, turn on the lowest energy laser (longest wavelength) at a very low power using the power supply. You only need enough power to just see the laser, no more.

* ***Be very careful when installing "up" mirrors.*** *These mirrors direct the laser toward eye level! *

e. Even on the platform, the laser beam's height will be too low. Install an "up" mirror immediately in front of the laser, and install an "over" mirror directly above the "up" mirror. Position this "over" mirror at the desired height of your light path.

* *Ideally, the majority of the light path should be built 4-6 inches above the table, parallel to the table top. Not only do we have the most hardware (posts, feet, etc.) to build the light path at this height, but this height also keeps he beam at a workable height that is safely below eye level! **

f. Construct a guide post by mounting a beam target onto a post at your desired light path height. Using this guide post, check to see if the beam leaving the initial "over" mirror is parallel to the table top by checking the beam's height in two positions (e.g. near to and far from the "over" mirror). Make adjustments to the "over" mirror as necessary until the beam is parallel.

g. "Rough in" the first laser's path to the TIRF spot using the dielectric mirrors. This path should include three final mirrors to be used by all of the lasers:

   i. An "up" mirror that directs the beam(s) from 4-6 inches above the table up toward

   ii. An "over" mirror taller than the microscope stage that directs the beam(s) toward

   iii. The final mirror that reflects the beam(s) at a shallow angle toward the objective.

** It is best to install the optics with the laser beam passing through the center and with any adjustment knobs at intermediate settings. Centering the beam prevents loss of intensity due to scatter, and both of these precautions ensure that there is adequate wiggle room for minor realignment later. **

h. Align the mirrors you have just installed so that the following criteria are met:
    i. Use the guide post to check (in two places) that the beam is parallel to the table.
    ii. Place a beam target on the face of each mirror to ensure that the beam hits each mirror near the center.

** While aligning the mirrors, it is best to first coarsely align the light path using the breadboard hardware (i.e. posts, pedestals, mounts, etc.). Only once everything is coarsely aligned to the best of your abilities and secured into place should you then use the adjustment knobs on the mirrors to finely align the light path. **

i. Continue installing the optical components (i.e. the half-wave plate, beamsplitting cube, and/or diffuser, as appropriate) specific to the first laser. Install them in the beam path paying special attention to their order. Leave space for, but do not install any dichroic mirrors at this point.

j. Repeat steps d.-i. with the next lowest energy laser using the same guide post to ensure the light path of all of the lasers are kept at the same height. Direct the second beam such that it intersects the first laser's light path at a right angle just before the final three mirrors.

k. At the point where the two beams cross, install the appropriate dichroic mirror (i.e. one with a critical wavelength between the wavelengths of the two lasers) such that the low energy beam is transmitted and the high energy beam is reflected. **CAREFULLY** align this mirror until:

i. The two beams are perfectly combined (i.e. overlapping perfectly) so that they hit each mirror in exactly the same spot and arrive together at a spot near the objective.

ii. The transmitted and reflected beams are centered on the mirror.

iii. The transmitted beam shows no reflection and the reflected beam shows no transmission (i.e. no stray beams are produced by the mirror).

** Note that the ability of a dichroic mirror to transmit and reflect light depends on the angle at which the beams hit the mirror. If the mirror is not properly oriented, some of the intensity may be lost. **

l. To add additional lasers, repeat steps j. and k. as needed in order of increasing energy. For each new laser, introduce a dichroic mirror to combine the beams prior to the final three mirrors.

** There are two major approaches to combining multiple lasers: from lowest to highest energy or from highest to lowest energy. **

m. Install an external shutter for each laser into the light path at a point where that laser has not yet been combined with the others. Plug the shutters into the control box, ensure that the cords are secured out of the light path.

3) Aligning the TIRF Spot

a. Use the coarse adjustment knob on the microscope to fully lower the objective. Adjust the final mirror of the light path until the laser spot(s) hits just to the right of the objective lens.

** To better see where the laser hits near the objective, put some water (around 50 µL) on the objective lens and then cover it with a scrap of lens paper (NOT A KIMWIPE!). This approach will give you a better view. **

b. Set the micrometers to an intermediate setting (i.e. around 5). Mount the TIRF focusing lens onto the micrometer setup. Adjust the TIRF focusing lens so that the beam passes through the center of the lens at a right angle. If the laser light is bent by the lens, use the micrometers to adjust the spot's position until it hits just to the right of the objective lens again (see below).

** Note that aiming off of the objective's center will compensate for both the refraction occurring at the air:quartz interface as well as the difference in height between the objective and a mounted slide. **

c. Ensure that all of the focused spots from all of the lasers overlap. If they do not, use the adjustment knobs on mirrors in the light path until they overlap.

** For this fine adjustment, it is best to use mirrors that are specific to each laser. Common choices include the "over" mirror just after the "up" mirror in front of each laser or the dichroic mirrors. **

d. Mount a prepared quartz slide filled with water. (See the Using the "DualView" TIRF Microscope protocol for details on how to mount a slide.)

e. Open the shutter to only one of the lasers. Look into the eyepieces. ***STOP LOOKING if the laser light is very bright!*** Don't ruin your eyes! Using the **fine adjustment knob**, **slowly** raise the objective closer to the slide until the TIRF spot comes into focus. (Spin it counterclockwise.)

