Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________________________ _______________

Julie E. Coats Date

Approval Sheet

Single-molecule Studies of DNA Mismatch Recognition

By
Julie E. Coats
Doctor of Philosophy
Physics

____
Ivan Rasnik
Advisor

____
Keith Berland
Committee Member

____
Paul Doetsch
Committee Member

____
Kurt Warncke
Committee Member

____
Eric Weeks
Committee Member

Accepted:

____
Lisa A. Tedesco, Ph.D. Dean of the James T. Laney School of Graduate Studies

____________________ Date

Single-molecule Studies of DNA Mismatch Recognition

By

Julie E. Coats

B.A., Hendrix College, 2006

Advisor: Ivan Rasnik, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

2012

Abstract

Single-molecule Studies of DNA Mismatch Recognition

By Julie E. Coats

The DNA mismatch repair system protects the genome from spontaneous mutations by recognizing and repairing DNA synthesis errors in a pathway that is highly conserved. The MutS family of proteins initiate DNA mismatch repair by specifically binding mismatched or extrahelical bases and communicating the presence of damage to downstream repair proteins in an ATP-dependent manner. Previous structural studies have implied that MutS-induced conformational changes on DNA are central to damage recognition. Because the conformational changes occur on the timescale of seconds, it is difficult to obtain kinetic information on this highly dynamic process with traditional ensemble techniques. In this work, we use single-molecule fluorescence resonance energy transfer to investigate the conformational dynamics of mismatched DNA substrates in the absence and presence of DNA mismatch recognition proteins. We present quantitative kinetic information on the dynamics of DNA substrates and on the rates of MutS binding and dissociation in a variety of buffer conditions.

Single-molecule Studies of DNA Mismatch Recognition

By

Julie E. Coats

B.A., Hendrix College, 2006

Advisor: Ivan Rasnik, Ph.D.

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

2012

Acknowledgements

Dr. Ivan Rasnik has been a great advisor and has given me every opportunity that I could ask for as a graduate student. During my time in his laboratory, Dr. Rasnik has provided me with interesting projects and necessary resources, as well as wonderful guidance and encouragement. I have been very fortunate to learn to carry out single-molecule FRET experiments from Dr. Rasnik, a person who has pioneered many of the methods that have made single-molecule FRET a useful and practical tool to study the conformational dynamics of biological molecules. Most importantly, Dr. Rasnik taught me to appreciate and cherish the opportunity to explore and wonder about nature.

Dr. Yuyen ‘Kent’ Lin has been a great colleague during my time in Dr. Rasnik’s lab, and he has taught me many biochemistry and molecular biology techniques. The *E. coli* MutS project would not have been possible without his oversight and expertise. Most importantly, Kent has been a great friend to me, and I’m grateful to know him.

The members of my dissertation committee, Dr. Keith Berland, Dr. Paul Doetsch, Dr. Kurt Warncke, and Dr. Eric Weeks, have provided great direction and encouragement to me during my graduate study.

It has been exciting to collaborate with Dr. Cynthia McMurray, Dr. Walter Lang, Dr. Jerzy Majka, and Dr. Greg Hura on the MSH2/MSH3 project.

I appreciate the technical and the administrative help of Jason Boss, Art Kleyman, Cody Anderson, Horace Dale, Jon Carr, Connie Copeland, Paulette Evans, and Calvin Jackson.

This work was supported by a NSF CAREER grant (# 0748642).

Table of Contents

Chapter 1: Introduction 1

Figure 1.1 10

Table 1.1 11

Chapter 2: Single-molecule fluorescence resonance energy transfer 12

Figure 2.1 28

Figure 2.2 28

Figure 2.3 29

Figure 2.4 30

Figure 2.5 30

Figure 2.6 31

Figure 2.7 32

Chapter 3: Discrete conformational dynamics in three-way DNA junctions 33

Figure 3.1 52

Figure 3.2 53

Figure 3.3 54

Figure 3.4 55

Figure 3.5 56

Figure 3.6 57

Figure 3.7 58

Figure 3.8 59

Figure 3.9 60

Chapter 4: Conformational trapping of Human Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops…………………………….. 61

Figure 4.1 ……………………………………………………………………… 83

Table 4.1 ……………………………………………………………………….. 84

Table 4.2 ……………………………………………………………………….. 84

Figure 4.2 ……………………………………………………………………… 85

Figure 4.3 ……………………………………………………………………… 86

Figure 4.4 ……………………………………………………………………… 87

Figure 4.5 ……………………………………………………………………… 88

Figure 4.6 ……………………………………………………………………… 89

Figure 4.7 ……………………………………………………………………….90

Figure 4.8 ……………………………………………………………………….91

Table 4.3 ……………………………………………………………………….. 92

Table 4.4 ……………………………………………………………………….. 92

Figure 4.9 ……………………………………………………………………….93

Figure 4.10 …………………………………………………………………….... 94

Figure 4.11 …………………………………………………………………….... 95

Chapter 5: The binding kinetics of *E. coli* MutS influence DNA mismatch selectivity……………………………………………………………… 96

Figure 5.1 ………………………………………………………………….... 112

Figure 5.2 ……………………………………………………………………113

Figure 5.3 ……………………………………………………………………114

Figure 5.4 ……………………………………………………………………115

Figure 5.5 ………………………………………………………………….... 116

Figure 5.6 ………………………………………………………………….... 117

Figure 5.7 ………………………………………………………………….... 118

Chapter 6: Conclusion …………………………………………………………… 119

Appendix A ……………………………………………...……………………… 126

Appendix B ……………………………………………………...……………… 155

References ……………………………………………………………………… 172

Author's Biography ………………………………………………………...…180

Chapter 1

Introduction

1.1 Introduction to single-molecule studies

The field of molecular biophysics seeks to understand the function of biological systems in terms of their structures and dynamics at various levels of complexity. One of the most exciting developments in the field of molecular biophysics in the past 20 years is the emergence of novel techniques to study biological macromolecules on the single-molecule level [1]. Single-molecule techniques can detect the structures, dynamics, and forces produced by individual molecules that cannot be measured using traditional ensemble techniques [1]. Observations made using single-molecule tools have enhanced the understanding of the functions of a number of biological macromolecules involved in cellular process ranging from DNA metabolism to intracellular motility [4].

Single-molecule fluorescence resonance energy transfer (FRET) is an extremely powerful single-molecule technique that is useful for measuring changes in intramolecular distances ranging from 20-80 Å [4-6]. This technique has proven very useful for studying the conformational dynamics of nucleic acids [7-9] and protein-nucleic acid systems [10-16]. In the work presented here, single-molecule FRET is used to study the conformational dynamics of non-canonical DNA substrates in the absence and presence of DNA mismatch recognition proteins.

1.2 Introduction to the DNA mismatch repair

Both prokaryotic and eukaryotic cells are capable of repairing mismatched bases and extra-helical loops that occur during DNA replication by the highly conserved DNA mismatch repair (MMR) pathway [17-19]. In order for DNA mismatch repair to occur, proteins involved in this pathway must work together to accomplish three basic tasks: (1) detect mismatched or extra-helical loops, (2) remove the lesion, and (3) re-synthesize the excised DNA. The DNA mismatch repair pathway of *E. coli* has been characterized in the most detail, and the components of the DNA mismatch repair pathway of *E. coli* have served as the prototype for understanding DNA mismatch repair in eukaryotes [17-19]. In *E. coli*, DNA mismatch repair is initiated by binding of the homodimeric mismatch recognition protein MutS to a base-base mismatch or a 1-4 nucleotide extra-helical loop. Then, MutS interacts with a second homodimer, MutL, in an ATP-dependent manner. Assembly of the MutS-MutL complex leads to activation of the endonuclease activity of MutH, which nicks the newly synthesized (unmethylated) strand at the d(GATC) sequence. This incision—which can occur on either side of the mismatch—confers strand specificity to the DNA mismatch repair pathway, directing repair exclusively to the newly synthesized strand containing the error. The MutH-generated incision is the entry point for MutL-dependent loading of DNA helicase II and binding of single-strand DNA binding protein (SSB). Working together, these proteins generate single-stranded DNA (ssDNA) that is digested by either 3' or 5' exonucleases, depending on the location of the nick relative to the mismatch. After the error is excised, DNA polymerase III re-synthesizes the excised DNA and DNA ligase seals the nick to complete the DNA mismatch repair pathway. A schematic detailing the DNA mismatch repair pathway of *E. coli* can be found in **Figure 1.1**.

The DNA mismatch repair pathway of eukaryotes is similar to that of *E. coli*, but the details of the pathway are less certain [17-19]. In eukaryotes, DNA mismatch repair is initiated by one of two heterodimeric MutS homologs, depending on the substrate for excision. MSH2/MSH6, like MutS, binds base-base mismatches and 1-4 nucleotide extra-helical loops, and MSH2/MSH3 binds extra-helical loops that are ≥ 1 nucleotide. Thus, MSH2/MSH6 is primarily responsible for detecting base-base mismatches, MSH2/MSH3 is primarily responsible for detecting long extra-helical loops, and both heterodimers share the responsibility of detecting 1-4 nucleotide extra-helical loops. Eukaryotes also have heterodimeric MutL homologs (MLH1-PMS2 and MLH1-MLH3), whose functions are largely unknown. Eukaryotes have no known homolog of *E. coli* MutH, so the origin of the entry point for strand excision and the mechanism of strand discrimination are not known. However, it has been hypothesized that strand discontinuities associated with replication may aid in strand discrimination and serve as the entry point for strand excision in eukaryotes. Also, no eukaryotic DNA helicase has been shown to participate in the repair of replication errors, so the proteins associated with strand separation are not known. As with the *E. coli* DNA mismatch repair system, strand excision is carried out by 3' or 5' exonucleases, and DNA re-synthesis is carried out by DNA polymerase δ. A table summarizing the names and functions of *E. coli* and eukaryotic DNA mismatch repair proteins is shown in **Table 1.1**.

1.3 Introduction to DNA mismatch recognition

The focus of this work is to understand details involved in the first step of the DNA mismatch repair system, DNA mismatch recognition. As mentioned in section 1.2, the DNA mismatch repair system is initiated by binding of homodimeric MutS proteins in *E. coli* or the heterodimeric MutS homologs MSH2/MSH6 or MSH2/MSH3 in eukaryotes. Structural studies of *E. coli* and *Taq* MutS [20-22], human MSH2/MSH6 [23], and yeast MSH2/MSH3 [24] show that mismatch recognition proteins specifically bind to base-base mismatches and extra-helical loops in a mismatch binding domain, and specific contacts between the protein and the mismatch cause the DNA to be bent at the site of the lesion. These structural studies have led to the notion that specific mismatch recognition occurs at the instant the mismatch recognition complex binds and bends the DNA, so it has been hypothesized that the local flexibility of a mismatch may be responsible for mismatch recognition and conformational changes confer specificity to the protein-DNA interaction [25].

In addition to their mismatch binding domain, each subunit of the MutS homodimer [20-22, 26-28] or the MSH2/MSH6 [23, 29] or MSH2/MSH3 [23, 30] heterodimer contains an ATPase domain that binds and hydrolyzes ATP, and the coordination of ATP binding and hydrolysis in each of the two subunits of the mismatch recognition complexes is essential to recruit and activate downstream repair factors [17-19]. The function of ATPase activity in mismatch recognition complexes has been the subject of considerable study [31-42], and two classes of mechanisms has emerged. One class of mechanisms is based on the observation that in the presence of ATP, mismatch recognition complexes can form sliding clamps which leave a mismatch by movement along the helix [37, 40, 43-46]. This movement is postulated to link mismatch recognition to activation of downstream events at the site that directs excision. Two sub-classes of this mechanism have been proposed to explain ATP-dependent movement of mismatch recognition complexes along the DNA duplex. One model proposes that movement depends on ATP hydrolysis by the DNA-bound protein [34, 36, 37, 44, 46]. The second model postulates that ATP binding by a mismatch-bound protein results in the formation of a sliding clamp that can freely diffuse along the helix, so the ATP binding acts as a molecular switch [39, 43, 45, 47]. The second class of mechanisms is based on the idea that MutS translocation to the site of a nick may be blocked by other DNA binding proteins, so intrastrand looping may be a more viable option for activation of the endonuclease activity of MutH [33, 48].

The importance of studying the molecular mechanisms by which DNA mismatch recognition complexes recognize mismatches and signal for repair is highlighted by associations between faulty DNA mismatch recognition and genetic mutations. More specifically, two classes of genetic mutations are associated with the failure of mismatch recognition complexes to initiate repair. In class one mutations, mutations in genes encoding mismatch recognition complexes or inactivation in genes encoding these proteins can lead to spontaneous mutations and a predisposition to cancer [17-19]. In class two mutations, mismatch recognition proteins participate in DNA transactions that destabilize genetic information which is associated with a number of hereditary neurological diseases [30]. As an example of a class two mutation, the frequency of large expansions of CAG repeats—which is the detrimental mutation in the genesis of Huntington's disease—is decreased in mice deficient in MSH2 and MSH3, indicating that mismatch recognition complexes actively contribute to the formation of large expansions [30]. It has been suggested that repeating sequences may slip out of the duplex during DNA replication to form secondary structures which MSH2/MSH3 cannot properly process [29], but this conjecture has not been supported by experimental data.

In this work, the fundamental aspects of DNA mismatch recognition proteins in complex with well-repaired and poorly-repaired non-canonical DNA substrates are investigated using single-molecule FRET. This approach allows us to compare productive protein-DNA interactions to non-productive protein-DNA interactions without population averaging, and the results of these studies have important implications for class one and class two mutations, respectively. In these studies, the following questions are asked: (1) Does the DNA substrate itself have conformational fluctuations—if so, how does this affect the protein-DNA interactions? (2) How is specific mismatch recognition linked to DNA bending (are binding and bending concerted or separate events)? (3) Does DNA bending affect the enzymatic activity of mismatch recognition complexes? A summary of the constituents of this work can be found in the following section.

1.4 Summary of *Single-molecule studies of DNA mismatch recognition*

In Chapter 2, *Single-molecule FRET studies*, the physical principles of fluorescence and FRET are discussed. Then, details describing how to detect single fluorophores are given. Finally, particulars about designing single-molecule FRET experiments, acquiring data, analyzing data, and the power of the single-molecule FRET approach are detailed.

In Chapter 3, *Discrete conformational dynamics in three-way DNA junctions*, single-molecule FRET is used to show that DNA substrates with hairpins formed from CAG and CTG repeats—which can be bound by human MSH2/MSH3—have intrinsic conformational fluctuations between discrete states. Further investigation into the dynamic nature of these three-way DNA structures reveals that the movement of these substrates originates from base pairing rearrangements at the branch point that are driven by the base-base mispairing of the hairpin and the tension imparted by the hairpin loop. Previous studies of three-way DNA junctions using ensemble techniques have noted that these DNA structures are conformationally flexible, but the conformers have a wide energy landscape of available states. Proteins that bind to a conformationally flexible molecule would be largely nonspecific from a conformational perspective, since the protein-DNA interactions would occur at a large range of DNA conformations. In contrast, this study looks at three-way DNA junctions at the single-molecule level and demonstrates that these non-canonical DNA structures fluctuate between discrete conformers, limiting the energy landscape and thus the available binding sites for proteins and conferring specificity for the protein-DNA interactions.

In Chapter 4, *Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops*, single-molecule FRET is used to observe binding and dissociation of MSH2/MSH3 proteins to and from extra-helical loop DNA substrates in real time. It is shown that MSH2/MSH3 binds/bends all extra-helical loops in a concerted manner with low nanomolar affinity. After binding, ATP-bound MSH2/MSH3 dissociates quickly from repair-competent loops, while repair-resistant loops (formed from CAG repeats) trap nucleotide-bound MSH2/MSH3, inhibiting its dissociation from the DNA. Since ATP-stimulated dissociation from mismatches is proposed to be an important mechanism for downstream signaling, we propose that the specific interactions between MSH2/MSH3 and the conformation of the extra-helical loop—not just the flexibility of the mismatch itself—governs whether MSH2/MSH3 can bind/hydrolyze ATP in order to move away from the mismatch and signal repair or whether the enzymatic activity of MSH2/MSH3 becomes stalled and the protein stabilizes the extra-helical loop which becomes a precursor for mutation.

In Chapter 5, *The binding kinetics of E. coli MutS influences DNA mismatch selectivity*, single-molecule FRET is used to detect binding and dissociation of *E. coli* MutS to and from DNA substrates in real time. MutS binds/bends both G/T and C/T mismatched DNA substrates in a concerted manner, but in the absence of ATP, the binding stability of MutS for the efficiently repaired G/T mismatch is much greater than the binding stability of MutS for a less efficiently repaired C/T mismatch. Experiments in the presence of ATP demonstrate that stable MutS-mismatch binding is required to allow ATP uptake which causes MutS to form the sliding clamp that can interact with downstream repair proteins. The significance of this study is the observation that MutS bends the efficiently repaired G/T mismatch and the less efficiently repaired C/T mismatch with similar angles, but the binding stability of the protein-DNA interactions is important for the formation of the ATP-bound form of MutS that is essential for downstream signaling.

In Chapter 6, *Conclusions*, the findings of this study are summarized and put into perspective as to what they reveal about specific DNA mismatch recognition, and—more generally—protein-DNA interactions. In addition, the value of using single-molecule FRET for studies of protein-DNA interactions is discussed.

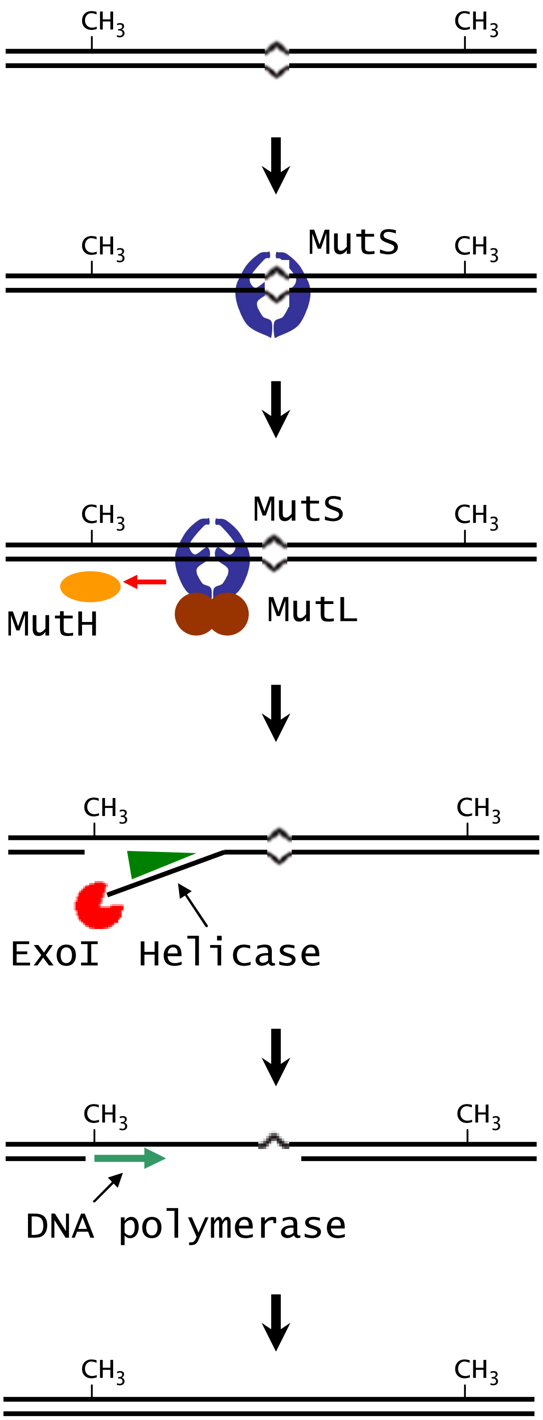


Figure 1.1
A schematic of the DNA mismatch repair pathway of *E. coli* is shown. Mismatches are generated during DNA synthesis, and the mismatches are initially recognized by a homodimer named MutS. MutS interacts with homodimeric MutL in an ATP-dependent process, and the MutS-MutL complex activates the endonuclease MutH, which nicks the daughter strand. A helicase unwinds the DNA, and an exonuclease removes the mismatch-containing strand. A DNA polymerase re-synthesizes the strand in the presence of single-strand DNA binding protein, and a DNA ligase seals the nick. This pathway results in a repaired duplex.

***E. coli* protein**	**Function**	**Homologs**	**Function**
MutS	Binds mismatches and 1-4 nucleotide loops	MSH2/MSH6	Binds single base-base and 1-4 nucleotide loops
		MSH2/MSH3	Binds some single-base loops and loops with ≥ 1 bases
MutL	Matchmaker that coordinates multiple steps in MMR	MLH1-PMS2	Matchmaker for coordinating multiple events in MMR
		MLH1-MLH3	Function of human heterodimer unknown
MutH	Nicks nascent unmethylated strand at hemimethylated GATC sites	None	
Helicase II	Unwinds DNA to allow excision of single-stranded DNA	None	
ExoI, ExoX, and ExoVII	Perform 3' to 5' excision of single-stranded DNA	EXOI	excision of double-stranded DNA
RecJ and ExoVII	Perform 5' to 3' excision of single-stranded DNA	3' exo of Pol δ and 3' exo of Pol ε	excision of ssDNA
DNA pol III	Accurate resynthesis of DNA	DNA pol δ	Accurate resynthesis of DNA
SSB	Participates in excision and DNA synthesis	RPA	Participates in excision and DNA synthesis
DNA ligase	Seals nicks after completion of DNA synthesis	DNA ligase	Seals nicks after completion of DNA synthesis

Table 1.1
The identities and functions of *E. coli* DNA mismatch repair proteins and their corresponding eukaryotic homologs are shown.

Chapter 2

Single-molecule FRET studies

Single-molecule fluorescence resonance energy transfer (FRET) is a powerful technique for studying the conformational dynamics of nucleic acids, proteins, and protein-nucleic acid systems. This chapter begins with a description of the physical principles of fluorescence and FRET. Then, details describing how to detect fluorescence and FRET in individual biomolecules are given. Finally, fine points about designing, carrying out, and analyzing single-molecule FRET experiments are provided.

2.1 Principle of fluorescence

Fluorescence is a physical phenomenon in which energy supplied by electromagnetic radiation is absorbed by a molecule and emitted as electromagnetic radiation [49, 50]. A molecule that is capable of undergoing the process of fluorescence is called a fluorophore. In its ground state, a fluorophore is in a relatively low energy, stable configuration, and it does not fluoresce [49, 50]. When light energy from an external source irradiates a fluorophore, the fluorophore can absorb the light energy, and if the absorbed energy is sufficient, the fluorophore reaches a higher energy state called the excited state. This process is known as excitation [49, 50]. Since the fluorophore is very unstable at high energy states, it then transitions to the lowest level excited state in a process called relaxation [49, 50]. Next, as the fluorophore transitions from the lowest level excited state back to the ground state, a photon is released in a process called emission [49, 50]. Because energy is lost in the relaxation step and other non-radiative processes during the fluorophores' excited state lifetime, the emitted light is of lower energy and thus longer wavelength than the absorbed light. From the ground state, the fluorophore can absorb light again and go through the entire process repeatedly [49, 50] (**Figure 2.1**).

A fluorophore absorbs light over a range of wavelengths, and each fluorophore has a characteristic excitation range. The wavelength that most effectively excites a fluorophore is called the excitation maximum. In addition, a fluorophore emits light over a range of wavelengths, and the wavelength at which a fluorophore emits light most intensely is called the emission maximum. Because energy is lost between the time the light is absorbed and emitted, a fluorophore emits lower energy (and thus longer wavelength) light than the light it absorbed. Since the excitation and emission wavelengths are different, the absorbed and emitted light is detectable as different colors on the visible spectrum (**Figure 2.2**).

2.2 Fluorescence resonance energy transfer (FRET)

Fluorescence resonance energy transfer (FRET) is a physical process that can be used to detect distance changes in the 20-80 Å range. In FRET, excitation energy from one (donor) fluorophore is transferred non-radiatively to another (acceptor) fluorophore via interaction between two induced dipoles (**Figure 2.3A**). The efficiency of energy transfer from the donor to the acceptor is inversely related to the distance between the dyes as described by the following relation:

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}$$

where R_0 is the distance at which $E = 0.5$ [51] and is ~60 Å for the classical FRET pair [52] (Cy3 and Cy5) conjugated to DNA [52]. Thus, when donor and acceptor are close together, the efficiency of energy transfer is high, and when the donor and acceptor are far apart, the efficiency of energy transfer is low (**Figure 2.3B**).

2.3 Single-molecule FRET

On the ensemble level, FRET can be used to detect the presence of conformational changes in biological molecules, but distributions of individual conformers and short-lived conformers are lost in the ensemble average. Single-molecule FRET measurements can readily determine the distribution of several conformations, not just the average of the conformations, making it possible to directly identify rarely-visited states and short-lived states [4-6]. Single-molecule FRET was first demonstrated in solution in 1996 [53]; since that time, methodological improvements have been made to optimize fluorescence microscopes for single fluorophore detection [5], improve fluorophore performance [54, 55], minimize surface effects [16, 56], increase data acquisition rate [5], and streamline data analysis procedures [3], and this technique has been adopted in many laboratories throughout the world to study conformational dynamics in nucleic acids, proteins, and macromolecular complexes [4].

In comparison to other single-molecule techniques such as single-molecule fluorescence localization (FIONA) [57, 58] and single molecule force methods (optical and magnetic tweezers) [59, 60], single-molecule FRET is less prone to environmental noise, as it is inherently a ratiometric technique (FRET between individual FRET pairs is determined by measuring the intensities of the donor and acceptor fluorophores—see section **2.8b**) and reports on the internal movements of the fluorophores in their center of mass frame. Therefore, small variations in excitation and emission intensities are tolerable, and relative drift of the molecule in the lab frame is much less of an issue. In addition, in single-molecule FRET, it is relatively easy to acquire data from several hundred molecules compared to optical and magnetic tweezers, so kinetic rates of biological events can be determined with the highest accuracy by screening out the intrinsic heterogeneity between single molecules with a large statistical distribution of events.

2.4 Single-molecule fluorescence microscopy

To detect single fluorescent molecules, two main components are required: (1) The background fluorescence must be minimized, and (2) the photon detection device should be as sensitive as possible, eliminating electronic noise. Prism-type total internal reflection (TIR) microscopy has been a major imaging tool for single-molecule FRET since it was first introduced [61]. In prism-type TIR microscopy, an inverted microscope is adapted to hold a fused silica prism on top of the sample chamber, and fluorescence is collected from an objective below the sample chamber. The laser beam (532 nm) used for excitation is focused by a lens onto the prism, passes through the prism and index-matching oil between the prism and a quartz[1] slide surface, and the beam is internally reflected at the quartz slide-water interface in the sample chamber. The fluorescence signal is collected using a long working distance water immersion objective (60x, 1.2 numerical apperature (N.A.)). Since only molecules within ~100 nm of the quartz slide surface are excited by the evanescent field, molecules not specifically attached to the slide surface are not excited, and the fluorescence background is low. After passing through the objective, the donor and acceptor fluorescence emission signals are collimated with a lens, spatially separated with a system of dichroic mirrors, and focused onto the screen of an EMCCD camera with a second lens. By exciting a large area (~0.05 mm^2 in size) and using EMCCD camera based detection, hundreds of individual molecules are imaged in parallel. Details on constructing and aligning the prism-type TIR microscope set-up can be found in **Appendix A**.

---

[1] It is necessary to use relatively expensive (but reusable) quartz slides to reduce light scattering and thus fluorescence background in this technique.

2.5 Single-molecule FRET experimental design

Single-molecule FRET experiments require careful experimental design to ensure the data is of high quality and reliability, as non-trivial decisions must be made about dye labeling, construct design, and imaging buffer constituents. This section discusses each of these design points in detail and outlines control experiments that should be done to ensure the smFRET observations do not contain photophysical artifacts.

a. Dye choice

Dyes for single-molecule fluorescence studies must be photostable (able to emit millions of photons before photobleaching and have little emission intensity fluctuations in the timescale of the biological events under study), be bright (have high extinction coefficients and quantum yields), be relatively small (so that they introduce minimal perturbation to the host molecule), be commercially available in a form that can be conjugated to biological molecules, and have excitation and emission maxima in visible wavelengths [4, 6]. A donor and acceptor pair (or FRET pair) for single-molecule FRET experiments should also have overlap between donor emission and acceptor absorption (required for energy transfer), large spectral separation between donor and acceptor emission (to reduce the amount of direct excitation of the acceptor by the laser and to reduce crosstalk between the donor and acceptor emission intensities), and comparable quantum yields for the donor and acceptor (to guarantee clearly anti-correlated intensity changes of the donor and acceptor in the single-molecule intensity trajectories) [4, 6]. Cy3 (donor) and Cy5 (acceptor) are the most popular FRET pair used in single-molecule FRET experiments because they are both photostable in an oxygen-free environment, they have similar quantum yields (~0.25), their spectral separation is large (~100 nm), and they are commercially available in a number of reactive forms. Alexa dyes—which have similar spectral properties to Cy3 and Cy5—have also been used in single-molecule FRET experiments, and in some biological systems of study in our lab Alexa647 is the preferred acceptor. Generally, though, Alexa dyes photobleach faster than Cy dyes, so Cy dyes allow longer observation times. TAMRA (tetramethylrhodamine) has similar spectral properties to Cy3, and has been used as a donor in single-molecule FRET experiments in the literature [62], but it has a lower extinction coefficient than Cy3 and has a tendency to change its emission intensity between two to three different levels [63], which can result in apparent changes in the FRET efficiency when no physical distance changes have occurred. Thus, TAMRA is not recommended for use in FRET experiments [63]. All single-molecule FRET experiments presented in this work are performed with Cy3- and Cy5- conjugated molecules.

b. Construct design

In single-molecule FRET experiments, donor, acceptor, and biotin molecules must be placed on specific locations of biological molecules. For nucleic acid only studies, the DNA or RNA molecule under study must contain the donor, acceptor, and biotin molecules, and these oligonucleotides can be readily obtained from several companies that sell custom-made DNA and RNA oligonucleotides that are conjugated to fluorescent dyes and biotin with ~100% labeling efficiency[2]. Likewise, for studies of protein-nucleic acid interactions, we prefer, whenever possible, to design experiments so that the nucleic acid contains the donor, acceptor, and biotin molecules to eliminate the need to label proteins. Because all the single-molecule FRET studies presented in this work have donor, acceptor, and biotin molecules attached to the DNA substrates, the remainder of this section will discuss details concerning the design of fluorescent DNA vectors.

Attaching the donor, acceptor, and biotin molecules at optimal locations is crucial to the success of single molecule FRET experiments. FRET is most sensitive (the relative FRET change per distance change is the largest) when the donor and acceptor are about R_0 apart (R_0 is estimated to be ~ 60 Å for Cy3 and Cy5 conjugated to double-stranded DNA), so the most sensitive change in distance measurements will be obtained when the donor and acceptor dyes are ~45-75 Å apart during a conformational change. When the dyes are too close together (<20 Å) or too far apart (>80 Å) before and after a conformational change, changes in conformation may not result in a FRET change that is measurable [4].

---

[2] Single-molecule FRET is relatively insensitive to incomplete labeling of a host molecule. If the biotin molecule is missing, the molecule is not tethered to the slide surface, and the molecule is not observed. If the donor is missing, the molecule is not observed because the acceptor is not directly excited. If the acceptor is missing, the donor-only species shows up as a zero-FRET population.

Although **Section 2.2** established that the FRET efficiency for a donor/acceptor FRET pair is inversely related to the distance between the fluorophores (**Figure 2.3B**), absolute distance measurements of fluorophore-conjugated nucleic acids contain experimental error because of the uncertainty in R_0 [64]. The curve in **Figure 2.3B** shows the relationship between the FRET efficiency and the distance R when $R_0 \sim 60$ Å. However, this curve is only an approximation of the actual distance R, as R_0 is difficult to determine precisely. Variables that affect R_0 can contribute to errors in distance measurements via FRET. R_0 depends on spectroscopic properties of the fluorophores and the medium, given by

$$R_0^6 = \frac{0.529 \cdot \kappa^2 \cdot \Phi_D \cdot J(\lambda)}{N \cdot n^4}$$

where κ^2 describes the relative orientation of the fluorophores, Φ_D is the quantum yield of the donor, $J(\lambda)$ is the spectral overlap integral, N is the Avogadro number, and n is the index of refraction of the medium [64]. The largest potential uncertainty in R_0 comes from the orientation term, κ^2, which can take values between 0 and 4 [49]. If the fluorophores undergo isotropic reorientation at a time shorter than their excited state lifetimes, then $\kappa^2 = 2/3$ [49]. However, if either dye is fixed or not free to rotate, then $\kappa^2 \neq 2/3$ [49]. For instance, fluorophores that are attached to double-stranded nucleic acids by a single covalent bond (i.e. the studies in this work) can stack onto the end of double-stranded DNA [65], and the stacking limits the fluorophores' orientational freedom [64]. The effect of fluorophore stacking on κ^2 has been studied in great detail [64], and it has been determined that the assumption that $\kappa^2 = 2/3$ could result in an error up to 12 Å in distance for the Cy3/Cy5 FRET pair commonly used in single-molecule FRET studies [64]. Therefore, unless dye orientation effects are taken into account, measuring the FRET efficiency between a donor and an acceptor is not an absolute measure of the distance between two fluorophores. However, for many situations (i.e. the studies in this work), a simple inverse-distance interpretation of FRET efficiency provides an adequate qualitative interpretation of FRET efficiency because the number of FRET states and the transitions between the FRET states are the critical observables for analysis of the number of conformers and the frequency of conformational fluctuations, respectively.

c. Imaging buffer

The cyclical process of fluorescence—of a fluorophore absorbing and emitting a photon—occurs until an irreversible chemical reaction occurs that changes the structure of the fluorophore. This process is known as photobleaching. In addition, conversion from a fluorophore's excited state to a triplet-state can interrupt fluorescence emission, causing the fluorophore to (reversibly) stop emitting photons in a process known as photo-blinking. In single-molecule FRET, a fluorophore should ideally last long (slow photobleaching) and not show temporal fluctuations of fluorescence intensity, but the photo-lifetime (time before photobleaching) and photostability of fluorophores is affected significantly by the constituents of the buffer solution. Molecular oxygen (O_2) is primarily responsible for photobleaching by photo-oxidation, but it is also an efficient triplet-state quencher. To prevent photobleaching, a method to remove O_2 from solution is required. However, once O_2 is removed from solution, transitions to the triplet-state will cause photo-blinking, so an alternate triplet-state quencher must be incorporated into the buffer solution.

To remove O_2 from solution and prevent photo-blinking, we combine an enzymatic oxygen scavenging system with a reducing agent, and this method has been shown to effectively extend the observation time of Cy fluorophores while preventing photo-blinking [55]. The oxygen scavenging system [66], which is composed of a mixture of glucose oxidase and catalase, converts glucose and O_2 into gluconic acid and water in a two step process that results in the net loss of O_2 in solution (**Figure 2.4**). To quench the triplet state, a saturated solution of TROLOX ® is incorporated into the imaging buffer solution [55]. Together, the oxygen scavenging system and the TROLOX ® solution work to increase the observation time of the average fluorophore to ~1 minute (at 32 ms time resolution) while suppressing photo-blinking. A protocol for preparing the imaging buffer can be found in **Appendix A**.

2.6 Surface preparation

smFRET experiments require surface immobilization of dye-labeled biomolecules so that the conformational changes of individual molecules may be observed over extended periods of time. Molecules are immobilized specifically to the surface of sample chambers that are constructed from quartz slides[3] and glass coverslips, and the chambers feature an inlet and an outlet hole for buffer exchange (**Figure 2.5A**). Sample chamber surfaces are prepared according to the study of interest.

a. BSA coated surface

For nucleic acid only studies, assembled sample chamber surfaces are coated with biotinylated BSA, and nucleic acids are immobilized via biotin-streptavidin bonds (**Figure 2.5B**). Nucleic acids can be immobilized specifically to BSA-biotin coated slides because quartz, BSA, and streptavidin are all negatively charged in neutral pH, so the negatively charged nucleic acids are repelled from the surface and do not have non-specific surface adsorption. A detailed protocol for preparing BSA-coated sample chambers is outlined in **Appendix A**.

b. Polyethylene glycol (PEG) coated surface

For studies involving proteins, BSA-coated surfaces are insufficient because the slide surface is too adhesive to many proteins [16]. To minimize protein interactions with the slide surface, quartz slides are coated with polyethylene glycol (PEG). A small fraction (~1%) of the PEG molecules are conjugated to biotin molecules so that DNA molecules can be immobilized to the slide surface specifically by biotin-streptavidin bonds. The PEG-coated surface reduces non-specific protein adsorption to an undetectable level (**Figure 2.5C**). Details for constructing PEG-coated surfaces are outlined in **Appendix A**.

---

[3] Quartz slides are used in single-molecule fluorescence microscopy because their uniform structure minimizes light scattering (compared with glass slides).

2.7 Immobilizing single molecules

Once the sample chamber surface is prepared, the slide surface should be imaged with imaging buffer before immobilizing nucleic acids to the surface. The ‘blank’ surface should have few fluorescent spots (**Figure 2.6A**) (any background fluorescence on the slide should be distinctly dimmer at the same excitation than surface-immobilized fluorophores). If this is not the case, an experiment should not proceed. The background fluorescence from slide contamination will interfere with the quality of the single molecule measurement, so the source of fluorescent contamination should be identified and corrected. If the slide is sufficiently clean, the slide surface should be tested for non-specific binding. To check for non-specific binding to the surface, donor-labeled DNA can be added to the slide and subsequently washed out, and the slide should be imaged with imaging buffer. No donor-labeled DNA molecules should be on the surface, since streptavidin is not present. After testing the chamber surface for background fluorescence and non-specific binding, specific single molecule immobilization is accomplished by adding streptavidin and then biotinylated, donor-labeled DNA/RNA to the sample chamber (see **Appendix A** for single molecule immobilization protocol). The surface density of the 25 μm × 50 μm imaging area should be ~200-300 fluorescent molecules (**Figure 2.6B**).

2.8 Imaging single molecules

The goal of taking single-molecule FRET data is to obtain time records of the relative distances between the donor and acceptor dyes. This information is then used to derive information on the average conformational populations at a given moment in time as well as to determine the transition rates for discrete conformational states of single molecules. In this section, details concerning data acquisition, processing, and analysis are presented.

a. Data acquisition and processing

In prism-type TIR microscopy, emission intensities of the donor and acceptor dyes are collected with an EMCCD camera, and the camera generates an electrical signal that encodes the information with the positions and intensities of the dyes (the image, **Figure 2.6B**). Then, the intensity data is saved in real time with software written in C++ with time resolution 30 ms or higher (specified by the user). The software records each frame of the movie in a single file that contains all the frames, where each pixel is encoded as a single byte. Each byte is 8 bits (i.e. 0-255), which represents how bright a signal is (the intensity). False colors are used to show the intensity in an effective manner. The image files generated are large—typically a 1 minute movie at 30 ms time resolution occupies 500 Mbytes of hard drive space. Intensity versus time information for individual donor/acceptor pairs (intensity time trajectories) are extracted from the recorded movie file using scripts written in IDL, and intensity time trajectories are used to calculate the FRET efficiency verses time (FRET efficiency time trajectories) (**Figure 2.7A**).