** **YOU MAY DAMAGE THE SCOPE** by ramming the objective into the slide. With the prism locked in place on the opposite side, you can permanently crack the expensive objective lens! Be very careful! **

NEVER USE THE COARSE ADJUSTMENT KNOB TO FOCUS ON THE TIRF SPOT!

f. Continue adjusting the focus until the TIRF spot comes into focus. ***STOP ADJUSTING*** the focus if:

i. The spots of light visible through the eyepieces stop moving. **You have hit the slide with the objective!** Lower the objective with the fine adjustment knob (spin it clockwise).

ii. The entire field of view is dark. Either the laser is too dim or the TIRF spot may be misaligned. First, try turning up the laser power; if you can see laser light, continue focusing. If not, try resettling the prism by adjusting the gold mount slightly to the left or right. Then, refocus on the TIRF spot.

IF YOU AREN'T SURE, <u>STOP</u> AND <u>ASK</u>!

g. The TIRF spot may never come into focus, as getting it centered just above the objective is difficult. If you do not see the TIRF spot, adjust the following (in order):

i. <u>The Focusing Lens Micrometers</u> – While looking through the eyepieces, adjust the micrometers and the fine adjustment focusing knob on the microscope. As you make these adjustments, watch for signs that the TIRF spot is near the field of view through the eyepieces. Remember that it may be out of focus!

** Note that the micrometers are calibrated. Before moving the micrometers, note their original position. That way, you can reset the scope to its previous settings if you need to start over. **

ii. <u>The Prism</u> – Resettle the prism slightly to the right or left and try to refocus on the spot.

** Note that the laser should travel through the prism, reflect off of the slide, travel through the prism again, and then exit cleanly out the other side. There should be little scattered light in the prism itself. If there is significant scatter, consider slightly moving the prism to the left or right until it clears. **

iii. <u>The Final Mirror in the Light Path</u> – If all else fails, take the slide off of the scope and readjust the laser beam to hit just to the right of the objective in a new position. Remount the slide, and check for a TIRF spot again.

** Note that, to some degree, these adjustments end up being largely trial-and-error. Keep making small adjustments, noting what does and does not seem to work. **

h. To align the TIRF spots of multiple lasers, open the other laser shutter(s). Carefully look at the spots to see if they are all in the center of the field of view and overlapping. ***STOP LOOKING if the laser light is very bright!*** Don't ruin your eyes! (Depending on the relative intensities of the lasers, you may need to turn the individual lasers off and on to see each spot.) If the spots are:

i. Centered and overlapping – the microscope is sufficiently aligned for experiments.

** Note that differently sized spots are still usable so long as the spots are centered and overlapping as much as possible. Molecules in the area where the spots overlap will be excited by both lasers. **

ii. Overlapping but not centered – use the micrometers to center both spots.

iii. Not overlapping – start by centering the one TIRF spot with the micrometers. Then, use the alignment knobs on the mirrors that are specific to the other lasers to move the other TIRF spots until they overlap with the first.

** It is possible that only one TIRF spot will be visible through the eyepieces, even with more than one laser shutter open. Using the micrometers, you can search for the other spot(s). If you find it, use the strategy described in iii. above to get the spots to overlap. If you cannot find the other spot(s), consider repeating steps e.-g. above using a different color laser. **

4) Confirming the Alignment

a. Following the "Using the Scope" protocol, mount a slide with known contents (e.g. the bead slide or a slide prepared with DNA). Assess the image as viewed through the camera.

b. If necessary, make small adjustments to the micrometers until sufficiently even and intense TIRF excitation is achieved at the TIRF spot(s).

APPENDIX C: smFRET DATA ANALYSIS PROTOCOL

Stage 1 – Extracting intensity time traces of single molecules from the movies.

I. Data collected in a given experiment will be a series of .pma files with the name "cascadeX.pma" where X is a number. Some of these (typically at least 5) should be images of beads. (Figure C.1A)

II. To begin the analysis, run the complexingui2.m script. Click "Open" and open a bead movie. (Figure C.1B)

   a. Note that the color map can be changed with the drop-down menu in the upper left.

III. Choose "Find Offset" from the drop-down menu. The X and Y offsets will appear in the Command Window. Open and assess the offsets of all of the bead movies. (Figure C.1C)

IV. Choose "Locate Leftside Molecules". This option will circle local maxima on the left half of the image as well as the mapped X and Y coordinates (accounting for the offset) on the right half of the image. Local maxima too close to each other or too close to the edges are discarded. Inspect the circled maxima, and ensure that the local maxima are near the centers of the circles. (Figure C.1D)

   a. If the maxima are systematically off center, choose a different bead movie to establish a more appropriate offset.

   b. If no offset is appropriate, particularly if there seems to be rotational offset rather than simple X and Y displacement between the two sides, consider realigning to DualView.

V. Choose the bead movie that best represents the average offset, locate the leftside molecules, and check the "Use current mapping function" check box.

VI. Once the offset is established, choose "Batch Analysis Leftside". This option will extract the time traces for all of the local maxima in all of the movies in a given folder. (Figure C.1E)

a. Ensure that all movies to be analyzed in one batch are in the same folder. Be sure that it is appropriate for all of the movies in a given folder to have the same offset applied (i.e. the data were collected on the same day).

b. Note that this will take quite a bit of time, especially for movies with a lot of molecules.