The IDL scripts locate the intensity peaks of the donor dyes by a peak-finding routine and use a calibration image to map the location of each corresponding acceptor. The calibration image is necessary because aberrations and imperfections in the alignment of the optical system slightly distort the donor and acceptor images. For a protocol on data acquisition, see Appendix A.

b. Calculating FRET efficiency

FRET efficiency is given by

$$E = I_A / (I_D + \gamma I_A)$$

where I_D is the intensity of the donor, I_A is the intensity of the acceptor, and γ is a parameter representing relative efficiencies and quantum yields of the donor and acceptor dyes and is determined from photobleaching events. For Cy3 and Cy5, it has been shown that $\gamma \approx 1$ [67].

c. FRET efficiency histograms

Construction and analysis of FRET efficiency histograms is the first step in the analysis of single molecule FRET data. FRET efficiency histograms are constructed by a program written in MATLAB that takes the average FRET efficiency of the first n frames of each molecule for thousands of molecules from multiple imaging areas (n is set to 10 for molecules with FRET states lasting seconds, but for molecules with changes in FRET that occur on the sub-second timescale, n is set to <10). A peak representing a single conformational state is usually less than 0.1 full-width-half-maximum wide, and statistical and instrumental noise inhibits further resolution of the peak. FRET histograms always have a population at $E = 0$ (which originate from molecules that have inactive acceptors—this peak is often called the 'donor only' peak) in addition to one or more populations at nonzero E values, depending on the number of conformational states of the system under study. When the area of the $E = 0$ ('donor only') peak is larger than the area of the other peak(s), the acceptor molecules are likely photobleaching too quickly, and the source of photobleaching needs to be addressed (**Figure 2.7B**).

d. Trajectory analysis

If the FRET efficiency time traces display temporal fluctuations, the data can be further analyzed to obtain more detailed information on the kinetics of the system using dwell-time analysis or hidden Markov modeling, depending on the type of fluctuation as described below. Regardless of the method of trajectory analysis, only molecules exhibiting single-step photobleaching are analyzed, as this is a unique signature of a FRET interaction between a single donor and single acceptor molecule (**Figure 2.7C**).

i. Dwell-time analysis

If there are two states, A and B, that are inter-converting on the timescale of tens of seconds, we can manually (via 'mouse clicking' using a program written in MATLAB) measure the dwell times of individual conformational states (or FRET states) from which we can determine average kinetic rates of the transitions.

ii. Hidden Markov Modeling

Hidden Markov modeling uses a step finding algorithm to automatically find states and measure times between changes in states in an unbiased way [3]. As with dwell-time analysis, this analysis method can be used for two-state systems with relatively slow dynamics (conformations lasting tens of seconds), but it is also useful for analyzing two-state systems with faster dynamics (conformations lasting a few seconds or less) and for systems with more than two conformational states. Hidden Markov modeling generates more reliable and reproducible transition rates, and is much less time-intensive than dwell-time analysis.

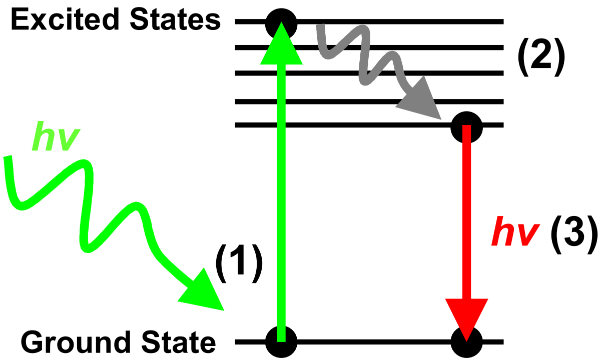


Figure 2.1
Transition state diagram (Jablonski diagram) for the fluorescence process. (1) Excitation of a fluorophore occurs through the absorption of light energy. (2) During the transient excited state lifetime, there is some loss of energy as the fluorophore relaxes to the lowest excited state. (3) Return of a fluorophore to its ground state accompanied by the emission of light. The light energy is always of a longer wavelength than the light energy absorbed (the fluorescence emission is shifted towards the red end of the visible spectrum) due to the energy lost during the transient excited state lifetime.

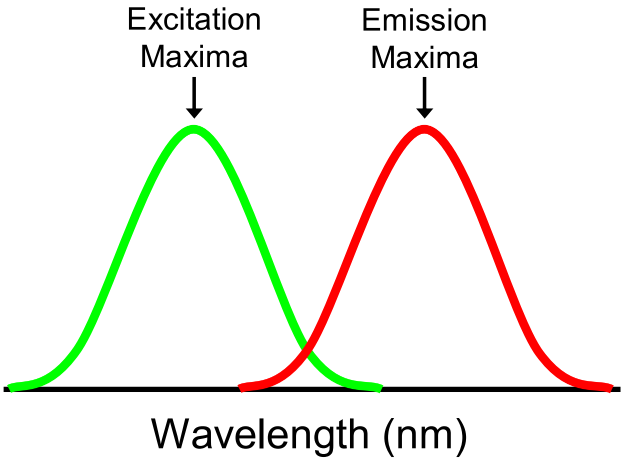


Figure 2.2
Generalized representation of the absorbance and emission spectra of a fluorophore.

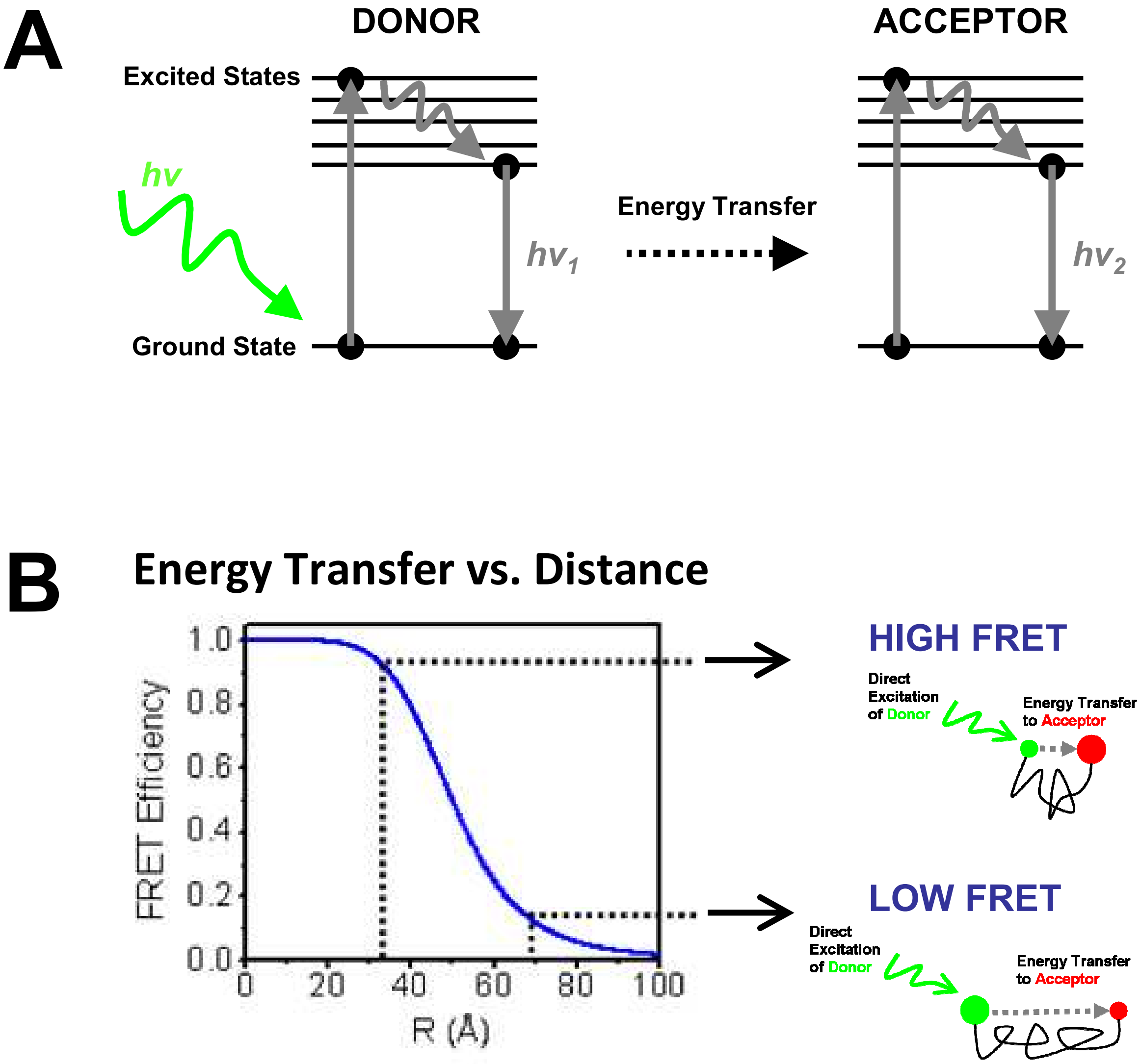

A
DONOR
ACCEPTOR
Excited States
hv
Ground State
hv1
Energy Transfer
hv2
B
Energy Transfer vs. Distance
FRET Efficiency
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
R (Å)
HIGH FRET
Direct Excitation of Donor
Energy Transfer to Acceptor
LOW FRET
Direct Excitation of Donor
Energy Transfer to Acceptor

Figure 2.3
(A) A donor fluorophore is directly excited by a light source (green line, hv), causing it to transition to an excited state. After relaxation, the fluorophore can either return to the ground state by emitting a photon (hv_1) or nonradiatively transfer energy to a nearby acceptor fluorophore. The nonradiative energy causes the acceptor fluorophore to reach an excited state, and after relaxation during the excited state lifetime, the fluorophore emits a photon (hv_2), allowing it to return to its ground state. (B) FRET efficiency is plotted as a function of distance *R* between the donor and acceptor (blue line). When the distance between the fluorophores attached to a biomolecule is small, the efficiency of energy transfer is high (high FRET), but when the distance between the fluorophores is large, the efficiency of energy transfer is low (low FRET) (green ball, donor; red ball, acceptor; black squiggle, biomolecule).

$$\text{glucose} + O_2 + H_2O \xrightarrow{\text{glucose oxidase}} \text{gluconic acid} + H_2O_2$$

$$H_2O_2 \xrightarrow{\text{catalase}} H_2O + \tfrac{1}{2}\, O_2$$

Figure 2.4
The oxygen scavenging system removes O_2 from the buffer in a two-step reaction.

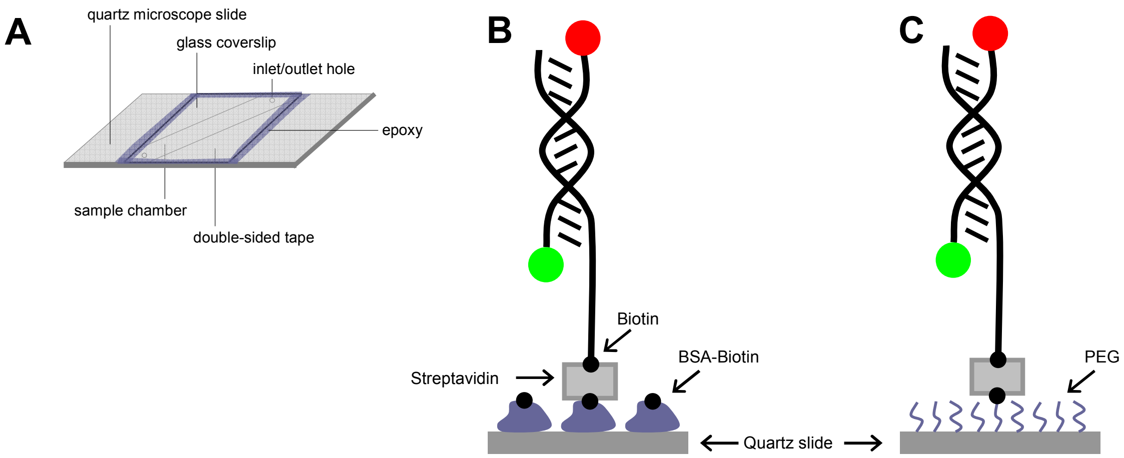

A
quartz microscope slide
glass coverslip
inlet/outlet hole
epoxy
sample chamber
double-sided tape
B
Biotin
BSA-Biotin
Streptavidin
C
PEG
Quartz slide

Figure 2.5
(A) A sample chamber is made by putting a quartz microscope slide and a glass coverslip together with double-sided tape and sealing the open ends with epoxy. Holes in the quartz slide are used for the inlet and outlet of solution exchange. For single-molecule FRET studies, surfaces were coated with (B) BSA-biotin or (C) 1% biotinylated polyethylene glycol (PEG) so that single DNA molecules (labeled with donor (green ball) and acceptor (red ball) molecules) could be immobilized via biotin-streptavidin bonds.

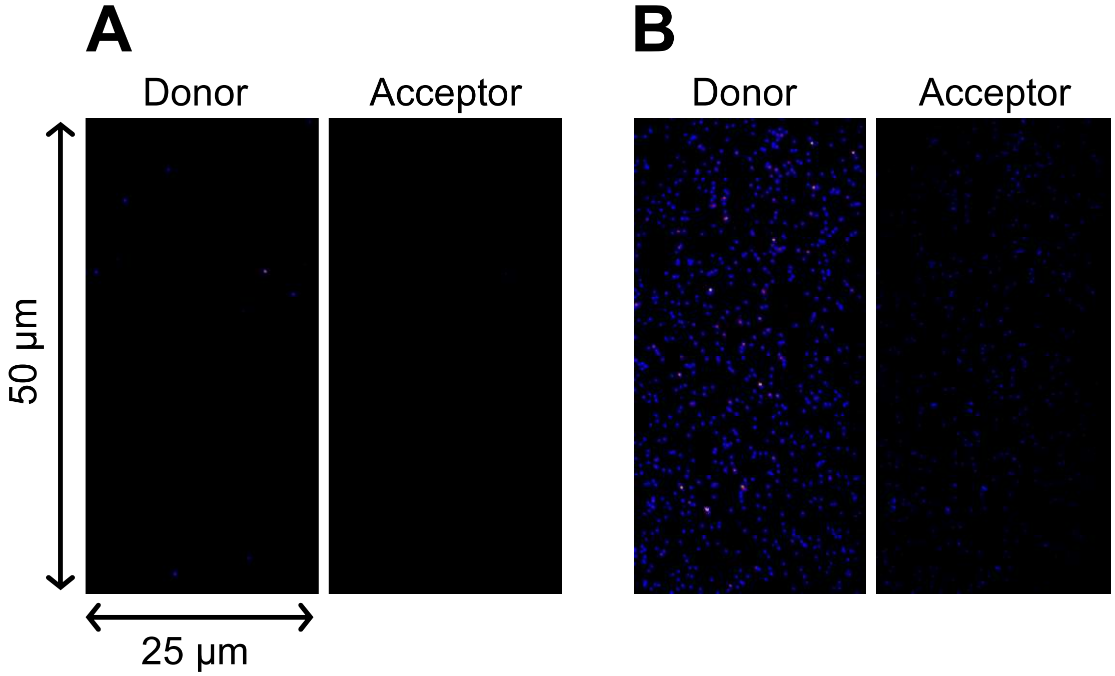


Figure 2.6
Single-molecule image. The image is split into donor and acceptor channels, each 25 µm × 50 µm. (A) The image from a blank sample chamber has few fluorescent spots. (B) After immobilizing DNA containing Cy3, Cy5, and biotin to the surface of a BSA-biotin or PEG coated slide, ~200-300 individual fluorescent spots of similar intensity are visible. The fluorescent intensities of the immobilized molecules are distinctly brighter than the debris molecules.

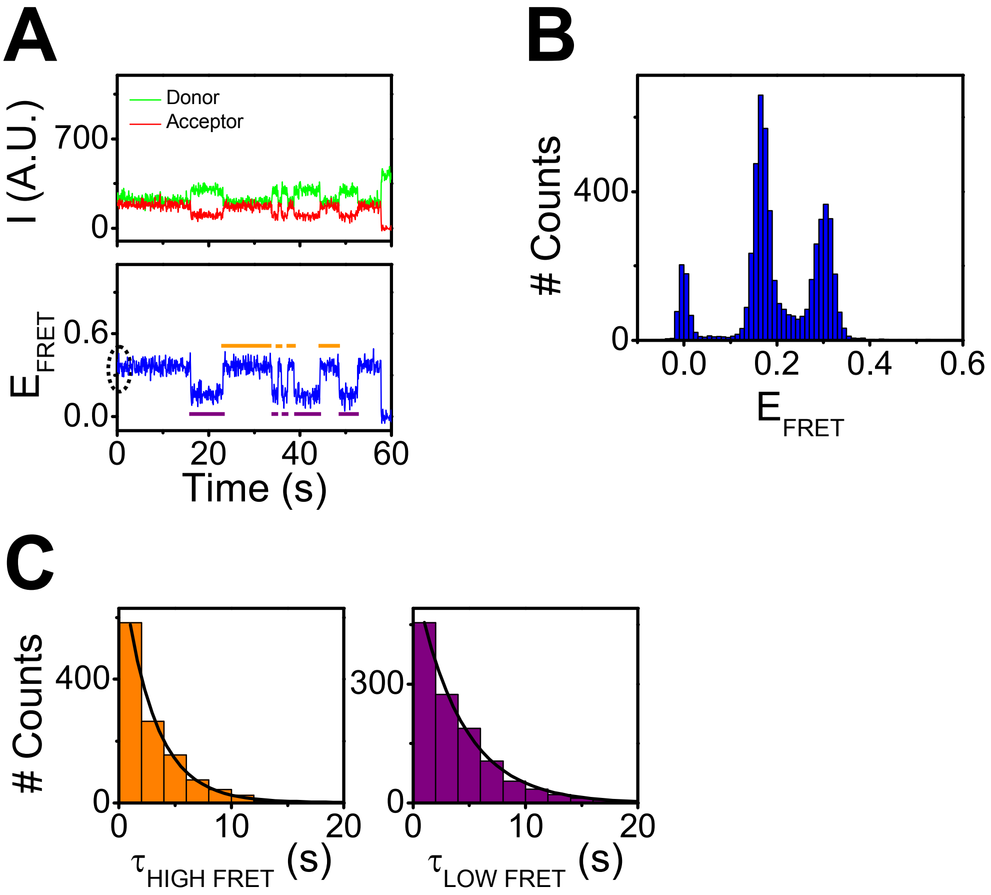


Figure 2.7
Example single-molecule FRET data. (A) Example single-molecule fluorescence intensity (I, in arbitrary units) time trace for donor (green) and acceptor (red) emission is shown. This intensity information was used to calculate the FRET efficiency (E_{FRET}) time trace (blue). Anti-correlated donor and acceptor intensity fluctuations correspond to changes in the calculated FRET efficiency and thus changes in the biomolecule's conformation. At ~60 s, single-step photobleaching of the acceptor molecule occurs (the emission intensity of the acceptor drops to zero and the emission intensity of the donor increases since energy is no longer being transferred to the acceptor), and the FRET efficiency drops to zero. The first n frames of the FRET efficiency traces (black dotted line in (A)) from thousands of single molecules are used to create (B) a FRET efficiency histogram, which reports on the populations of all observed molecules at a given point in time. The FRET efficiency histogram has a peak at $E_{FRET} = 0$ (which is the population of molecules that have donors but inactive acceptors), at $E_{FRET} = 0.17$, and at $E_{FRET} = 0.30$. Thus, the biomolecule in this data can exist in two discrete conformational states. Since the single-molecule FRET time trace in (A) shows that the biomolecule interconverts between the two conformational states rapidly, this data demonstrates that this biomolecule operates in a highly dynamic two-state manner. (C) Single-molecule trajectory analysis was used to determine the dwell times of high FRET ($E_{FRET} \sim 0.30$) and low FRET ($E_{FRET} \sim 0.17$) states for hundreds of FRET efficiency traces like that shown in (A) (in (A), orange dashes represent measured dwell times of high FRET states, while purple dashes represent measured dwell times of low FRET states), and the dwell times (τ) for high FRET (orange histogram) and low FRET states (purple histogram) were plotted. These histograms fit well to a single-exponential, meaning that the conformational change results from a single kinetic step.

Chapter 3

Discrete conformational dynamics in three-way DNA junctions[4]

In the study presented here, single-molecule FRET was used to show that DNA substrates with $(CAG)_{13}$ and $(CTG)_{13}$ hairpins (which are three-way DNA junctions that are recognized by DNA mismatch recognition proteins) are highly dynamic molecules, and the dynamics of these substrates originate from base pairing rearrangements at the branch point that are driven by the base-base mispairing and the tension imparted by the hairpin loop. Since the interaction of proteins with DNA molecules—and branch structures in particular—is highly dependent on the local conformation of the DNA substrate, the dynamic characteristics of these DNA substrates is likely to be important for understanding the interactions of DNA mismatch recognition proteins with these structures.

3.1 Background: Three-way DNA junctions

Three-way junctions are composed of three double helices (or branches) that intersect at a common branch point (**Figure 3.1**). Junctions with three helical branches are widespread in RNA structures [68], and three-way DNA junctions can form when complementary repeating sequences on a single-strand pair up and form thermodynamically stable hairpins [69-71]. In the latter case, the formation of 'slipped-strand DNA structures' is believed to contribute to the expansion of nucleotide repeat tracts which are associated with the development of many hereditary and anticipative neurodegenerative diseases [69-73]. The mechanistic details of nucleotide repeat expansions are unclear, but a detailed understanding of the geometry, stability, and dynamics of three-way DNA junctions at their branch points may illuminate this pathway.

---

[4] This work has been submitted for publication:
J.E. Coats, Y. Lin, and I. Rasnik. 'Discrete conformational dynamics in three-way DNA junctions.' *Nucleic Acids Research* (2012).

The basic conformation of three-way DNA junctions has been previously studied both as a model system to understand fundamental principles of nucleic acid folding and as a tool for the construction of synthetic DNA nano-machines. Perfect three-way DNA junctions with full base pairing (3H junctions) (**Figure 3.1A**) have been studied using comparative gel electrophoresis [74-76], chemical probing [74-76], ensemble fluorescence resonance energy transfer (FRET) [77], and atomic force microscopy (AFM) [69, 78]. In the absence of magnesium ions, 3H junctions possess trigonal pyramidal geometry that minimizes electrostatic interactions and is compatible with steric constraints at the point of branch exchange. When magnesium is present, some symmetry is lost, but the structures remain extended without the helix-helix stacking that occurs for four-way DNA junctions under these conditions [79]. The presence of additional formally unpaired bases at the branch point ($3HS_n$ junctions) (**Figure 3.1B**) increases the conformational flexibility and stability of the structures [80]. In the absence of divalent ions, these bulged junctions have an unstacked, extended conformation, but when ions are added, two of the arms become coaxially stacked in one of two isomeric configurations [80-91]. The particular isomer conformation adopted by $3HS_n$ junctions is highly sequence-dependent [85, 87].

In addition to conformation, the dynamical nature of three-way DNA junctions has been investigated. AFM studies of three-way DNA junctions with a triplet repeat sequence hairpin [92] and with inverted repeat sequences [78] have shown that the junctions adopt a trigonal pyramidal geometry with a wide range of inter-arm angles, suggesting the junctions are conformationally flexible. Junction flexibility has also been suggested by studies of enzymatic ligation of 3H junction structures [93] and ensemble time-resolved FRET experiments of $3HS_n$ junctions. Other studies have shown the presence of discrete conformational states, using single molecule FRET (with 3H junctions) [94, 95] and ensemble FRET (with $3HS_n$ junctions) [84].

In the study presented here, single-molecule FRET is used to study the conformational dynamics of three-way DNA junctions with a triplet repeat sequence on one branch. The experiments show that the presence of mispaired bases on one branch of a three-way junction has significant effects on the conformational dynamics of the junction. In addition, the sterical constraints imposed by the presence of the hairpin further contribute to the observed conformational dynamics. The results given here are consistent with the "dynamical junction model" proposed previously for 3H junctions [89] and have important implications for the prediction of interactions between proteins and these structures at the point of branch exchange.

3.2 Results

Single-molecule experiment design

For the single-molecule FRET experiments, three-way DNA vectors were constructed by annealing two single-stranded oligonucleotides to form three-way DNA junctions with two homoduplex branches (that were 9 and 10 bp in length) flanked by a hairpin branch formed from the partial self-pairing of thirteen triplet repeats of either CAG (cytosine-adenine-guanine) (**Figure 3.2A**) or CTG (cytosine-thymine-guanine) (**Figure 3.2B**). In these constructs, the hairpin branches contain base-base mismatches every three base pairs, for a total of six mismatches in each template. The homoduplex branches of both substrates were identical in sequence, so the only structural differences in the two DNA constructs is the base-base mispairing on the hairpin branch. The ends of the shorter strands were labeled with donor (Cy3, 5' end) and acceptor (Cy5, 3' end) fluorophores. For both templates, the longer strand at the 3' end contains a poly-dT extension and a biotin tag for immobilization on a streptavidin-coated surface for single-molecule observation. The extension was designed to prevent potential interactions of the fluorophores with the streptavidin surface. In these constructs, single-molecule FRET was used to monitor the relative changes in distance between the ends of the labeled (homoduplex) branches.

Individual junctions have intrinsic conformer transitions

Single-molecule donor and acceptor intensity time traces of the $(CAG)_{13}$ and $(CTG)_{13}$ hairpin substrates show rapid, anti-correlated fluctuations of donor and acceptor signals which correspond to rapid changes in the FRET efficiency of the junctions (**Figure 3.2C,D**). The fluctuations of the $(CAG)_{13}$ hairpin occur many times per second, so the time resolution of the microscope (32 ms) is a limiting factor in determining accurate rates of the transitions (**Figure 3.2C**). The $(CTG)_{13}$ hairpin, however, fluctuates between two discrete conformational states ($E_{FRET} \sim 0.22$ and 0.43), each conformational state lasts seconds, and the substrate displays almost no bias between the two conformational states (**Figure 3.2D**). Since the $(CAG)_{13}$ and $(CTG)_{13}$ hairpin junctions are identical except for the base-base mispairing in the hairpin branches, the difference in the conformations and dynamics of the two junctions must arise from the structural instabilities in each hairpin branch, however the nature of the structural dynamics in these junctions is not known. To investigate the origin of the conformational dynamics in DNA hairpin substrates, we carried out a systematic study of the $(CTG)_{13}$ hairpin junction because this junction has clearly resolved conformational transitions that are much longer than the time resolution of the experiment.

Influence of magnesium chloride and sodium chloride on $(CTG)_{13}$ hairpin conformer transitions

The conformations of branched DNA and RNA structures are highly dependent on the ionic strength and/or the presence of divalent ions in solution [68], so studies of these structures often probe the effect of sodium chloride and magnesium chloride on the conformations and dynamics of these structures [7, 68, 86, 87, 89, 96]. Increases in the ionic strength of the branched nucleic acid structures decreases the electrostatic repulsion between the branches, and divalent ions have specific interactions with nucleic acids that can facilitate coaxial stacking of helices [7, 68, 86, 87, 89, 96].

First, the dependence of the conformational states and transition rates of the $(CTG)_{13}$ hairpin on magnesium chloride ($MgCl_2$) concentration was determined in physiologically relevant sodium chloride conditions (110 mM NaCl). In the absence of $MgCl_2$, the FRET efficiency histograms of the $(CTG)_{13}$ hairpin substrate show a narrow distribution of FRET values centered at E_{FRET} ~ 0.26 (the peak at E_{FRET} ~ 0 corresponds to substrates with inactive acceptor fluorophores and will be ignored from this point on) (**Figure 3.3A, top**). The corresponding FRET efficiency time traces (**Figure 3.3B, top**) do not show changes in FRET efficiency over time, indicating that there are no conformational dynamics in the $(CTG)_{13}$ hairpin substrate when no $MgCl_2$ is present. Addition of $MgCl_2$ to the $(CTG)_{13}$ hairpin causes two discrete FRET states to appear in the FRET efficiency histograms and the corresponding FRET efficiency time traces (**Figure 3.3A,B, +$MgCl_2$**). The FRET efficiency histograms show that the FRET values of the two conformers are dependent on the concentration of $MgCl_2$. As the concentration of $MgCl_2$ increases, the FRET value of the 'low FRET' conformer decreases and the FRET value of the 'high FRET' conformer increases, but the population of the conformers remains ~1:1, regardless of $MgCl_2$ concentration (**Figure 3.3A,B, +$MgCl_2$**). Thus, the angle between the two conformers increases with increasing $MgCl_2$ concentration. The single-molecule FRET efficiency time traces reveal that the individual $(CTG)_{13}$ hairpin molecules undergo spontaneous transitions between two FRET states (so the two FRET efficiency populations observed in the FRET efficiency histograms are not from two different annealed species). In agreement with the FRET efficiency histograms, the FRET efficiency time traces show an increase in the amplitude of the FRET transition with increasing $MgCl_2$ (because the FRET value of the 'low FRET' conformer decreases and the FRET value of the 'high FRET' conformer increases), but the transition rates (determined by hidden Markov analysis) for the junctions in the presence of $MgCl_2$ are not affected by $MgCl_2$ concentration (the transition rate from the low FRET to the high FRET state ($k_{low\text{-}high}$) is ~2.6 s^{-1} at each $MgCl_2$

concentration, and the transition rate from the high FRET to the low FRET state ($k_{\text{high-low}}$) is ~2.4 s^{-1} at each $MgCl_2$ concentration) (**Figure 3.3C**). Thus, $MgCl_2$ ions are required for the dynamics of the $(CTG)_{13}$ hairpin substrate, but changing the $MgCl_2$ concentration does not alter the activation energy nor the relative stability of the two conformers.

Next, the dependence of the conformational states and transition rates of the $(CTG)_{13}$ hairpin on sodium chloride (NaCl) concentration was determined in physiologically relevant $MgCl_2$ conditions (1 mM $MgCl_2$). In all NaCl conditions tested (10 mM – 1 M), the FRET efficiency histograms show two discrete states that are similar in FRET efficiency value (**Figure 3.3D**), and the single-molecule FRET efficiency time traces show that the FRET values of the two conformers exchange rapidly (**Figure 3.3E**), with similar transition rates for the junctions in NaCl concentrations ranging from 110 mM – 1 M (**Figure 3.3F**).

Together, the experiments probing the effects of $MgCl_2$ and NaCl on the conformational dynamics of the $(CTG)_{13}$ hairpin substrate show that specific interactions by $MgCl_2$ stabilize the conformers of the junction. In particular, comparing the conformational dynamics of the junction in 0 mM $MgCl_2$ and 110 mM NaCl (**Figure 3.3A,B,C**; 0 mM $MgCl_2$)—which has no dynamics and has an ionic strength of $I = 0.17$—to the activity of the junction at 1 mM $MgCl_2$ and 10 mM NaCl (**Figure 3.3D,E,F**)—which has dynamics and has an ionic strength of $I = 0.013$—shows that a minimum ionic strength is not the driving force in the conformational dynamics. Thus, the conformers of the $(CTG)_{13}$ hairpin are formed by the specific interactions of divalent ions with the junction, but the number of divalent ions present in solution does not alter the transition rates among the conformers.

Influence of temperature on $(CTG)_{13}$ hairpin conformer transitions

Since changing the concentration of monovalent and divalent ions does not affect the transition rates between the $(CTG)_{13}$ hairpin conformers, we investigated the effect of solution temperature on the conformers' transition rates. Unlike 4H junctions, where the conformer transitions do not necessarily involve the disruption of base pairing at the branch point [96], 3H junctions are more sterically constrained, so any conformational rearrangement of a 3H junction requires breaking of basepairs near the branch point. Thus, the frequency of the conformational rearrangements of the $(CTG)_{13}$ hairpin substrate should increase with increasing temperature, since high temperatures promote melting of base pairs.

The influence of temperature on the conformational dynamics of the $(CTG)_{13}$ hairpin was studied by regulating the temperature of the sample chamber with a circulating water bath and monitoring the temperature of the sample with a thermocouple. In buffer containing 110 mM NaCl and 1 mM $MgCl_2$, changing the temperature between 15-35°C does not alter the relative populations of the low and high FRET states. The low FRET state does not change conformation with temperature (E_{FRET} ~0.20 for temperatures 15-35°C), and the high FRET state changes conformation slightly with temperature (E_{FRET} ~0.43 at 35°C, E_{FRET} ~0.38 at 15°C) (**Figure 3.4A**). As expected, both of the transition rates ($k_{low\text{-}high}$ and $k_{high\text{-}low}$) increase with increasing temperature, so changing the temperature changes the barrier between the conformational states but does not change the relative free energy of the conformers (**Figure 3.4B**). A global fit to an Arrhenius plot of the transition rates at each temperature (**Figure 3.4C**) reveals that the activation energy of the transition from low FRET to high FRET is 9.3 kcal/mol and the transition from high FRET to low FRET is 9.3 kcal/mol.

Influence of photophysical artifacts on the observed dynamics of the (CTG)$_{13}$ hairpin

To eliminate the possibility that the FRET populations in the (CTG)$_{13}$ hairpin data were photophysical artifacts of the local fluorophore environment, measurements were performed on a substrate with a 19 bp sequence that was identical to the homoduplex portion of the (CTG)$_{13}$ hairpin substrate (**Figure 3.5A**). The FRET efficiency histograms of the homoduplex substrate in both the absence and presence of $MgCl_2$ show a single narrow population of conformers at E_{FRET} ~ 0.17 (**Figure 3.5B**), and the FRET efficiency time traces indicate the absence of dynamics in both conditions (**Figure 3.5C**). Because the local environment of the fluorescent dyes is identical for the two substrates, the results show that the FRET states observed for the (CTG)$_{13}$ hairpin substrate must be associated with the conformational dynamics of the hairpin junction.

Conformational dynamics of perfectly paired hairpin

Since the (CAG)$_{13}$ and (CTG)$_{13}$ hairpin junctions had dynamics with different conformations and transition rates, we next investigated the effect of the base-base mismatches on the observed conformational transitions of the (CTG)$_{13}$ hairpin junction. To determine the influence of the T-T mismatches in the (CTG)$_{13}$ hairpin substrate on the intrinsic dynamics of the junction, the six T-T mismatches on the hairpin loop of the (CTG)$_{13}$ hairpin substrate were replaced with A-T base pairs to create a substrate named (CTG)$_7$(CAG)$_6$ hairpin (**Figure 3.6A**). In the absence of magnesium, FRET efficiency histograms of the $(CTG)_7(CAG)_6$ hairpin show a narrow population of conformers at E_{FRET} ~0.21 (**Figure 3.6B, -Mg**), and single-molecule FRET efficiency time traces show no conformational dynamics for this substrate in these conditions (**Figure 3.6C, -Mg**). In the presence of magnesium, two conformational populations appear, one at E_{FRET} ~0.18 and another at E_{FRET} ~0.38 (**Figure 3.6B, +Mg**). The single-molecule FRET efficiency time traces show that (like for the $(CTG)_{13}$ hairpin substrate) the two FRET efficiency populations correspond to transitions between two states for individual molecules with transition rates of 0.017 s^{-1} and 0.1 s^{-1} for the low to high ($k_{low\text{-}high}$) and high to low ($k_{high\text{-}low}$) transitions, respectively (**Figure 3.6C, +Mg**). Therefore, the $k_{low\text{-}high}$ rate of the $(CTG)_7(CAG)_6$ hairpin is ~100x lower than that of the $(CTG)_{13}$ hairpin substrate, and the $k_{high\text{-}low}$ rate of the $(CTG)_7(CAG)_6$ hairpin is ~20x lower than that of the $(CTG)_{13}$ hairpin substrate. Thus, the T-T mismatches in the $(CTG)_{13}$ hairpin substrate are not required for conformational fluctuations, but the base-base mismatches do affect the conformational state (the distances between the homoduplex arms), the relative population of high and low FRET conformers, and the frequency at which the substrate interconverts between conformers.

Effect of the hairpin loop on the dynamics of the $(CTG)_{13}$ hairpin

Next, to determine the influence of the hairpin loop on the dynamics of the $(CTG)_{13}$ hairpin, a substrate named $(CTG)_6(CG)(CTG)_6$ three-way was designed that was identical to the $(CTG)_{13}$ hairpin substrate except the hairpin loop was replaced by an open duplex end (**Figure 3.7A**). The FRET efficiency histogram of the $(CTG)_6(CG)(CTG)_6$ three-way substrate shows a single conformational population at E_{FRET} ~ 0.23 in the absence of magnesium (**Figure 3.7B, -Mg**), and the single-molecule FRET efficiency time traces show no conformational dynamics for this substrate in the absence of magnesium (**Figure 3.7B, -Mg**). In the presence of magnesium, the FRET efficiency histograms of the $(CTG)_6(CG)(CTG)_6$ three-way substrate show a major conformational population at E_{FRET} ~0.20, and there is a small shoulder to the right of the major FRET efficiency peak in the histogram (**Figure 3.7B, +Mg**). This single-molecule FRET efficiency time trace shows that there are transient fluctuations from the major conformational population (at E_{FRET} ~0.20) to a higher FRET conformation (**Figure 3.7C, +Mg**). The transition rate from the low FRET state to the high FRET state ($k_{low-high}$) was 0.22 s^{-1}, and the transition rate from the high FRET state to the low FRET state ($k_{high-low}$) was 3.8 s^{-1}. Since $k_{low-high}$ is similar for the $(CTG)_{13}$ hairpin and the $(CTG)_6(CG)(CTG)_6$ three-way substrates, the base-base mismatches must be primarily responsible for the transition from the high FRET state to the low FRET state. However, the $k_{high-low}$ for the $(CTG)_6(CG)(CTG)_6$ three-way substrate is an order of magnitude larger than the $k_{high-low}$ for the $(CTG)_{13}$ hairpin substrate, so the hairpin loop must play a significant role in stabilizing the high FRET conformation.

Effect of the hairpin loop on the dynamics of the $(CTG)_7(CAG)_6$ hairpin

To determine the influence of the hairpin loop on the dynamics of the $(CTG)_7(CAG)_6$ hairpin, a substrate named $(CTG)_6(CG)(CAG)_6$ three-way was designed that was identical to the $(CTG)_7(CAG)_6$ hairpin substrate except that the hairpin loop was replaced by an open duplex end (**Figure 3.8A**). The FRET efficiency histogram of the $(CTG)_6(CG)(CAG)_6$ three-way substrate shows a single conformational population at

E_{FRET} ~ 0.21 in the absence of magnesium and a single conformational population at E_{FRET} ~ 0.18 in the presence of magnesium (**Figure 3.8B**), and the FRET efficiency time traces look identical in the absence and presence of magnesium (**Figure 3.8C**). Thus, removing the hairpin from the $(CTG)_7(CAG)_6$ hairpin substrate eliminates the intrinsic conformational dynamics of this substrate, and the hairpin loop of the $(CTG)_7(CAG)_6$ hairpin must contribute to the observed conformational fluctuations.

Effect of the sequence of the homoduplex branches on the dynamics of $(CTG)_{13}$ hairpin

Previous studies of 3H junctions and $3HS_n$ junctions have demonstrated that the observed conformations and dynamics of these structures are dependent on the sequences near the branch point [74-76, 80, 86, 87, 89, 90], and such studies led to the proposal of a dynamical junction model to describe the dynamics of three-way DNA junctions [89]. In this model, each of the base pairs flanking the branch can break transiently, allowing the bases to associate with other unpaired bases. According to this model, repeated melting and re-annealing in different conformations would account for conformational heterogeneity in these types of junctions. If the conformational dynamics of the $(CTG)_{13}$ hairpin substrate can be described by the dynamical junction model, then we expect that changing the sequence of the structure relative to the hairpin will change the conformational dynamics of the substrate.

To determine the effect of the homoduplex branches' sequences on the observed dynamics of the $(CTG)_{13}$ hairpin (and thus test the validity of the dynamical junction model for this junction), a substrate named $(CTG)_{13}$ hairpin-B was designed. $(CTG)_{13}$ hairpin-B is identical to $(CTG)_{13}$ hairpin except that the homoduplex arms have been swapped (compare **Figure 3.9A** with **Figure 3.2A**), and this change results in a new sequence relative to the hairpin arm. In the absence of magnesium, the FRET efficiency histogram of the $(CTG)_{13}$ hairpin-B substrate has two peaks—one at E_{FRET} ~0.31 and another at E_{FRET} ~0.49 (**Figure 3.9B,** -Mg). Transitions between these two FRET states are seen in the corresponding FRET efficiency histograms (**Figure 3.9C,** -Mg) with rates of 4.0 s^{-1} and 1.5 s^{-1} for the low to high and high to low FRET transitions, respectively. In the presence of magnesium, the FRET efficiency of the two observed FRET states changes. The FRET efficiency histogram shows that the two conformational states of the junction are a E_{FRET} ~ 0.26 and E_{FRET} ~ 0.61 (**Figure 3.9B,** +Mg). The transition rates between these two FRET states were 2.4 s^{-1} and 0.2 s^{-1} for the low to high and high to low FRET transitions, respectively (**Figure 3.9C, +Mg**).

Similar to the $(CTG)_{13}$ hairpin junction, changing the temperature of the sample chamber holding the $(CTG)_{13}$ hairpin-B junction between 15°C and 35°C does not alter the change in the relative populations between the low and high FRET states (**Figure 3.9D**). The FRET efficiency value of the high FRET state shifted slightly from E_{FRET} ~0.61 to ~0.56 when the temperature was increased from 15° to 35°, but no shift was observed for the low FRET state as a function of temperature. Like for the $(CTG)_{13}$ hairpin substrate, increasing the temperature lowers the energy barrier between the two conformational states. A global fit to an Arrhenius plot of the transition rates at each temperature (**Figure 3.9F**) reveals that the activation energy of the transition from low FRET to high FRET is 11.5 kcal/mol and the transition from high FRET to low FRET is 11.5 kcal/mol.