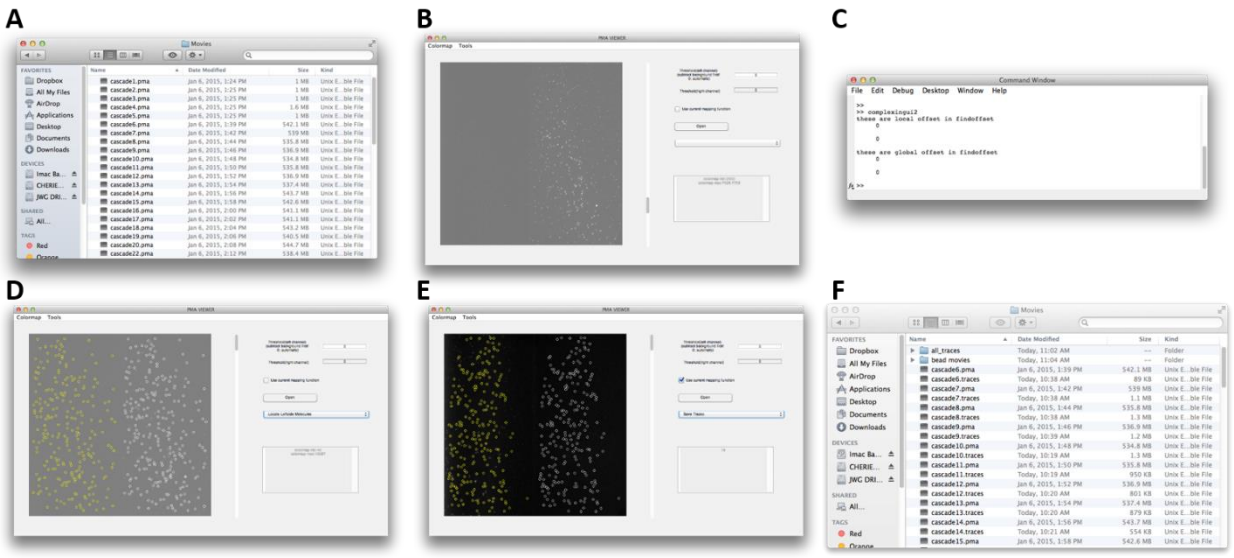

Figure C.1: Batch analysis of movies.

Stage 2 – Calculating FRET and detecting transitions from the single molecules.

I. The .traces files extracted from the .pma files should be in the same folder. Move the bead movies out of this folder, and create a new folder named "all_traces". (Figure C.1F)

II. Move the .traces files and the empty all_traces folder into the MATLAB folder. (Figure C.2A) The MATLAB folder must contain the following .m files necessary for this stage of the analysis:

a. align_extract_trans_data.m

b. baddata.m

c. batch_trace_analysis.m

d. CKedgedetection.m

e. FRET_ck.m

f. GKedgedetection.m

III. For the next stage of analysis, ensure the following .m files are also in the MATLAB folder:

a. new_bkgrd_TA.m

b. new_edges_TA.m

c. new_tmpts_TA.m

d. transition_analysis.m

IV. To begin the batch analysis, run the batch_trace_analysis.m script.

V. A dialog box will open with the title "Select the starting .traces file." Choose the .traces file you'd like to analyze with the lowest number. A second dialog box will open with the title "Select the final .traces file." Choose the .traces file you'd like to analyze with the highest number. (Figure C.2B)

VI. Three dialog boxes containing default analysis parameters will open successively. (Figure C.2C) These boxes allow you to input your own parameters for the following analyses:

a. The Chung-Kennedy Filter is an edge preserving filter used to smooth the donor and acceptor traces using the following parameters:

i. Windows – Used to calculate the averages ahead of and behind the data.

ii. Predictor Windows – Used to find the standard deviation ahead of and behind the data. Note that standard deviation increases when a transition occurs within the window.

iii. Exponents – Used to exaggerate the weight assigned to the averages. Note that the averages are weighted more when the standard deviation of the predictor window is low, and the exponent controls the extent to which this occurs.

b. The Gaussian Kernel is a mathematical method of detecting edges based on finding inflection points in the data (i.e. maxima/minima in the second derivative) using the following parameters:
    i. Thresholds – Used to determine how if maxima/minima are significant enough to count.
    ii. Scales – Used to convolve the data with a Gaussian function of different breadths.
    iii. Transition Width – Used to establish how separated transitions must be.
c. The Chung-Kennedy Edge Detection method uses the standard deviations of the predictor windows, which increase during transitions, to detect edges using the following parameters:
    i. Windows – Used to find the standard deviation ahead of and behind the data.
    ii. Percentile – Used to determine which peaks in the predictor window standard deviations qualify as a transition.
    iii. Separation – Used to establish how separated transitions must be.

VII. Once the parameters are entered, the batch_trace_analysis.m script will analyze each .traces file. Updates on its progress will appear in the command window. The following analysis is done with each single molecule's fluorescence intensity time traces:
a. The gamma and leakage corrections are applied to the donor and acceptor traces, respectively.
b. The corrected donor and acceptor traces are smoothed using the Chung-Kennedy filter by the FRET_ck.m script.

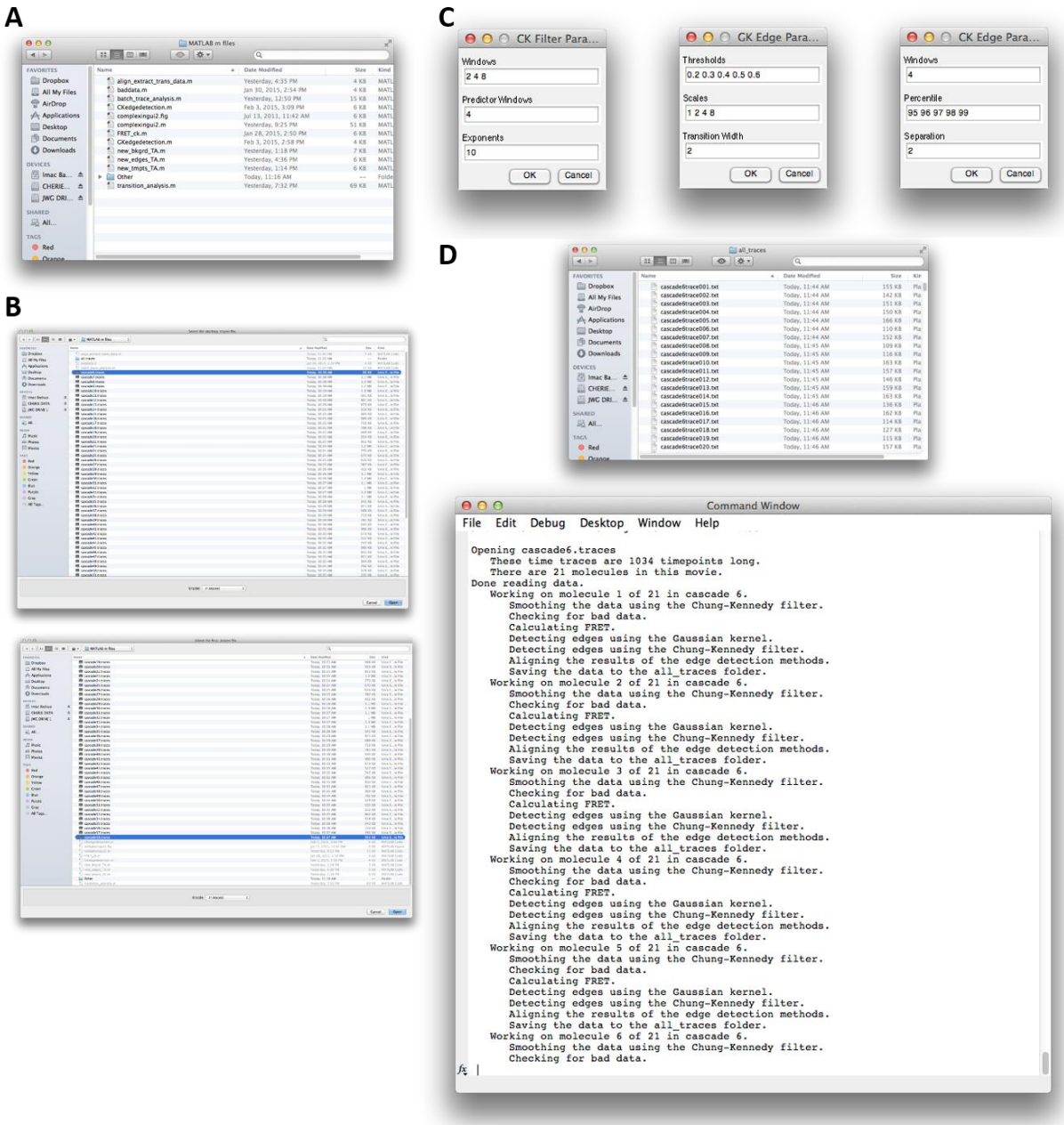

Figure C.2: Batch analysis of .traces files.

c. The data are checked for bleaches and blinks by the baddata.m script, which does the following:
    i. A 5 point boxcar average is applied to the data, and a 95% confidence interval is determined. If this interval contains zero, the data are considered “bad”.
    ii. Data points where the donor intensity and/or acceptor intensity are below 100 or above 3000 are considered “bad”.
    iii. Short-lived “good” or “bad” data are set to the quality of their neighbors (i.e. a single “bad” point surrounded by “good” points is set as “good”).
    iv. Molecules found to be all “bad” are discarded and not analyzed further.
d. FRET is calculated for the “good” data from both the raw and smoothed donor and acceptor traces. Residuals between the raw and smoothed FRET are calculated.
e. A histogram of the FRET at each time point is calculated.
f. Edges are detected by both the Gaussian kernel and the Chung-Kennedy method using the GKedgedetection.m and CKedgedetection.m scripts.
g. In the align_extract_trans_data.m script, the time resolution of the transitions detected by both methods is aligned using a weighted average. The averages of the donor, acceptor, and FRET between each transition are calculated. The standard deviations are also calculated and are used to determine 95% confidence intervals.
h. A .txt file containing the results of these analyses is saved into the “all_traces” folder. (Figure C.2D)

Stage 3 – Confirming the analysis, transitions, and “bad” data.

I. The .txt files within the “all_traces” files can be read in Excel (if desired) for individual assessment. For more interactive analysis, open the transition_analysis.m script.