3.3 Discussion

Three-way DNA junctions formed from the self-pairing of triplet repeat sequences are believed to be mutagenic intermediates in triplet repeat expansion, and their dynamic characteristics are likely to be important for the interactions of proteins with these structures. The interaction of proteins with DNA molecules—and branched structures in particular—is highly dependent on the local conformation of the DNA substrate. The conformation of three-way DNA junctions has been studied extensively for perfectly paired junctions as well as for junctions with formally unpaired bases at the junction, but junctions that contain hairpins with triplet repeat sequences have been studied to a lesser extent.

In the study presented here, single-molecule FRET was used to show that DNA substrates with $(CAG)_{13}$ and $(CTG)_{13}$ hairpins are highly dynamic molecules. Further investigation into the dynamic nature of the $(CTG)_{13}$ hairpin—which interconverts between two discrete states—reveals that $MgCl_2$ plays a defining role in the conformations that the $(CTG)_{13}$ hairpin junction assumes, and the rate of both the low to high FRET ($k_{low\text{-}high}$) and high to low FRET ($k_{high\text{-}low}$) transitions of the $(CTG)_{13}$ hairpin increase as the temperature of the buffer is raised.

The dependence of the transition rates of the $(CTG)_{13}$ hairpin on the buffer temperature suggest that the stability of the junction is a primary driving force for the dynamics of the junction. Indeed, the measured $(CTG)_{13}$ hairpin transitions are both ~10 kcal/mol, which is, remarkably, the amount of energy needed to destabilize the hairpin arm, as computed by UNAfold [97]. Thus, the experiments with the $(CTG)_{13}$ hairpin indicate that the hairpin arm plays a major role in driving the conformational dynamics of the junction.

To see how the hairpin arm affected the junction as a whole, we designed a series of perturbations of the $(CTG)_{13}$ hairpin substrate. Removing the T-T mismatches of the $(CTG)_{13}$ hairpin decreases the frequency of the conformational dynamics of the junction relative to the dynamics of the $(CTG)_{13}$ hairpin junction in identical buffer conditions, indicating that the T-T mismatches affect but are not solely responsible for the observed conformational dynamics. Replacing the end of the hairpin loop with a blunt duplex end greatly decreases the transition rate from the high to low ($k_{high\text{-}low}$) FRET, but the transition rate from low to high ($k_{low\text{-}high}$) FRET is similar to that of the $(CTG)_{13}$ hairpin. This latter result indicates that while the transition from low to high ($k_{low\text{-}high}$) FRET is primarily dependent on the T-T mismatches, the transition from the high to low ($k_{high\text{-}low}$) FRET is very unstable in the absence of a hairpin loop. Changing the T-T mismatches to A-T base pairs in the three-way junction with the open duplex end eliminates the conformational dynamics altogether.

Together, these results indicate that the conformational dynamics of the three-way DNA junctions arise from melting of base pairs near the branch point and subsequent rearrangement/re-annealing of the bases to a new conformation. Later on, melting of base pairs near the branch point of the new conformer allow the base pairs of the initial conformer to re-hybridize, so the conformational changes are reversible. A similar 'dynamical junction model' was proposed previously by Zhong et al. [89] to explain the existence of a manifold of interconverting conformers in perfectly paired three-way DNA junctions. Changing the sequence of the homoduplex arms and leaving the hairpin intact resulted in drastic changes in the observed conformational dynamics of the junction, which is consistent with a dynamical junction model in which different base-pairings will be favored depending on the overall sequence.

It should be noted that the results presented here do not provide detailed information on the precise conformational nature of the observed conformers, as the focus of this work was on understanding the underlying reasons for the conformational dynamics observed in the $(CAG)_{13}$ and $(CTG)_{13}$ hairpin substrates. Future studies with multiple vector labeling and three-color FRET experiments will allow a detailed analysis of the specific geometries of each conformer.

3.4 Experimental procedures[5]

Single-molecule FRET

The single-molecule FRET experiments were performed using clean quartz slides and class coverslips to create sample chambers, and chamber surfaces were prepared for DNA immobilization by coating the surfaces with BSA-biotin and streptavidin as described in Chapter 2. As a control, experiments were performed on polyethylene glycol (PEG) coated slide surfaces, and the results of experiments on PEG coated slides were indistinguishable from the results of the experiments performed on BSA-coated slides. Unless otherwise noted, measurements were made at 22°C in a single-molecule imaging buffer containing 25 mM HEPES pH 8.1, 0.1 mM EDTA, 10% glycerol, 0.4% glucose, 0.1 mg/mL glucose oxidase, 0.02 mg/mL catalase, and saturated Trolox ® [55]; $MgCl_2$ and NaCl were included as indicated in the text. The experiments were carried out on a prism-type total internal reflection (TIR) microscope, and data was collected with 32 ms time resolution. For experiments requiring temperature regulation, a water-circulating bath (NESLAB) was connected to the microscope stage, and the sample temperature was monitored directly using a thermocouple. For the data analysis, a two frame average of the FRET efficiency for each individual molecule was used to determine histograms representing the population distribution for each data set. The time traces of donor emission, acceptor emission, and calculated FRET efficiency were analyzed when appropriate to extract transition rates between FRET states using hidden Markov analysis [3].

---

[5] All single-molecule FRET experiments were done by J.E.C.

Oligonucloetides

All DNA strands were purchased from Integrated DNA Technologies (IDT). The DNA substrates were formed by annealing the strands together in a ratio of 1:1.2 (two-stranded substrates) or 1:1.5:1.6 (three-stranded substrates) in an annealing buffer containing 20 mM Tris pH 8.0, 5 mM $MgCl_2$, and 1 mM NaCl by heating the strands at 80°C for three minutes followed by slow cooling to room temperature. All annealing reactions were done in the dark and the annealed DNA was stored at -20°C.

The sequences of the oligonucleotides for each DNA construct are as follows:

$\underline{(CTG)_{13}\text{ hairpin}}$

5’ Cy3 TCTGCCTCAAGACGGTAGT Cy5 3’

5’ACTACCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTTT

GAGGCAGATTTTTTTTTTTTTTT Biotin 3’

$\underline{(CAG)_{13}\text{ hairpin}}$

5’ Cy3 TCTGCCTCAAGACGGTAGT Cy5 3’

5’ACTACCGTCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT

TTGAGGCAGATTTTTTTTTTTTTTT Biotin 3’

(CTG)$_7$(CAG)$_6$ hairpin

5’ Cy3 TCTGCCTCAAGACGGTAGT Cy5 3’

5’ACTACCGTCACAGCAGCAGCAGCAGCTGCTGCTGCTGCTGCTGTT

TGAGGCAGATTTTTTTTTTTTTTT Biotin 3’

(CTG)$_6$CG(CTG)$_6$ 3way

5’ Cy3 TCTGCCTCAAGACGGTAGT 3’

5’ Cy5 ACTACCGTCACTGCTGCTGCTGCTGC 3’

5’ GCTGCTGCTGCTGCTGCTGTTTGAGGCAGATTTTTTTTTTTTTTT Biotin 3’

(CTG)$_6$CG(CAG)$_6$ 3way

5’ Cy3 TCTGCCTCAAGACGGTAGT 3’

5’ Cy5 ACTACCGTCACAGCAGCAGCAGCAGC 3’

5’ GCTGCTGCTGCTGCTGCTGTTTGAGGCAGATTTTTTTTTTTTTTT Biotin 3’

(CTG)$_{13}$ hairpin-B

5’ Cy5 TGATGGCAGAACTCCGTCT Cy3 3’

5’ Biotin TTTTTTTTTTTTTTTAGACGGAGTTCTGCTGCTGCGCTGCTGCTGCTG

CTGCTGCTGCTGCTGACTGCCATCA 3’

Homoduplex

5’ Cy5 TGATGGCAGAACTCCGTCT Cy3 3’

5’ Biotin TTTTTTTTTTTTTTTAGACGGAGTTCTGCCATCA 3’

A$_2$ bulge

5’ Cy5 TGATGGCAGAACTCCGTCT Cy3 3’

5’ Biotin TTTTTTTTTTTTTTTAGACGGAGTTAACTGCCATCA 3’

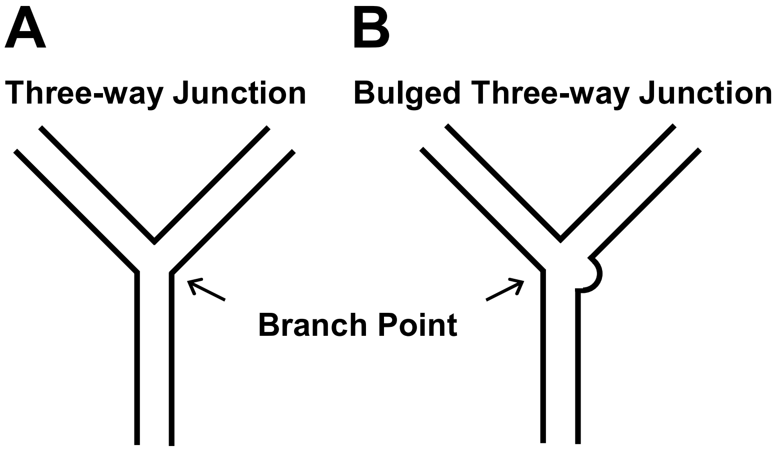

A
Three-way Junction
B
Bulged Three-way Junction
Branch Point

Figure 3.1
(A) Three-way DNA junctions are composed of three helices that share a common branch point. (B) Three-way DNA junctions with one or more unpaired bases at the branch point have increased conformational flexibility that allows them to undergo coaxial stacking of two of the helices.

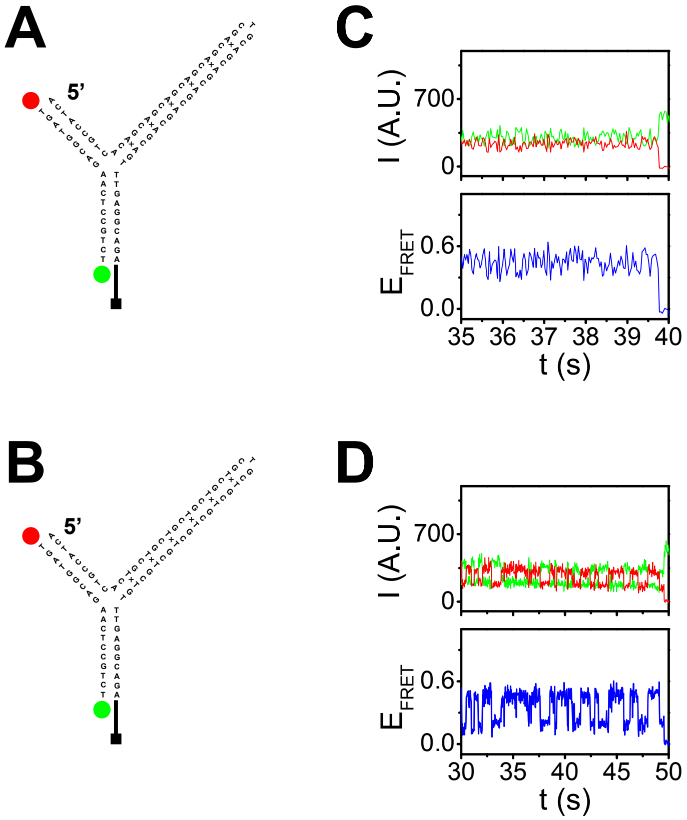


Figure 3.2

Schematic representations of the (A) $(CAG)_{13}$ hairpin and (B) $(CTG)_{13}$ hairpin are shown. Each junction is labeled on the short 'bottom' strand with donor (5' end, green ball) and acceptor (3' end, red ball) fluorophores for fluorescence observations. In addition, each junction contains a 3' poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). The ×'s denote base-base mismatches. Representative single-molecule time traces for the (C) $(CAG)_{13}$ hairpin and (D) $(CTG)_{13}$ hairpin substrates are shown. In each figure, the top traces are the emission intensity time traces for the donor (green) and acceptor (red), and the bottom traces show the corresponding FRET efficiency values. In (C) and (D) the buffer contains 110 mM NaCl and 20 mM $MgCl_2$.

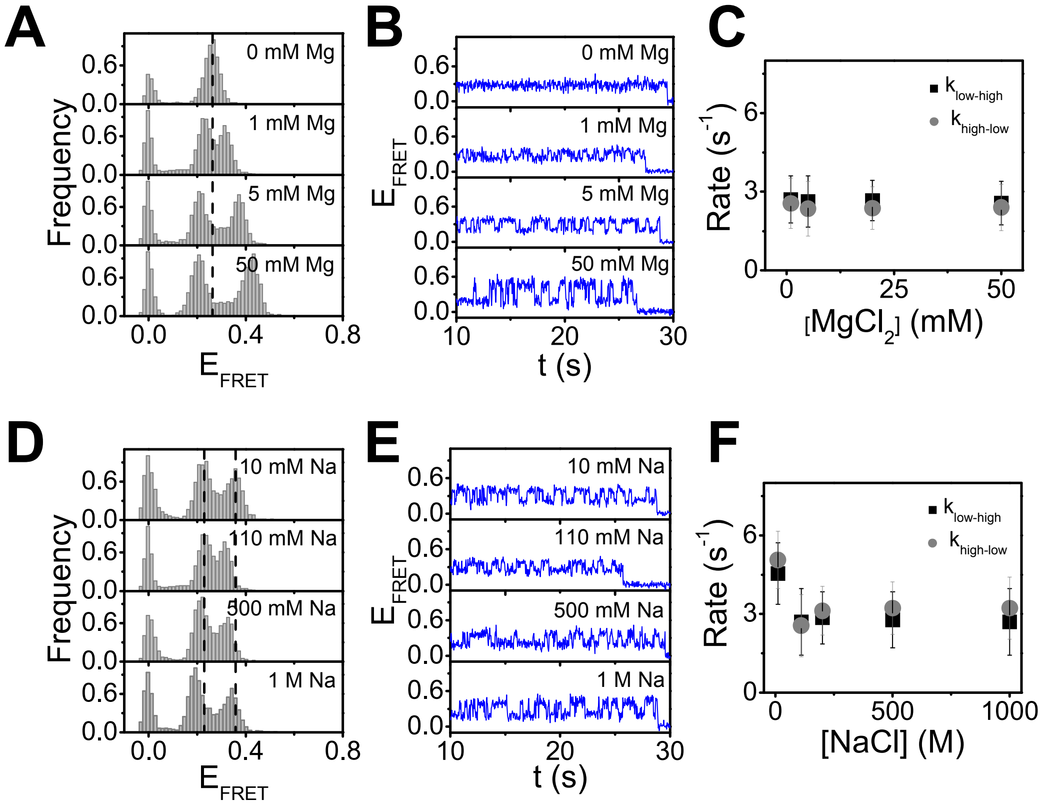


Figure 3.3
(A) The FRET efficiency histograms for the $(CTG)_{13}$ hairpin in 110 mM NaCl and the indicated amount of $MgCl_2$ are shown. (B) Representative FRET efficiency histograms for the $(CTG)_{13}$ hairpin at 110 mM NaCl and the indicated amount of MgCl2 are shown. (C) Transition rates for the low FRET to high FRET and high FRET to low FRET transitions shown in (B) are given. (D) The FRET efficiency histograms for the $(CTG)_{13}$ hairpin in 1 mM $MgCl_2$ and the indicated amount of NaCl are shown. (E) Representative FRET efficiency histograms for the $(CTG)_{13}$ hairpin at 1 mM $MgCl_2$ and the indicated amount of NaCl are shown. (F) Transition rates for the low FRET to high FRET and high FRET to low FRET transitions shown in (E) are given.

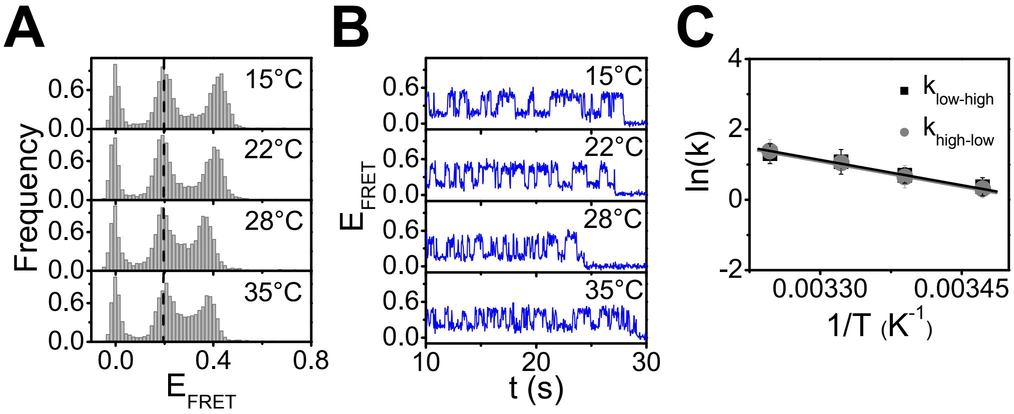


Figure 3.4
(A) The FRET efficiency histograms for the $(CTG)_{13}$ hairpin in 110 mM NaCl and 1 mM $MgCl_2$ are shown at temperatures ranging from 15 – 35 °C. (B) Representative FRET efficiency time traces for the $(CTG)_{13}$ hairpin in 110 mM NaCl and 1 mM $MgCl_2$ are shown at temperatures ranging from 15 – 35 °C. (C) Arrhenius plot of the transition rates for the low FRET to high FRET and high FRET to low FRET transitions shown in (B) are given.

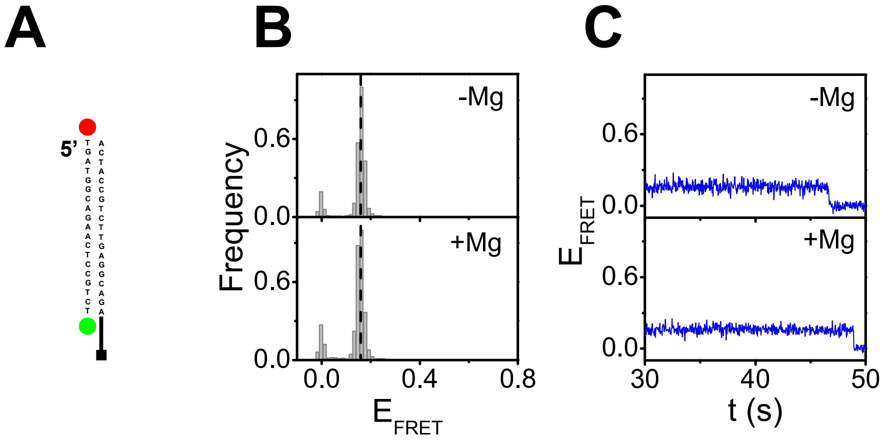


Figure 3.5
(A) A schematic of the homoduplex substrate is shown. As with the three-way junction substrates, the homoduplex substrate is labeled with donor (green ball) and acceptor (red ball) fluorophores. In addition, the substrate contains a poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). (B) The FRET efficiency histograms for the homoduplex substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (C) Representative FRET efficiency time traces for the homoduplex substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown.

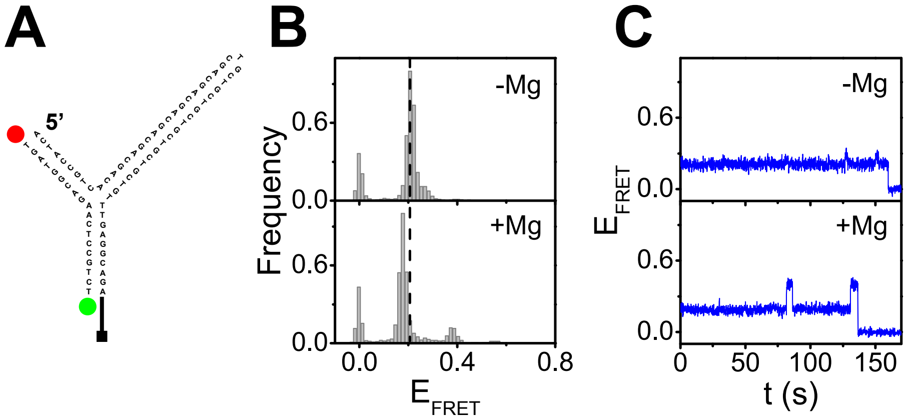


Figure 3.6
(A) A schematic of the $(CTG)_7(CAG)_6$ hairpin substrate is shown. As with the $(CTG)_{13}$ hairpin substrate, the homoduplex substrate is labeled with donor (5’, green ball) and acceptor (3’ red ball) fluorophores. In addition, the substrate contains a 3’ poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). (B) The FRET efficiency histograms for the $(CTG)_7(CAG)_6$ hairpin substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (C) Representative FRET efficiency time traces for the $(CTG)_7(CAG)_6$ hairpin substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown.