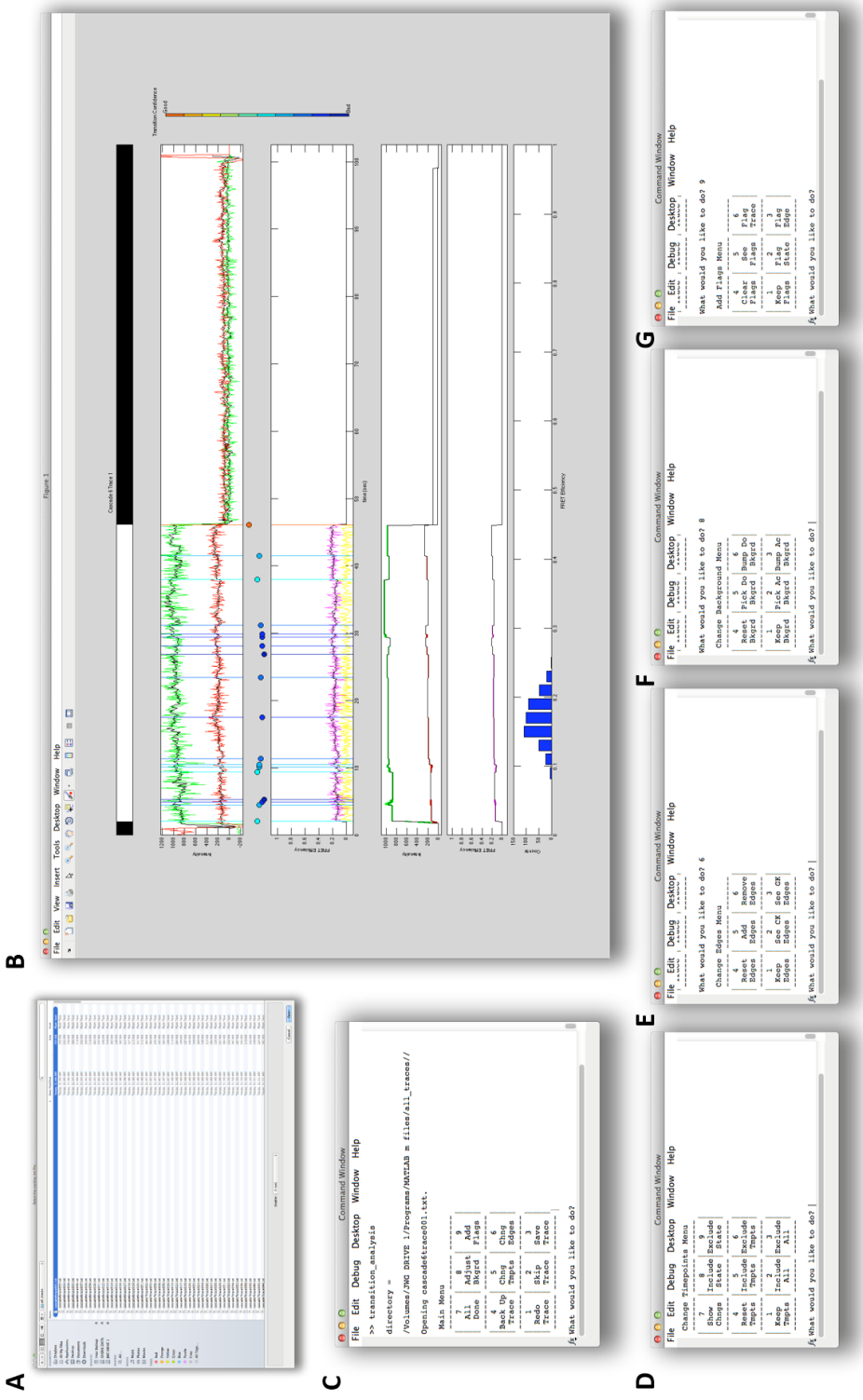

Figure C.3: User-interface for trace analysis.

II. A dialog box will open with the title “Select the starting .traces file.” Choose the .traces file you’d like to analyze with the lowest cascade and trace number. (Figure C.3A)

III. A figure will open with the following plots: (Figure C.3B)

a. A bar that is black where there is “bad” data and white where there is “good” data.

b. A line plot showing the raw donor (green) and acceptor (red) traces, the smoothed donor (black) and acceptor (black) traces, and the automatically detected transitions (vertical lines).

i. Note that the strength of the edge detection is denoted by the color of the vertical line according to the color bar to the right (red = obvious transition, blue = vague transition).

c. A series of dots for each transition, colored an arranged vertically according to the strength of the transition. These large dots are used to more easily select the transitions with the mouse.

d. A line plot showing the FRET calculated from the raw (magenta) and smoothed (black) donor and acceptor traces, the residuals (yellow) between the raw and smoothed FRET, and the automatically detected transitions (vertical lines).

e. A line plot showing the average donor and acceptor intensities (black) between the each transition and the 95% confidence intervals for the donor (green) and acceptor (red) averages.

f. A line plot showing the average FRET (black) between the each transition and the 95% confidence interval for the FRET (magenta) average.

g. A histogram of the (non-zero) smoothed FRET efficiency at each time point.

IV. A user interface to be used with a number pad and a mouse will appear in the command window. This user interface has the following menus and submenus:

a. Main Menu (Figure C.3C)

i. Redo Trace: Used to plot the current molecule using the .txt file.

ii. Skip Trace: Used to discard the current molecule and move to the next molecule.

iii. Save Trace: Used to accept the current molecule as analyzed (i.e. the edges, background, and "bad" data determination are all good), save the analysis results, and move to the next molecule.

iv. Back Up Trace: Used to back up to the most recently analyzed molecule.

v. Change Time Points: Opens a submenu to change which data are considered "bad". (Figure C.3D)

    1. Keep Time Points: Used to accept the changes to the "bad" data and analyze the data accordingly.
    2. Include All: Used to include all time points in the analysis.
    3. Exclude All: Used to exclude all time points in the analysis.
    4. Reset Time Points: Used to reset the "bad" data to that determined by the batch analysis script.
    5. Include Time Points: Used to include time points in a range chosen by the user using mouse left clicks. (Right click to escape.)
    6. Exclude Time Points: Used to exclude time points in a range chosen by the user using mouse left clicks. (Right click to escape.)
    7. Show Changes: Used to plot both the original and new "bad" data.
    8. Include State: Used to include a range of time points between two transitions as chosen by the user with a mouse left click. (Right click to escape.)
    9. Exclude State: Used to exclude a range of time points between two transitions as chosen by the user with a mouse left click. (Right click to escape.)