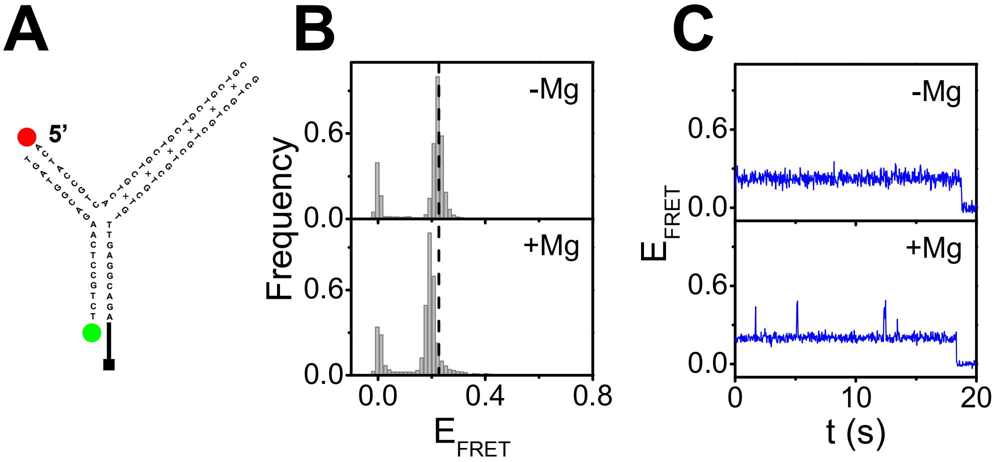


Figure 3.7
(A) A schematic of the $(CTG)_7(CAG)_6$ hairpin substrate is shown. As with the $(CTG)_{13}$ hairpin substrate, the homoduplex substrate is labeled with donor (5’, green ball) and acceptor (3’ red ball) fluorophores. In addition, the substrate contains a 3’ poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). (B) The FRET efficiency histograms for the $(CTG)_7(CAG)_6$ hairpin substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (C) Representative FRET efficiency time traces for the $(CTG)_7(CAG)_6$ hairpin in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown.

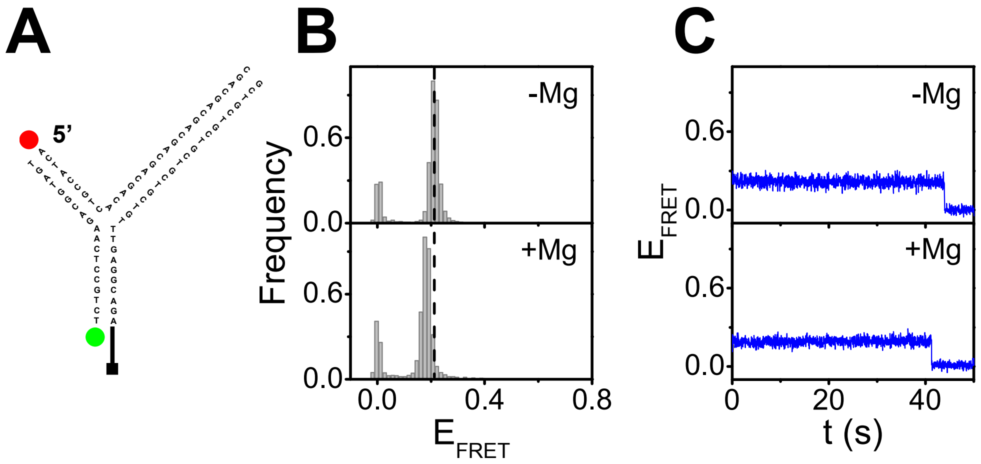


Figure 3.8
(A) A schematic of the $(CTG)_6(CG)(CAG)_6$ three-way substrate is shown. The substrate is labeled with donor (5’, green ball) and acceptor (3’ red ball) fluorophores. In addition, the substrate contains a 3’ poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). (B) The FRET efficiency histograms for the $(CTG)_6(CG)(CAG)_6$ three-way substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (C) Representative FRET efficiency time traces for the $(CTG)_6(CG)(CAG)_6$ three-way substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown.

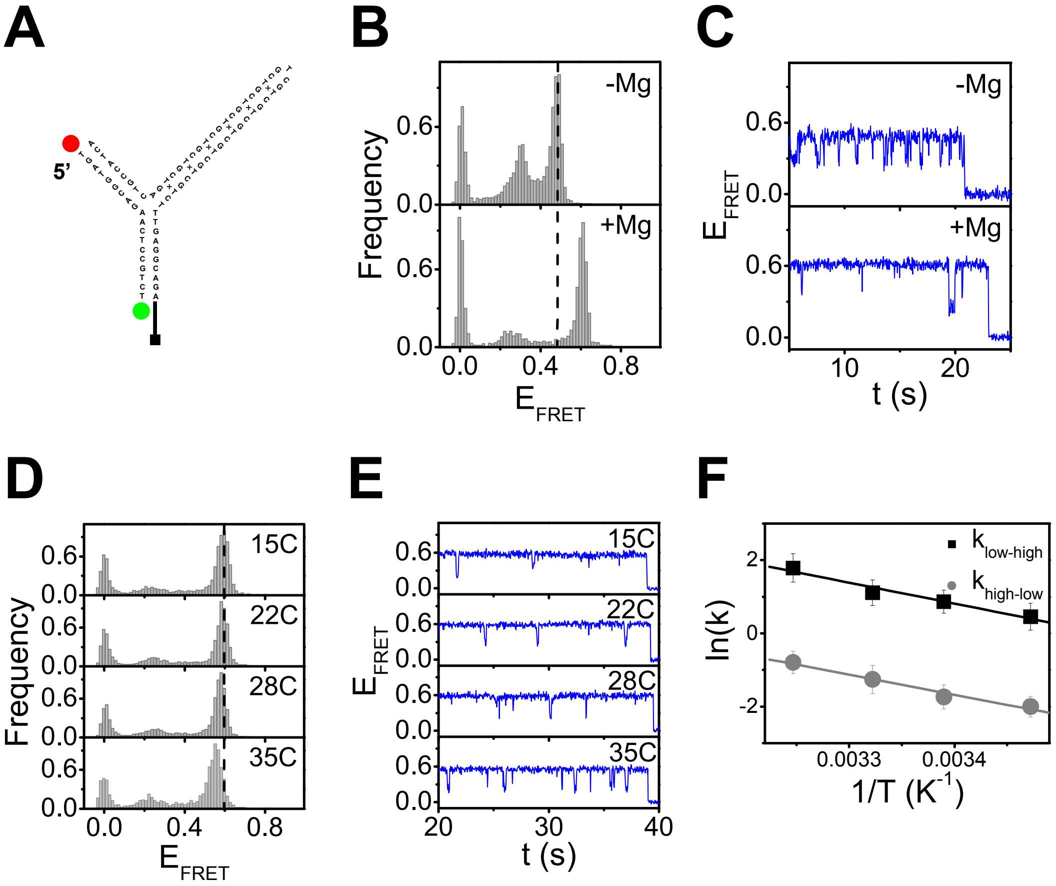


Figure 3.9
(A) A schematic of the $(CTG)_{13}$ hairpin-B substrate is shown. The substrate is labeled with donor (3', green ball) and acceptor (5', red ball) fluorophores. In addition, the substrate contains a 5' poly-dT extension (bold, black line) that is conjugated to a biotin molecule (black square). (B) The FRET efficiency histograms for the $(CTG)_{13}$ hairpin-B substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (C) Representative FRET efficiency time traces for the $(CTG)_{13}$ hairpin-B substrate in 110 mM NaCl and either 0 or 20 mM $MgCl_2$ are shown. (D) The FRET efficiency histograms for the $(CTG)_{13}$ hairpin-B substrate in 110 mM NaCl and 1 mM $MgCl_2$ are shown at temperatures ranging from 15 – 35 °C. (E) Representative FRET efficiency time traces for the $(CTG)_{13}$ hairpin-B substrate in 110 mM NaCl and 1 mM $MgCl_2$ are shown at temperatures ranging from 15 – 35 °C. (F) Arrhenius plot of the transition rates for the low FRET to high FRET and high FRET to low FRET transitions shown in (E) are given.

Chapter 4

Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops[6]

In this chapter, single-molecule FRET is used to observe binding and dissociation of MSH2/MSH3 proteins to and from DNA substrates in real time. We find that MSH2/MSH3 binds, bends, and dissociates from repair-competent loops, while repair-resistant loops trap nucleotide-bound MSH2/MSH3, inhibiting its dissociation from the DNA. We propose that the specific interactions between MSH2/MSH3 and the extrahelical loop governs whether the loop is removed by the DNA mismatch repair system or escapes repair and becomes a precursor for mutation.

4.1 Introduction

Insertion or deletion of small extrahelical loops is one of the most common mutations in human cancers [98-100], but the mechanism by which they occur is unknown. The Human DNA Mismatch Recognition Complex MSH2/MSH3 functions to bind and signal the repair of extrahelical loops [18, 19, 101-105], but the biochemical activities of MSH2/MSH3 and the mismatch repair efficiencies are loop-dependent [29,

---

[6] This work has been published as a paper:
W. H. Lang*, J. E. Coats*, J. Majka, G. L. Hura, Y. Lin, I. Rasnik, and C. T. McMurray. 'Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops.' *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108 no. 42 E837-E844 (2011). (*Equal contribution)

105-107]. For this reason, defects in repair mediated by MSH2/MSH3 may be a major source of insertion/deletion mutations.

Small $(CA)_4$ loops of DNA can be faithfully repaired by MSH2/MSH3 both *in vitro* [105, 106] and *in vivo* [107-109]. In contrast, hydrogen bonded CAG hairpin loops are not excised, and confer genomic instability through insertion and amplification of CAG repetitive tracts [29, 110, 111]. Thus, although both $(CA)_4$ loops and CAG hairpins both have extrahelical loops (a.k.a. three-way junctions), MSH2/MSH3 interacts with them differently [29, 105, 106]. It is not known why one substrate is repaired more efficiently than the other, but the consequence is remarkable: Inefficient repair of CAG loops results in mutations that underlie more than 20 hereditary neurodegenerative or neuromuscular diseases [29, 110, 112, 113].

Here, the underlying basis for discriminating repair-competent and repair-resistant DNA loops by human MSH2/MSH3 is examined. We find that MSH2/MSH3 binds with similar affinity to repair-competent $(CA)_4$ loops and to repair-resistant CAG hairpins. However, the three-way DNA junction formed from CAG hairpins adopts a conformational state that traps nucleotide-bound MSH2/MSH3 and inhibits its dissociation from the hairpin. Together, our biochemical and single-molecule FRET results imply that repair-resistant CAG hairpins provide a unique but nonproductive binding site for nucleotide-bound MSH2/MSH3, which fails to effectively couple DNA binding with downstream repair signaling. We envision that conformational regulation of small loop repair occurs at the level of junction dynamics.

4.2 Results

Ensemble analysis of MSH2/MSH3 activity

The purified human MSH2/MSH3 protein (hereafter referred to as MSH2/MSH3) used in the studies presented here was of high quality. The full-length MSH2/MSH3 was expressed and copurified as a heterodimer, and each subunit, when resolved by PAGE, migrated as a single band according to the expected molecular mass (**Figure 4.1A**).

The DNA binding affinity and nucleotide binding properties of MSH2/MSH3 bound to a $(CA)_4$ loop or CAG hairpin loops of 7 or 13 triplet repeats (denoted $(CAG)_7$ and $(CAG)_{13}$, respectively) were probed using ensemble techniques. Both the loop and hairpin templates were constructed from annealing two single-stranded oligonucleotides, and the substrates were designed so that neither the $(CA)_4$ loop nor the CAG stems had complementary sequences within the duplex portion of the template. Thus, the junction templates folded into stable extrahelical loops which have been previously characterized in solution [29, 106]. Folding of the CAG loops creates A/A mismatches every third base pair in the stem, for a total of three mismatches in the $(CAG)_7$ template or a total of six mismatches in the $(CAG)_{13}$ template (**Figure 4.1B**). The DNA templates were analyzed by gel electrophoresis to (i) ensure that the absence of any traces of single-stranded DNA and (ii) demonstrate that the DNA loops were intact, as judged by an increase in the loop size and gel mobility. Unless specifically noted, all DNA templates were synthesized containing a duplex base of 18 bases.

When MSH2/MSH3 was pre-bound to a repair-competent $(CA)_4$ loop or repair-resistant $(CAG)_7$ and $(CAG)_{13}$ hairpin templates, little difference in nucleotide affinity was observed. As measured by UV-crosslinking (**Figure 4.1C,D**), the affinity of ADP or

ATP to either subunit of DNA-bound MSH2/MSH3 was substantially weaker when MSH2/MSH3 was bound to DNA (**Table 4.1**). However, the reduction in nucleotide affinity for MSH2/MSH3 did not display striking differences among $(CA)_4$ loop, $(CAG)_7$ hairpin, and $(CAG)_{13}$ hairpin templates. Thus, MSH2/MSH3 binds nucleotides with high affinity at both repair-competent and repair-resistant templates.

To test whether nucleotide binding to MSH2/MSH3 altered its association with DNA, each DNA template was labeled with fluorescein at the 5' end of the bottom (shorter) strand, and the DNA binding affinity was determined via fluorescence anisotropy. In the absence of bound nucleotide, the apparent affinity of MSH2/MSH3 for both the $(CA)_4$ loop and the $(CAG)_7$ and $(CAG)_{13}$ hairpin templates was in the low nanomolar range (**Table 4.2**), and the apparent affinity was in good agreement with previous measurements [29, 105, 106]. The presence of magnesium decreased the affinity of ATP-bound MSH2/MSH3 to any template by about 10-fold, but, in general, DNA binding of ADP- or ATP-bound MSH2/MSH3 did not distinguish the repair-competent $(CA)_4$ loop from the repair-resistant $(CAG)_7$ hairpin or $(CAG)_{13}$ hairpin.

Single-molecule FRET DNA constructs

To study the conformational differences of mismatch recognition between the well-repaired $(CA)_4$ loop and the poorly-repaired $(CAG)_{13}$ DNA, we measured the protein-induced conformational dynamics of each substrate using single-molecule FRET. DNA substrates were prepared that were identical to those used in the biochemical DNA-binding experiments, except that the bottom strand was labeled with Cy3 (on the 5' end) and Cy5 (on the 3' end). Thus, for each template used in the single-molecule FRET experiments, the local environment of the fluorophores was identical. For both templates, the top strand at the 3’ end contains a poly-dT extension and a biotin tag for immobilization on a streptavidin-coated surface for observation. The extension was designed to prevent potential interactions of the fluorophores with the streptavidin surface. In these constructs, single-molecule FRET was used to probe the relative changes in distance between the Cy3 and Cy5 fluorophores.

MSH2/MSH3 Stabilizes a High FRET State When Bound to the Repair-Competent (CA)$_4$ Loop

In the absence of protein, there is a narrow distribution of E_{FRET} values for the (CA)$_4$ loop at $E_{FRET} \sim 0.31$ (**Figure 4.2A**). Following the dynamics of each individual FRET pair by plotting the time traces of the donor (Cy3) and acceptor (Cy5) emission and each corresponding FRET efficiency time trace reveals that there were no observable transitions within the time resolution of the experiment. To determine if transitions faster than the time resolution of the experiment were present, we measured the recovery rate of the acceptor dye intensity from the transitions to nonfluorescent states in the presence of 2-mercaptoethanol. We compared results for the (CA)$_4$ loop to the results for the homoduplex DNA (the sequences of the two substrates were identical, except for the (CA)$_4$ loop so that the two substrates had identical local environments for the fluorophores). The recovery of the intensity for the (CA)$_4$ loop was an order of magnitude faster than for the homoduplex DNA (**Figure 4.3**). Since recovery of fluorescence is facilitated by the donor and acceptor fluorophores coming into close proximity (2-3 nm) with each other [2], this experiment revealed that the ends of the

$(CA)_4$ loop substrate came into close proximity, confirming that there are conformational fluctuations in the $(CA)_4$ loop substrate even though they were not directly observable within the time resolution of our single-molecule FRET experiments.

Addition of MSH2/MSH3 to the $(CA)_4$ loop template, in the absence of nucleotides, led to the appearance of a new E_{FRET} peak (a bound state) with a higher FRET value of $E_{FRET} \sim 0.4$ (**Figure 4.2A**). The population in the high FRET state increased with protein concentration until the entire population had shifted to $E_{FRET} \sim 0.4$. This result is consistent with the DNA binding affinities determined by fluorescence anisotropy, which show that MSH2/MSH3 has high affinity for this substrate (**Table 4.2**). Thus, MSH2/MSH3 forms a stable complex with the repair-competent $(CA)_4$ loop template in which the two ends of the heteroduplex loop were positioned more closely, suggestive of protein-induced bending.

To monitor the conformational dynamics of the transitions, individual donor and acceptor emission intensity traces were followed, and the corresponding FRET efficiency traces were calculated. Experimental observation times were limited typically to less than 60 s by photodestruction of the acceptor. There were no observable dynamics of the bound, bent FRET state within the time resolution of the experiment, but a few of the single molecule time traces captured conformational transitions that were consistent with a protein binding/dissociation event (**Figure 4.2B**). However, the lifetime of the bound state was longer than our maximum observation time (**Figure 4.4**), so binding and dissociation events could not be seen on the same time trace. Similar results were obtained with MSH2/MSH3 bound to a comparably labeled AA bulge (**Figure 4.2A,B**), which is also a template for MSH2/MSH3-dependent repair.

As a negative control, we investigated the conformational properties of MSH2/MSH3 in complex with homoduplex DNA, since MSH2/MSH3 has weak affinity and rapidly dissociates from homoduplex DNA. No high FRET population was observable when MSH2/MSH3 was added to homoduplex DNA, even at very high protein concentration (**Figure 4.2E,F**).

As a positive control, we tested the ability of the human mismatch recognition protein MSH2/MSH6 to bend a substrate specific for it: A G-T mismatched template. Both MSH2/MSH6 and bacterial MutS have been shown to bend G-T mismatched DNA at the site of a mismatch [20, 21, 23], and from those studies, it has been inferred that bending is an important component of mismatch recognition in general. Human MSH2/MSH6 was purified and added to a labeled, surface-immobilized G-T mismatched DNA. Indeed, for the MSH2/MSH6 complex, we observed a high FRET shift that was similar in magnitude to the shift induced by MSH2/MSH3 on the $(CA)_4$ loop (**Figure 4.2G**). The single molecule traces indicated that the high FRET state was stable (**Figure 4.2H**). MSH2/MSH6 does not bind to the $(CA)_4$ loop, and no high smFRET population was observable when MSH2/MSH6 was added to that template, even at very high protein concentration. Thus, MSH2/MSH6 and MSH2/MSH3 complexes displayed similar transitions when bound to their preferred repair-competent templates with an estimated bending angle of 40-45° (**Figure 4.5**).

Nucleotide Binding Increases the Dissociation of MSH2/MSH3 from the $(CA)_4$ Loop Under Hydrolytic Conditions

The MutS family of proteins couple DNA binding and ATP hydrolysis to initiate downstream repair, so we next tested the effects of ATP binding and hydrolysis on the conformational dynamics of the MSH2/MSH3-bound $(CA)_4$ loop substrate. ATP was added to a complex containing MSH2/MSH-bound $(CA)_4$ loop DNA in the absence (-Mg) or presence (+Mg) of magnesium, and single-molecule FRET was used to determine the conformational dynamics of the complexes in both non-hydrolyzing (-Mg, **Figure 4.6A**) and hydrolyzing (+Mg, **Figure 4.6B**) conditions. Generation of the high FRET (bent) state by DNA-bound MSH2/MSH3 was independent of magnesium and nucleotide binding. ATP (+Mg) binding, however, weakened the affinity of MSH2/MSH3 for the $(CA)_4$ loop DNA, and more nucleotide-bound MSH2/MSH3 was required to saturate the high FRET shift compared to the absence of nucleotide. Inspection of the single-molecule intensity and FRET efficiency time trajectories reveal that ATP binding, under hydrolytic conditions, resulted in a striking alteration in the dynamics of MSH2/MSH3 binding. In the presence of ATP (+Mg), multiple transitions between high FRET states and low FRET states were obvious in the single-molecule traces (**Figure 4.6C**), as the lifetime of the nucleotide-bound MSH2/MSH3 bound to the $(CA)_4$ loop only seconds (compared with a lifetime of minutes for MSH2/MSH3 bound to the $(CA)_4$ loop in the absence of nucleotide). Thus, MSH2/MSH3 binding was sufficient to stabilize the high FRET state, and binding of ATP under hydrolytic conditions increased dissociation of MSH2/MSH3 from the $(CA)_4$ loop.

To determine the binding and dissociation kinetics, we applied a hidden Markov model to hundreds of time traces for several MSH2/MSH3 concentrations to generate robust measures of the average transition times. We found that the transition rate to the high FRET state increased with protein concentration, but the transition rate back to the initial FRET state was independent of protein concentration (**Figure 4.6D**). Collectively, these findings indicated that the shift to the high FRET state depended on MSH2/MSH3 binding to the $(CA)_4$ loop, while the reverse transition rate arose from concentration-independent MSH2/MSH3 dissociation.

Binding of MSH2/MSH3 to the Repair-Deficient $(CAG)_{13}$ Hairpin Results in the Appearance of an Unique Conformational Population

Ensemble experiments show that MSH2/MSH3 binds well to both the $(CA)_4$ loop and the $(CAG)_{13}$ hairpin, but only the $(CA)_4$ loop is accurately excised and repaired *in vitro* and *in vivo*. Therefore, we compared the conformational dynamics of the $(CAG)_{13}$ hairpin to those of the $(CA)_4$ loop to determine if the conformational dynamics of mismatch recognition contributes to mismatch selectivity.

The $(CAG)_{13}$ hairpin DNA (**Figure 4.7A**) alone displayed a relatively broad distribution of FRET efficiencies, centering around E_{FRET} ~0.3 (**Figure 4.7B**, DNA only). Remarkably, binding of MSH2/MSH3 to the $(CAG)_{13}$ hairpin resulted in two new FRET populations, one at E_{FRET} ~0.43 (similar to that of the $(CA)_4$ loop) and another at E_{FRET} ~0.20 (**Figure 4.7B**). Each FRET state was stable, with an average lifetime longer than 30 s (**Figure 4.7C,D**). Thus, in contrast to binding $(CA)_4$ loops in a single conformation, MSH2/MSH3 binds $(CAG)_{13}$ hairpin in two conformational populations.

Surprisingly, the high FRET state (E_{FRET} ~0.43) largely disappears when $(CAG)_{13}$-bound MSH2/MSH3 was occupied with nucleotide. Under hydrolytic conditions (+Mg), addition of ATP to the $(CAG)_{13}$-bound MSH2/MSH3 shifted the equilibrium populations towards the low FRET state (**Figure 4.7E**). Analysis of the single molecule traces indicated that ATP (+Mg) occupancy of MSH2/MSH3 significantly shortened the average lifetime for the high FRET state to around 5-10 seconds (**Figure 4.7G**). Under the same conditions, the low FRET state was stable, and dissociation was rarely observed (**Figure 4.7F**). Thus, the repair-deficient $(CAG)_{13}$ hairpin template differed from the repair-competent $(CA)_4$ loop in that nucleotide occupancy of MSH2/MSH3 promoted its dissociation from the high FRET state and the nucleotide-bound MSH2/MSH3 was, instead, trapped in the low FRET state.

The differences between the MSH2/MSH3-mediated conformational dynamics on the repair-resistant $(CAG)_{13}$ hairpin and the repair-competent $(CA)_4$ loop were striking. The shift to the low FRET state on the $(CAG)_{13}$ hairpin could not be explained by differential affinity of nucleotide-bound MSH2/MSH3: The biochemical data indicated that nucleotide binding and DNA-binding affinities of MSH2/MSH3 on the $(CAG)_{13}$ hairpin and the $(CA)_4$ loop were similar. The differences in MSH2/MSH3-induced DNA conformational dynamics could also not be explained by oligomerization of MSH2/MSH3 on the DNA templates, since it had been previously reported that the stoichiometry of MSH2/MSH3 on both the $(CA)_4$ loop and the $(CAG)_{13}$ hairpin is one heterodimer per DNA molecule as measured by sedimentation equilibrium analysis. Thus, models in which the low and high FRET states were stabilized by two or more MSH2/MSH3 heterodimers appeared unlikely.

Given the reported data on the conformational dynamics of MSH2/MSH3 in complex with the repair-competent $(CA)_4$ loop and the repair-deficient $(CAG)_{13}$ hairpin loop, two models seemed possible. First the low and high FRET states could arise if two distinct nucleotide-bound forms of MSH2/MSH3 were able to bind to the $(CAG)_{13}$ hairpin and induce distinct conformations. Alternatively, the low and high FRET states might arise if the $(CAG)_{13}$ hairpin template itself formed two major DNA conformations that were able to bind MSH2/MSH3. We considered both possibilities.

The High and Low FRET states of the $(CAG)_{13}$ Hairpin Loop Do Not Arise from Binding of Distinct Nucleotide-Bound Forms of MSH2/MSH3

The MSH2 and MSH3 subunits of MSH2/MSH3 bind nucleotides stochastically, and efficient hydrolysis results in formation of ADP-MSH2/MSH3-empty and empty-MSH2/MSH3-ADP in solution. Only ADP-MSH2-MSH3-empty stably binds to the $(CA)_4$ loop DNA. However, it was possible that the altered conformation of the $(CAG)_{13}$ hairpin template permitted binding of both ADP-bound forms of MSH2/MSH3 (**Figure 4.8A**). In such a model, binding of the two distinct ADP-bound forms of MSH2/MSH3 to the $(CAG)_{13}$ hairpin template would result in the observed high and low FRET states.

To test this hypothesis, mutants of MSH2/MSH3 were created in which only one of the Walker motifs was competent to bind nucleotides (**Figure 4.8B**). The lysine that is critical for binding nucleotides in the Walker A site was changed to methionine in the MSH2, MSH3, or both subunits. These mutants are referred to as sgl2 (mutation in MSH2 only), sgl3 (mutation in MSH3 only), or dbl (both subunits mutated), depending on the site(s) of the amino acid change(s) (**Figure 4.8B**). The amino acid substitutions had no effect on the expression of the protein relative to the wild-type protein, and each subunit was expressed at stoichiometric levels (**Figure 4.8C**). Thus, each mutant heterodimer was purified and characterized with respect to DNA and nucleotide binding activities.

The mutant MSH2/MSH3 proteins had the anticipated nucleotide binding properties. As judged by cross-linking experiments, neither [α-32P]-ADP nor [α-32P]-ATP (+Mg) bound to MSH2/MSH3 dbl (**Figure 4.8D**, lane 4), while sgl2 and sgl3 bound nucleotides only in their intact site (**Figure 4.8D**, lanes 2 and 3). Wild-type MSH2/MSH3 bound both sites equally (**Figure 4.8D**, lane 1). Anisotropy measurements revealed that both sgl2 and sgl3 bound ADP with the same affinity as wild-type MSH2/MSH3 (**Table 4.3**). Thus, mutation in one site did not influence the nucleotide affinity in the other. Consequently, each mutant MSH2/MSH3 heterodimer was able to form a single nucleotide-bound complex which varied only in the nucleotide-bound subunit.

Next, the affinity of the mutant MSH2/MSH3 proteins for $(CA)_4$ loop or $(CAG)_{13}$ hairpin DNA was tested via fluorescence anisotropy (**Table 4**). With the exception of dbl, the wild-type and mutant MSH2/MSH3 substrates had good affinity for the $(CAG)_{13}$ hairpin in the presence of nucleotide (**Table 4**). Then, the nucleotide affinity for the DNA-bound wild-type and mutant MSH2/MSH3 was measured using fluorescently-labeled ATP and ADP (**Figure 4.8E,F**). ATP bound well to the MSH2 or MSH3 subunit of wt, sgl2, or sgl3 as free heterodimers (**Figure 4.8E**, open symbols), but none of these ATP-bound complexes adopted a stable ATP-bound state on the $(CAG)_{13}$ hairpin DNA (**Figure 4.8E**, closed symbols). In the presence of ADP(+Mg), wild-type and sgl3 MSH2/MSH3 retained high affinity for $(CAG)_{13}$ hairpin, but the sgl2 mutant had low affinity for the $(CAG)_{13}$ hairpin substrate in these conditions (**Figure 4.8F**, solid inverted triangles). Thus, nucleotide-bound MSH2/MSH3 associated to the $(CAG)_{13}$ hairpin only when nucleotide occupied the MSH2 subunit and the MSH3 subunit was empty (sgl3), yet both the high and low FRET states were observed (**Figure 5G**). These experiments argued against a model where the high and low FRET states arose from binding of two distinct nucleotide-bound MSH2/MSH3 complexes to the $(CAG)_{13}$ hairpin.

The Repair-Resistant $(CAG)_{13}$ Template Traps MSH2/MSH3 in the Low FRET State

For the $(CAG)_{13}$ hairpin DNA template (**Figure 4.9A**) in the absence of protein, increasing magnesium concentrations allowed the observation of the broad FRET efficiency peak at ~0.24 into two closely spaced DNA populations at E_{FRET} ~0.21 and ~0.24 (**Figure 4.9B**). The single-molecule intensity and FRET efficiency time traces showed that these two conformational populations were rapidly interconverting (**Figure 4.9C,D**).

Structurally, the $(CAG)_{13}$ hairpin DNA can be characterized as a three-way DNA junction with two homoduplex arms and one heteroduplex arm (the $(CAG)_{13}$ stem). Perfectly paired three-way DNA junctions have been shown to form a single stable conformation, so we hypothesized that the unpaired A-A mispaired bases in the stem of the $(CAG)_{13}$ hairpin might allow rearrangement of the junction into two major conformational populations of DNA. If the $(CAG)_{13}$ DNA intrinsically adopted high and low FRET states, then MSH2/MSH3 might preferentially bind to one of the states.

To test this idea, we stabilized the junction by converting the two A-A pairs closest to the junction of the $(CAG)_{13}$ hairpin template into A-T pairs (AT-$(CAG)_9$ hairpin) (**Figure 4.9E**). Introduction of the two A-T pairs at the base of the junction "locked" it into a single narrow distribution which did not show two populations upon increasing the concentration of magnesium in solution (**Figure 4.9F**). Moreover, AT-$(CAG)_9$ adopted a single stable state as shown in the single molecule traces (**Figure 4.9G,H**). We next added MSH2/MSH3 to the AT-$(CAG)_9$ hairpin and tested whether MSH2/MSH3 would bind this template in the same high and low conformations that it bound the $(CAG)_{13}$ hairpin template. Remarkably, binding of MSH2/MSH3 to the AT-$(CAG)_9$ hairpin resulted in only a low FRET conformational population (**Figure 4.10A**). Adding ATP and MSH2/MSH3 under hydrolyzing conditions lowered the affinity of MSH2/MSH3 to the AT-$(CAG)_9$ hairpin but did not alter the overall conformation of the $(CAG)_{13}$ hairpins (**Figure 4.10B**). In fact, few conformational transitions were observed even when MSH2/MSH3 was in the nucleotide-bound state (**Figure 4.10C,D**). Thus, MSH2/MSH3 bound stably to the AT-$(CAG)_9$ junction and did not dissociate readily from the low FRET conformation. When bound to MSH2/MSH3, the AT-$(CAG)_9$ hairpin adopted a single stable low FRET state which was not observed for the $(CA)_4$ loop template under any condition tested.

4.3 Discussion

How insertion and deletion mutations arise in the genome and why some loops are repaired more efficiently than others are unknown. Here, we show that MSH2/MSH3 discriminates between a repair-competent and repair-resistant loop by sensing the conformational dynamics of the DNA substrate at the branch point of the three-way junction. We propose that the conformational properties of the substrate govern whether a loop is removed or becomes a precursor for mutation. We find the repair-competent

$(CA)_4$ loop substrate is intrinsically a flexible hinge with dynamics that are faster than our time resolution. MSH2/MSH3 binds and stabilizes the bent state (**Figure 4.10E**, bending), presumably to verify the lesion. Upon nucleotide binding, the enzyme undergoes a series of rapid nucleotide-dependent steps and eventually dissociates to signal downstream repair (**Figure 4.10E**, sliding). Indeed, the single-molecule FRET results imply that the substrate dynamics induced by nucleotide-bound MSH2/MSH3 at the $(CA)_4$ loop have nonexponential dwell-time distributions, consistent with the presence of more than one kinetic step (**Figure 4.11**). Rapid association and dissociation puts the protein complex in position to verify a mismatch and move away from the lesion to initiate interactions necessary for downstream signaling.

The repair-resistant $(CAG)_{13}$ hairpin junction intrinsically adopts two discrete conformational states as indicated by the two-state FRET distribution (most noticeable at high magnesium concentrations) (**Figure 4.9**). Unliganded MSH2/MSH3 recognizes both of these conformational states with similar affinity and further separates them (**Figure 4.7**). MSH2/MSH3 can convert some of the hairpin junctions into a repair-competent bent state. However, upon nucleotide binding, MSH2/MSH3 dissociates from the bent state and, instead, is trapped by a junction configuration from which it cannot dissociate (**Figure 4.10**, trapped). The nucleotide-bound protein becomes “stuck” on the lesion and likely cannot carry out the steps leading to ADP/ATP exchange, which is critical for dissociation and downstream repair. These findings imply that the repair-resistant CAG-hairpins provide a unique but nonproductive binding site for nucleotide-bound MSH2/MSH3 which fails to effectively couple DNA binding with ATP hydrolysis.

The AT-$(CAG)_9$ hairpin junction differs only by two nucleotides relative to the $(CAG)_{13}$ hairpin loop, but only one intrinsic conformation is available for MSH2/MSH3 binding. Similar to the repair-resistant $(CAG)_{13}$ hairpin loop, ATP(+Mg)-bound MSH2/MSH3 cannot bend the AT-$(CAG)_9$ hairpin and then readily move away from the lesion. Instead, MSH2/MSH3 binds to the AT-$(CAG)_9$ hairpin in an unbent conformation, and the residence of MSH2/MSH3 on the AT-$(CAG)_9$ hairpin is long lived, whether or not the protein is bound with nucleotides (**Figure 4.10C,D**). Thus, dynamics of the junction is an active participant in directing loop conformation. We envision that conformational dynamics of small loop repair occurs at the level of the junction dynamics.

Mutations in the MMR (Mismatch Repair) machinery lead to an increase in spontaneous mutation rate, which is typically referred to as a mutator phenotype [98-100]. For example, about 15% of patients with hereditary nonpolyposis colorectal cancer have widespread genome instability, characterized by single base changes or changes in copy number at repetitive tracts [98-100]. The mutational spectrum in this class of MMR deficits reflects the inability of the mutated MMR machinery to correct postreplicative errors throughout the genome [98-100]. Our data provide a plausible mechanism for a second class ("class two") of MMR defects in which the lesion itself prevents its processing by the normal repair machinery [113]. Defective repair arises when the repair-resistant loops trap the MMR proteins during recognition of the lesion and they remain uncorrected. The resulting insertion and deletion mutations, in this case, will be "site-specific" in that they are limited to particular locations where the repair-resistant lesions reside. The properties of trinucleotide expansion characterize this type of mutation.

We do not know yet whether the unusual junction dynamics provides a general mechanism underlying all “class two” insertion/deletion mutations or whether the unusual dynamics are restricted to only some junctions. However, our results provide, at the structural level, a glimpse into why some loops are recognized differently by MSH/MSH3 and imply that the junction dynamics is at least one component in a complex process that leads to mutation.

Together, the biochemical and single molecule FRET data presented here clarifies two key issues bearing on the expansion of triplet repeats. First, the role of MSH2/MSH3 ATP hydrolysis activity in causing expansion has been unclear. A G674A Walter A site mutation in the MSH2 subunit suppresses the CTG (cytosine-thymine-guanine) expansion in mice [114], and prevents GAA (guanine-adenine-adenine) deletion in yeast [115], implying that ATP hydrolysis in the MSH2 subunit is a requisite step in expansion. However, we observe in our biochemical measurements that the G674A Walker A site mutant in the MSH2 subunit binds ATP poorly, if at all, in the context of MSH2/MSH3. Thus, the G674A Walker A site mutation does not block hydrolysis per se, but failure to bind ATP in the MSH2 subunit prevents formation of ADP-bound MSH2/MSH3, the major lesion-binding form [106]. Second, the effect of lesion binding on the biochemical properties of MSH2/MSH3 has been unclear. While ATP hydrolysis is reduced in MSH2/MSH3-bound CAG hairpins, the apparent nucleotide affinity and the k_{cat}/K_M for ATP hydrolysis are similar for MSH2/MSH3 when bound to a repair-competent $(CA)_4$ loop and the repair-resistant $(CAG)_{13}$ hairpin [29, 116]. In the study presented here, single molecule FRET is used to resolve the conformational dynamics of MSH2/MSH3 in complex with both lesions. The data provides definitive evidence that MSH2/MSH3 captures a distinct conformation of the $(CAG)_{13}$ hairpin, which significantly lengthens the lifetime of the protein-$(CAG)_{13}$ hairpin complex relative to MSH2/MSH3 bound to the repair-competent $(CA)_4$ loop. Because MSH2/MSH3 binds with equal apparent affinity to the $(CA)_4$ loop and the $(CAG)_{13}$ hairpin templates, the kc_{at}/K_M is expected to be similar, but the altered recognition properties of MSH2/MSH3 on the low FRET population cannot be resolved in bulk measurements [29, 116]. The time scale of the changes requires sensitive, high-resolution techniques to observe them. Because DNA binding inhibits ATP hydrolysis for MSH2/MSH6 [47, 117] and MSH2/MSH3 [29, 116], the longer lifetime of the MSH2/MSH3 on the repair-resistant template implies a reduction of ATP binding and/or hydrolytic activity in the straightforward conformation [29]. Collectively, our proposed model (**Figure 4.10E,F**) provides a basis for how an intact MMR complex can become inefficient when bound to particular types of loops. The junction dynamics are poised to be a pivot point for coupling DNA loop binding and ATP hydrolysis by an intact MSH2/MSH3 to outcomes of mutation or repair.

4.4 Experimental procedures[7]

Protein purification

His-tagged human MSH2/MSH3 was over-expressed in SF9 insect cells using pFastBac dual expression system (GIBCO-BRL) and purified as described previously [29, 105, 106]. SF9 insect cell pellets expressing human MSH2/MSH3 were obtained under contract with the University of Colorado Cancer Center's Cell Culture Core Facility (Denver, CO). All purification steps were performed at 4°C. Cells were resuspended in

---

[7] The single-molecule FRET experiments were done by J.E.C. and Y.L. The MSH2/MSH3 and MSH2/MSH6 protein was purified and ensemble experiments where done by W.H.L. and J.M.

Lysis buffer (25 mM HepesNaOH, pH 8.1, 300 mM NaCl, 20 mM Imidazole, 20% glycerol (v/v) containing a protease inhibitor cocktail (Roche)) and lysed by repeated passage through a 25 gauge needle. After centrifugation for 1 hr 20 min at 40,000g, the supernatant was loaded onto a 5 ml HiTrap chelating column (GE Healthcare) charged with Nickel and equilibrated with Lysis buffer. The bound proteins were then eluted with a 25 ml 20 to 200 mM imidazole gradient. The peak fractions containing the MSH2/MSH3 complex eluted at 140 mM imidazole and were then loaded onto a MonoP and HiTrap Heparin column (both GE Healthcare) connected in tandem and equilibrated in column buffer (25mM Hepes pH 8.1, 0.1 mM EDTA, 20% glycerol (v/v), 1 mM DTT) containing 300 mM NaCl. After collecting the flow through, the MonoP column was removed and the proteins bound to the Heparin column were eluted with a NaCl gradient from 300 mM to 1 M. The MSH2/MSH3 proteins eluted at 450 mM NaCl. The MSH2/MSH3 containing fractions were then applied to MonoQ (GE Healthcare) equilibrated in column buffer containing 100 mM NaCl and eluted with a 20 ml gradient form 100 mM to 1 M NaCl. The MSH2/MSH3 complex eluted at 320 mM NaCl. MSH2/MSH3 containing fractions were then aliquoted and frozen at -80°C.

Oligonucloetides

Oligonucleotides used in the binding studies were obtained from Operon or IDT. All oligonucleotides were purified by the manufacturer by two HPLC steps. The oligos were annealed (20 mM Tris pH 8.0, 500 mM NaCl) by heating the strands to 80°C for 3 minutes and cooling them slowly (3 hours) to room temperature. After annealing, the double stranded oligonucleotide was separated from residual single stranded oligonucleotide by gel filtration on SuperDex200 (GE Health care).