vi. Change Edges: Opens a submenu to change the transitions. (Figure C.3E)
    1. Keep Edges: Used to accept the current edges and analyze the data accordingly.
    2. See Chung-Kennedy Edges: Used to show only the edges detected by the Chung-Kennedy edge detection method.
        a. Note that you are given the option to keep only these edges, if desired.
    3. See Gaussian Kernel Edges: Used to show only the edges detected by the Gaussian Kernel.
        a. Note that you are given the option to keep only these edges, if desired.
    4. Reset Edges: Used to reset the edges to those found by the batch analysis script.
    5. Add Edges: Used to add edges using mouse left clicks. (Right click to escape.)
    6. Remove Edges: Used to remove edges using moues left clicks on the large dots between the intensity and FRET traces. (Right click to escape.)
vii. All Done: Used to compile all of the analyzed molecules, produce histograms of the FRET states and the dwell times, and make a transition density plot.
viii. Adjust Background: Opens a submenu to change the donor and acceptor traces' background subtractions. (Figure C.3F)
    1. Keep Background: Used to accept the current data corrections and analyze the data accordingly.
    2. Pick Acceptor Background: Used to set the acceptor intensity of a state between two transitions as zero. The user chooses this state with a mouse left click. (Right click to escape.)

3. Bump Acceptor Background: Used to bump the acceptor intensity up or down 10 intensity units by clicking above or below the data, respectively. (Right click to escape.)
4. Reset Background: Used to reset the data to those in the batch analysis script.
5. Pick Donor Background: Used to set the donor intensity of a state between two transitions as zero. The user chooses this state with a mouse left click. (Right click to escape.)
6. Bump Donor Background: Used to bump the donor intensity up or down 10 intensity units by clicking above or below the data, respectively. (Right click to escape.)

ix. Add Flags (Figure C.3G)

1. Keep Flags: Used to save the current flags.
2. Flag State: Used to flag the region between two transitions.
    a. Note that the flag can be no longer than 20 characters.
3. Flag Edge: Used to flag the transitions.
    a. Note that the flag can be no longer than 20 characters.
4. Clear Flags: Used to clear the current flags.
5. See Flags: Used to display the current flags.
6. Flag Trace: Used to flag the entire trace.

V. To analyze each molecule, use the following example workflow:

a. Decide if the molecule is worth analyzing at all. If not, skip to the next molecule.

i. Note that the "bad" data bar can help in determining if a molecule is worth analyzing, as it regions of data with intensities that are too low or too high will already be identified. (Figure C.4A vs. B)

b. If a bleach or blink is present, ensure that the background subtraction correctly places these intensities as zero. If not, use the "Change Background" submenu to correct the baseline. (Figure C.4B → C)

c. After applying any necessary background corrections, ensure that the time points included in the analysis are satisfactory. If not, use the "Change Time Points" submenu to correctly assign which data are "bad" data. (Figure C.4C → D)

d. Next, verify that the edges detected by the program are correct. Remove or add edges as appropriate using the "Change Edges" submenu. (Figure C.4D → E)

e. If desired, add flags at this time. (Figure C.4F)

f. When satisfied with the molecule's analysis, choose "Save Trace" to move to the next molecule. This creates a new .txt file named "analysis.txt" which contains the results of the analysis and the flags.

g. Continue analyzing until you have finished with all of the molecules. Then, choose "All Done".

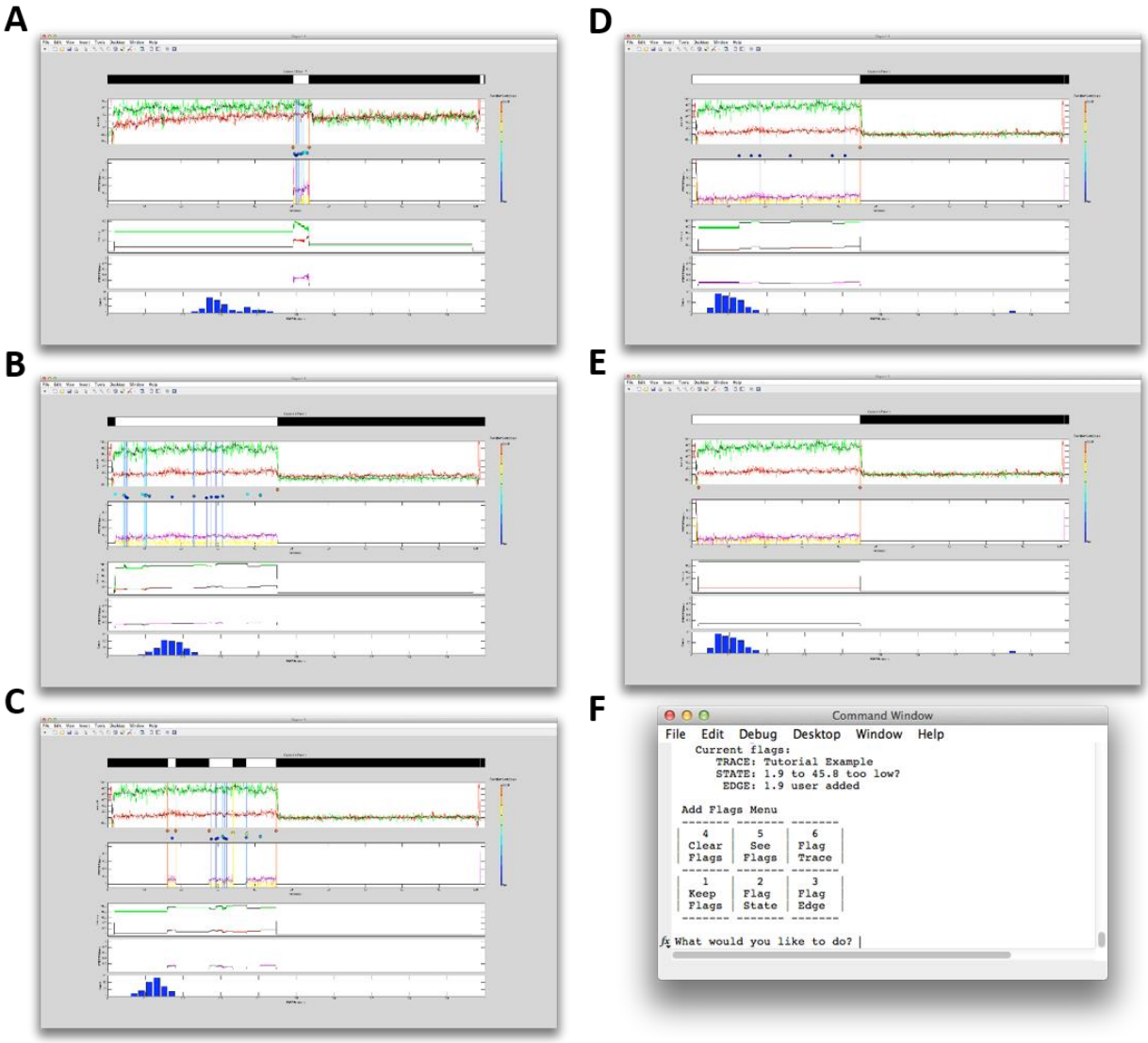

Figure C.4: Example analysis

REFERENCES

Alani, E., Lee, J. Y., Schofield, M. J., Kijas, A. W., Hsieh, P., & Yang, W. (2003). Crystal structure and biochemical analysis of the MutS.ADP.beryllium fluoride complex suggests a conserved mechanism for ATP interactions in mismatch repair. *The Journal of Biological Chemistry*, *278*(18), 16088–94. http://doi.org/10.1074/jbc.M213193200

Ban, C., Junop, M., & Yang, W. (1999). Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*, *97*(1), 85–97. http://doi.org/10.1016/S0092-8674(00)80717-5

Chung, S. H., & Kennedy, R. a. (1991). Forward-backward non-linear filtering technique for extracting small biological signals from noise. *Journal of Neuroscience Methods*, *40*(1), 71–86. http://doi.org/10.1016/0165-0270(91)90118-J

Constantin, N., Dzantiev, L., Kadyrov, F. a., & Modrich, P. (2005). Human mismatch repair: Reconstitution of a nick-directed bidirectional reaction. *Journal of Biological Chemistry*, *280*(48), 39752–39761. http://doi.org/10.1074/jbc.M509701200

DeRocco, V., Anderson, T., Piehler, J., Erie, D. a, & Weninger, K. (2010). Four-color single-molecule fluorescence with noncovalent dye labeling to monitor dynamic multimolecular complexes. *BioTechniques*, *49*(5), 807–16. http://doi.org/10.2144/000113551

Derocco, V. C., Sass, L. E., Qiu, R., Weninger, K. R., & Erie, D. A. (2014). Dynamics of MutS-mismatched DNA complexes are predictive of their repair phenotypes. *Biochemistry*, *53*(12), 2043–2052. http://doi.org/10.1021/bi401429b

Erie, D. a., & Weninger, K. R. (2014). Single molecule studies of DNA mismatch repair. *DNA Repair*, *20*, 71–81. http://doi.org/10.1016/j.dnarep.2014.03.007

Förster, T. (1959). Transfer mechanisms of electronic excitation. *Discussions of the Faraday Society*, *27*(10), 7. http://doi.org/10.1039/df9592700007

Gell, C., Brockwell, D., & Smith, A. (2006). *Handbook of single molecule fluorescence spectroscopy*. *Annals of Physics*. Retrieved from http://books.google.com/books?hl=en&lr=&id=_OcPA8B4We8C&oi=fnd&pg=PR11&dq=Handbook+of+Single+Molecule+Fluorescence+Spectroscopy&ots=5nh1f92WlL&sig=6_fachVS801Blrg04__5PRTsUZk

Genschel, J., Bazemore, L. R., & Modrich, P. (2002). Human Exonuclease I Is Required for 5’ and 3' Mismatch Repair. *Journal of Biological Chemistry*, *277*(15), 13302–11. http://doi.org/10.1074/jbc.M111854200

Haran, G. (2004). Noise reduction in single-molecule fluorescence trajectories of folding proteins. *Chemical Physics*, *307*(2-3 SPEC.ISS.), 137–145. http://doi.org/10.1016/j.chemphys.2004.05.017

Hsieh, P. (2001). Molecular mechanisms of DNA mismatch repair. *Mutation Research*, *486*(2), 71–87. http://doi.org/10.1016/S0921-8777(01)00088-X

Hsieh, P., & Yamane, K. (2008). DNA mismatch repair: Molecular mechanism, cancer, and ageing. *Mechanisms of Ageing and Development*, *129*(7-8), 391–407. http://doi.org/10.1016/j.mad.2008.02.012

Iyer, R. R., Pluciennik, A., Burdett, V., & Modrich, P. L. (2006). DNA mismatch repair: Functions and mechanisms. *Chemical Reviews*, *106*(2), 302–323. http://doi.org/10.1021/cr0404794

Jeong, C., Cho, W.-K., Song, K.-M., Cook, C., Yoon, T.-Y., Ban, C., … Lee, J.-B. (2011). MutS switches between two fundamentally distinct clamps during mismatch repair. *Nature Structural & Molecular Biology*, *18*(3), 379–385. http://doi.org/10.1038/nsmb.2009

Jiricny, J. (2006). The multifaceted mismatch-repair system. *Nature Reviews. Molecular Cell Biology*, *7*(5), 335–46. http://doi.org/10.1038/nrm1907

Kadyrov, F. a, Genschel, J., Fang, Y., Penland, E., Edelmann, W., & Modrich, P. (2009). A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(21), 8495–8500. http://doi.org/10.1073/pnas.0903654106

Kadyrov, F. A., Dzantiev, L., Constantin, N., & Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. *Cell*, *126*(2), 297–308. http://doi.org/10.1016/j.cell.2006.05.039

Kadyrov, F. A., Holmes, S. F., Arana, M. E., Lukianova, O. A., O'Donnell, M., Kunkel, T. A., & Modrich, P. (2007). Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease. *The Journal of Biological Chemistry*, *282*(51), 37181–90. http://doi.org/10.1074/jbc.M707617200

Kaur, G., Masoud, A., Raihan, N., Radzi, M., Khamizar, W., & Kam, L. S. (2011). Mismatch repair genes expression defects & association with clinicopathological characteristics in colorectal carcinoma. *The Indian Journal of Medical Research*, *134*(2), 186–92. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3181019&tool=pmcentrez&rendertype=abstract

Kunkel, T. a, & Erie, D. a. (2005). DNA mismatch repair. *Annual Review of Biochemistry*, *74*, 681–710. http://doi.org/10.1146/annurev.biochem.74.082803.133243

Lakowicz, J. R. (2006). *Principles of fluorescence spectroscopy*. *Principles of Fluorescence Spectroscopy*. http://doi.org/10.1007/978-0-387-46312-4

Latouche, M., Fragner, P., Martin, E., El Hachimi, K. H., Zander, C., Sittler, A., … Stevanin, G. (2006). Polyglutamine and polyalanine expansions in ataxin7 result in different types of aggregation and levels of toxicity. *Molecular and Cellular Neuroscience*, *31*, 438–445. http://doi.org/10.1016/j.mcn.2005.10.013

Leak, R. K. (2014). Heat shock proteins in neurodegenerative disorders and aging. *Journal of Cell Communication and Signaling*. http://doi.org/10.1007/s12079-014-0243-9

Longley, M. J., Pierce, A. J., & Modrich, P. (1997). DNA polymerase δ is required for human mismatch repair in vitro. *Journal of Biological Chemistry*, *272*(16), 10917–10921. http://doi.org/10.1074/jbc.272.16.10917

McCann, J. J., Choi, U. B., Zheng, L., Weninger, K., & Bowen, M. E. (2010). Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophysical Journal*, *99*(3), 961–970. http://doi.org/10.1016/j.bpj.2010.04.063

McMurray, C. T. (2010). Mechanisms of trinucleotide repeat instability during human development. *Nature Reviews. Genetics*, *11*(11), 786–799. http://doi.org/10.1038/nrg2917

Modrich, P. (1987). DNA mismatch correction. *Annual Review of Biochemistry*, *56*, 435–466. http://doi.org/10.1146/annurev.biochem.56.1.435

Obmolova, G., Ban, C., Hsieh, P., & Yang, W. (2000). Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, *407*(6805), 703–10. http://doi.org/10.1038/35037509

Pluciennik, A., Dzantiev, L., Iyer, R. R., Constantin, N., Kadyrov, F. A., & Modrich, P. (2010) PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proceedings of the National Acadamy of Sciences of the United States of America*, *107*(37), 16066-71.

Qiu, R., Derocco, V. C., Harris, C., Sharma, A., Hingorani, M. M., Erie, D. a, & Weninger, K. R. (2012). Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *The EMBO Journal*, 1–13. http://doi.org/10.1038/emboj.2012.95

Qiu, R., Sakato, M., Sacho, E. J., Wilkins, H., Zhang, X., & Modrich, P. (2015). MutL traps MutS at a DNA mismatch, (919). http://doi.org/10.1073/pnas.1505655112

Roy, R., Hohng, S., & Ha, T. (2008). A practical guide to single-molecule FRET. *Nature Methods*, *5*(6), 507–516. http://doi.org/10.1038/nmeth.1208

Sass, L. E., Lanyi, C., Weninger, K., & Erie, D. a. (2010). Single-molecule FRET TACKLE reveals highly dynamic mismatched DNA-MutS complexes. *Biochemistry*, *49*(14), 3174–90. http://doi.org/10.1021/bi901871u

Schofield, M. J., & Hsieh, P. (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annual Review of Microbiology*, *57*, 579–608. http://doi.org/10.1146/annurev.micro.57.030502.090847

Sondheimer, N., & Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. *Molecular Cell*, *5*(1), 163–172. http://doi.org/10.1016/S1097-2765(00)80412-8

Tessmer, I., Yang, Y., Zhai, J., Du, C., Hsieh, P., Hingorani, M. M., and Erie, D. (2008) *Mechanism of MutS searching for DNA mismatches and signaling repair*, *283*(52), 36646-54.

Vermulst, M., Denney, A. S., Lang, M. J., Hung, C., Moore, S., Mosely, A. M., … Erie, D. A. (2015). Shorten Cellular Lifespan. *Nature Communications*, *6*, 1–10. http://doi.org/10.1038/ncomms9065

Wang, H., Yang, Y., Schofield, M. J., Du, C., Fridman, Y., Lee, S. D., … Erie, D. a. (2003). DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(25), 14822–7. http://doi.org/10.1073/pnas.2433654100

Yang, Y., Sass, L. E., Du, C., Hsieh, P., & Erie, D. a. (2005). Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. *Nucleic Acids Research*, *33*(13), 4322–34. http://doi.org/10.1093/nar/gki708