Nucleotides

Nucleotides of the highest grade were purchased from Sigma, prepared as 100 mM stock solutions at pH 7.0 and immediately frozen in small single use aliquots. [α32P]-ATP was purchased from Perkin Elmer, and [α32P]-ADP was derived by incubation of [α32P]-ATP with hexokinase. All preparations of nucleotides used contained less than 1% contamination of other nucleotides.

DNA and Nucleotide binding experiments

Experiments were performed as described previously [106]. Magnesium was included at 5 mM and nucleotides at 100 μM final concentration, unless indicated otherwise. All binding experiments were performed at room temperature under the following buffer conditions: 25 mM Hepes pH 8.1, 110 mM NaCl, 1.0 mM DTT, 0.1 mM EDTA, 100 μg/ml BSA, and 10% glycerol. Magnesium chloride was added to 5.0 mM final concentration, where indicated. For nucleotide binding experiments in the absence of magnesium, EDTA was added to 2 mM final concentration. After all components were added, reactions were incubated for 10 minutes before measurement.

UV-Cross-linking

Experiments were performed at a concentration of 100 nM protein and 500 nM DNA. UV Cross linking of labeled nucleotides to MSH2/MSH3 was done in a

Stratalinker 1800 (Stratagene) for 7.5 minutes. Crosslinking products were then analyzed by SDS-PAGE and signals quantified with a phosphoimager. Results were analyzed and binding constants determined with the aid of the curve fitting program GraphPad Prism.

Fluorescence Anisotropy

Anisotropy was measured using an Infinite M 1000 plate reader (Tecan Group Ltd.) at ambient temperature. Binding was monitored by either bodipy-labeled nucleotides (ADP, ATP, or AMP-PNP) or fluorescein-labeled DNA. Experiments were done at a concentration of 1 nM labeled nucleotide or 1nM labeled DNA. The G-factor was determined using 1 nM fluorescein as recommended by the manufacturer. All anisotropy measurements for K_D determinations were performed in triplicates.

Single-molecule FRET

The single-molecule FRET experiments were performed on a prism-type total internal reflection microscope which features 532 nm excitation from a Nd:YAG laser (50 mW, CrystaLaser) [6, 55]. This wide-field microscope system permits the simultaneous observation of ~200-300 DNA molecules that are immobilized on the surface of a quartz slide, allowing 1000's of molecules to be observed within minutes for each reaction condition. Cy3 and Cy5 emission intensities were detected with 32 ms time resolution using an electron multiplying charge-coupled device (CCD) camera (iXon, Andor Technology). Quartz slides (Chemglass) and glass coverslips (Fisherbrand) were coated with polyethylene glycol (PEG) (Laysan Bio, Inc.) and assembled with doublesided tape and epoxy to create sample chambers 5 mm wide, 20 mm long, and 100 μm tall. Two holes in the quartz slides—one on each end of the sample chamber—facilitated fluid exchange. The fluorescently-labeled DNA (10-20 pM) was flowed into the sample chamber and attached to the surface of the quartz slide by way of biotin-streptavidin bonds, and experiments were performed in a single-molecule imaging buffer containing 5 mM HEPES pH 8.1, 110 mM NaCl, 0.1 mM EDTA, 10% glycerol, 0.4% glucose, 0.1 mg/ml glucose oxidase, 0.02 mg/ml catalase, and saturated Trolox ®.

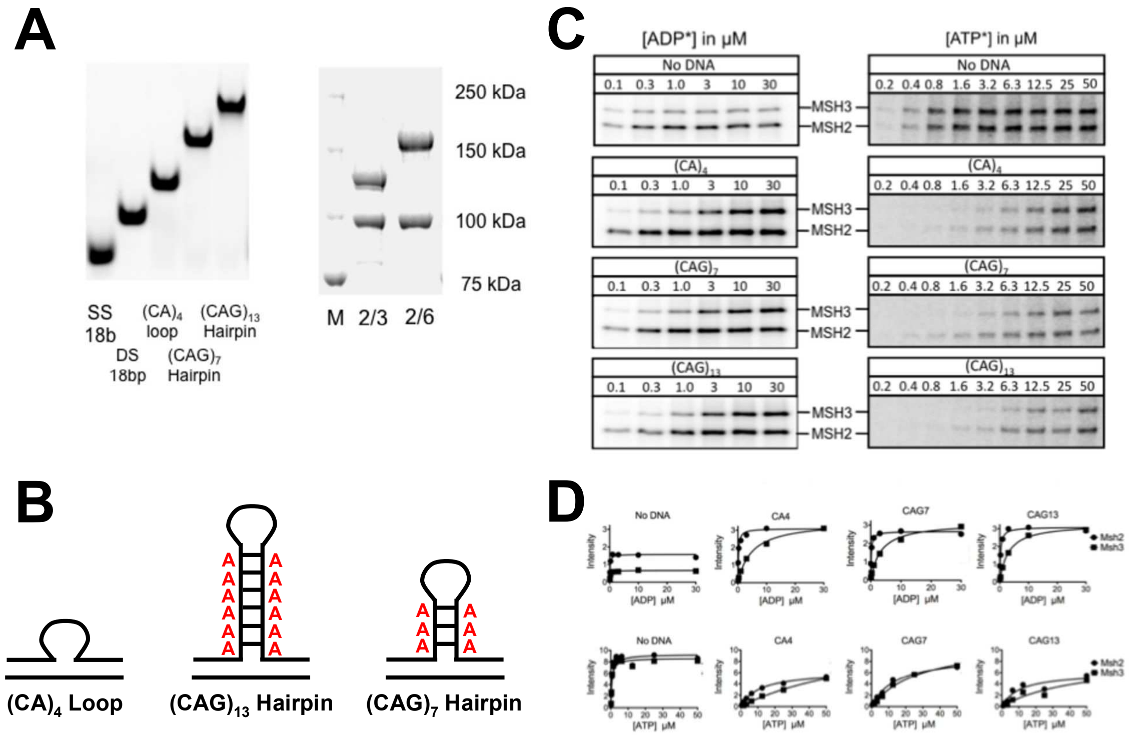


Figure 4.1

Ensemble analysis of MSH2/MSH3 activity. (A) (Left) The purified proteins resolved on native polyacrylamide gel and visualized by ethidium bromide staining. SS is the 18 nt single strand DNA that is the complementary strand for the looped templates, DS is the 18 bp homoduplex DNA, and the heteroduplex looped substrates are as labeled. (Right) Resolution of purified MSH2/MSH3 (middle lane) and MSH2/MSH6 (right lane) proteins by SDS-PAGE. (B) Schematic structures of the three-way DNA junction templates. The (CA)$_4$ loop and CAG hairpins are centrally located in the upper strand. The duplex base contains 18 bp, and the bottom strands were labeled with 5’ fluorescein for the fluorescence anisotropy experiments and 3’ Cy3/5’ Cy5 for the single-molecule FRET experiments. (C) Nucleotide binding of MSH2/MSH3 to repair-competent (CA)$_4$ loop and to the repair-resistant CAG hairpins. UV-cross-linking of (left) [α-^{32}P]-ADP or (right) [α-^{32}P]-ATP to the MSH2 and to the MSH3 subunits of MSH2/MSH3 in the presence of magnesium. The star indicates the labeled nucleotide, and the total concentration of nucleotide (in μM) is indicated above the gel boxes. The DNA templates are indicated. (D) Each cross-linked product was resolved on a PAGE gel after irradiation, and the signal intensity for each band were quantified and plotted versus nucleotide concentration. The K_D’s were obtained by fitting these data using Graph-Pad Prism and are summarized in Table 1.

Subunit	Ligand	No DNA	CA4	CAG7	CAG13
MSH2	ADP($-Mg^{2+}$)	210 ± 90	250 ± 90	190 ± 50	530 ± 120
	ADP($+Mg^{2+}$)	150 ± 60	150 ± 120	250 ± 50	440 ± 130
	ATP($-Mg^{2+}$)	730 ± 240	5,300 ± 700	4,800 ± 500	4,300 ± 800
MSH3	ADP($-Mg^{2+}$)	nq	nq	nq	nq
	ADP($+Mg^{2+}$)	160 ± 90	7,200 ± 1,000	4,000 ± 1,100	3,700 ± 540
	ATP($-Mg^{2+}$)	550 ± 190	53,800 ± 15,100	11,400 ± 4,000	59,200 ± 15,100

Nq: not quantifiable—lack of signal.

Table 4.1
Nucleotide binding affinities of MSH2/MSH3 subunits determined by cross-linking in the absence or presence of DNA templates (K_D in nM).

Template	CA4	CAG7	CAG13
(−Mg)	2.5 ± 0.5	3.8 ± 0.6	6.4 ± 0.7
(+Mg)	2.6 ± 0.7	5.1 ± 0.6	5.3 ± 0.9
ADP(−Mg)	1.6 ± 0.06	1.6 ± 0.06	2.8 ± 0.04
ADP(+Mg)	41.9 ± 5.3	41.6 ± 6.2	27.7 ± 2.4
ATP(−Mg)	3.0 ± 0.05	4.2 ± 0.04	6.6 ± 0.05
ATP(+Mg)	39.7 ± 4.3	70.9 ± 11.6	43.0 ± 5.3

Table 4.2
DNA-binding affinities of wild-type MSH2/MSH3 determined by fluorescence anisotropy in the presence or absence of nucleotides (K_D in nM).

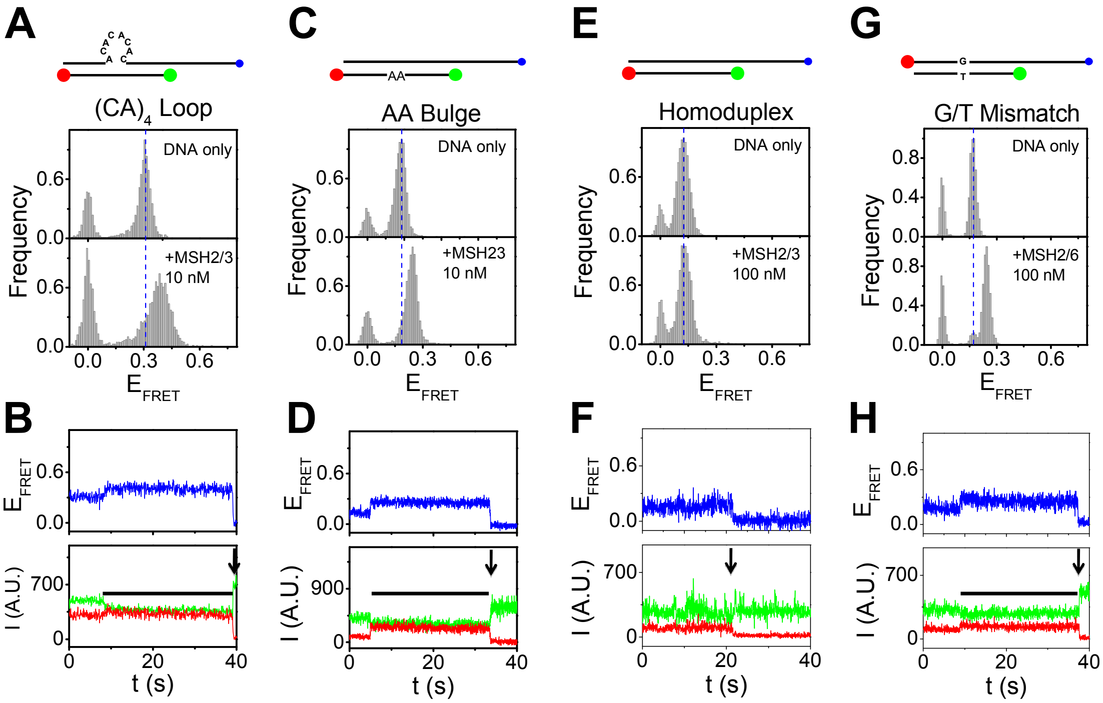


Figure 4.2
Binding of MSH2/MSH3 and MSH2/MSH6 to their preferred repair substrates causes substrate bending. (A) Single-molecule FRET efficiencies for the $(CA)_4$ loop substrate without (top) or with (bottom) MSH2/MSH3. In the substrate schematic, the green ball represents the Cy3 fluorophore, the red ball represents the Cy5 fluorophore, and the blue ball represents the biotin label. The protein concentration is indicated. (B) Representative time traces for donor (green, Cy3) and acceptor (red, Cy5) fluorescence intensities in arbitrary units and the corresponding FRET efficiency (blue). The black line indicates the time of the $(CA)_4$ loop in the high FRET state. The time of acceptor photobleaching is indicated by a black arrow. (C,D) Same as (A,B) for the AA bulge substrate. (E,F) Same as (A,B) for the homoduplex substrate. (G,H) Same as (A,B) for binding of MSH2/MSH6 to a G/T mismatched substrate.

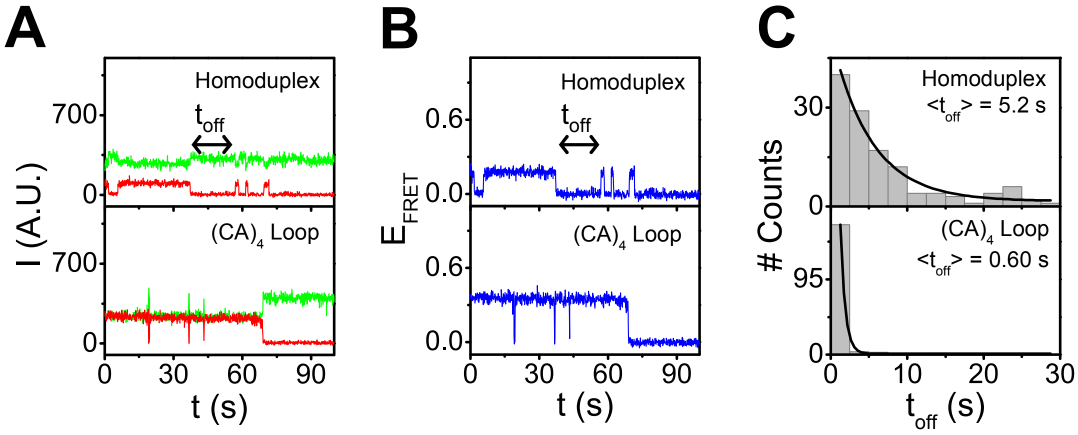


Figure 4.3

Typical single-molecule (A) intensity and (B) FRET efficiency time traces for homoduplex and (CA)$_4$ loop DNA in the absence of protein. The traces were taken in the presence of 2-mercaptoethanol, which induces photophysical effects in the acceptor. Recovery of the acceptor dye from triplet state occurs by close proximity (2-3 nm) of the donor dye [2]. In this experiment, the local environment of the donor and acceptor dyes on each substrate was identical. The single-molecule intensity traces show reversible transitions of the acceptor to a non-fluorescent state (t_{off}). The FRET efficiency of the t_{off} state is E ~ 0. (C) Histograms of the duration of the "off" state are plotted for both substrates and fitted to a single exponential. The duration of the t_{off} state is an order of magnitude longer for the homoduplex substrate. Thus, the donor and acceptor in the (CA)$_4$ loop substrate must come into close proximity.

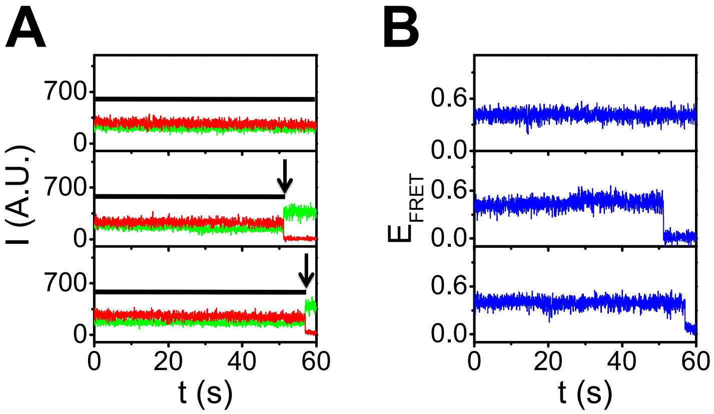


Figure 4.4
Typical traces showing the $(CA)_4$ loop bound to MSH2/MSH3. (A) The fluorescence emission intensities of the donor (green) and acceptor (red) show stable binding for >60 seconds (binding depicted by black bar), and the measurements are limited by acceptor photobleaching (marked with black arrows). (B) The calculated FRET efficiency is shown for each trace given in (A).

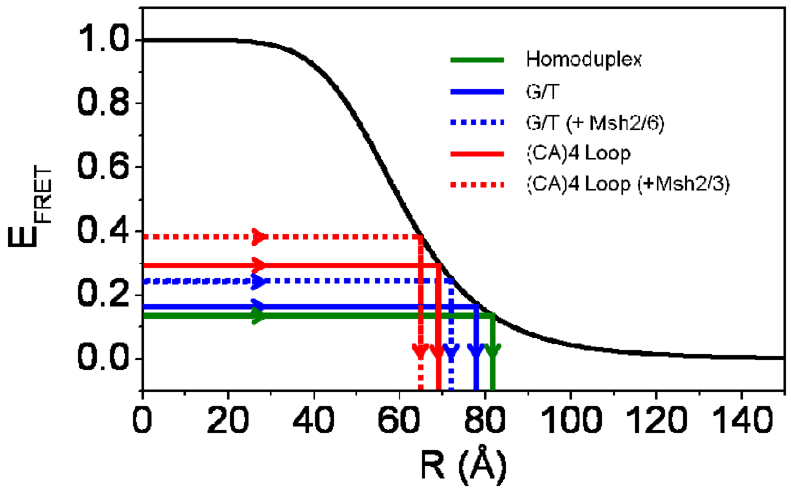


Figure 4.5
The FRET efficiency curve (black) shows the relationship between the measured FRET efficiency E_{FRET} and the distance between the donor and acceptor fluorophores R. The curve was generated from $E_{FRET} = (1+(R/R_0)^6)^{-1}$, using an R_0 value of 60.1 Å. The change in distance measurements for the $(CA)_4$ loop and G/T mismatch substrates were used to estimate the bending angle caused when MSH2/MSH3 and MSH2/MSH6 bind, respectively.

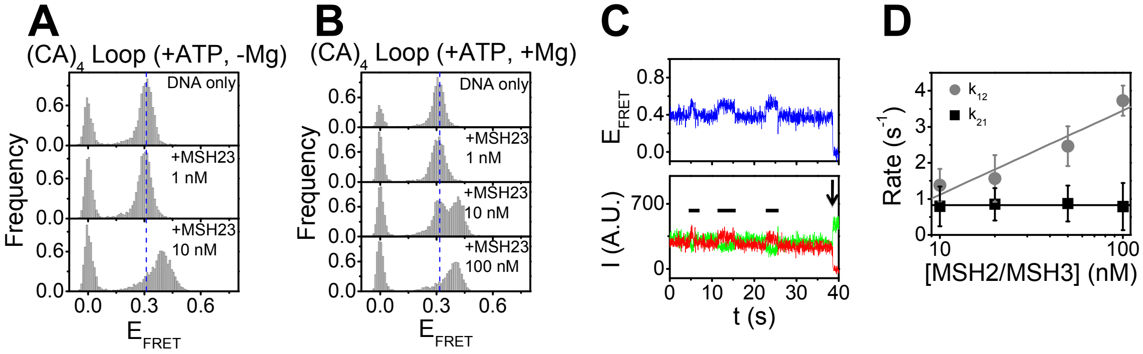


Figure 4.6

ATP increases dissociation of MSH2/MSH3 from a $(CA)_4$ loop substrate. FRET efficiency histograms show the effect of MSH2/MSH3 binding to the $(CA)_4$ loop substrate at different protein concentrations in the presence of 100 μM ATP (A) without magnesium and (B) with 5 mM magnesium. MSH2/MSH3 concentration is as indicated. (C) The dynamics of the $(CA)_4$ loop substrate upon MSH2/MSH3 binding in (B). Time traces of representative donor fluorescence (green, Cy3) and acceptor fluorescence (red, Cy5) and the corresponding calculated FRET efficiency (blue) are shown. The black lines denote the times of MSH2/MSH3 binding. The time of acceptor photobleaching is indicated by a black arrow. (D) Time traces like the one shown in (C) were analyzed for the two-state system using a hidden Markov model [3] to determine the average transition rates from initial to high and from high to initial FRET states. The transition rates from the initial to high FRET state depend on protein concentration (grey balls), while the transition rates from the high to the initial FRET state were independent of the protein concentration (black squares).

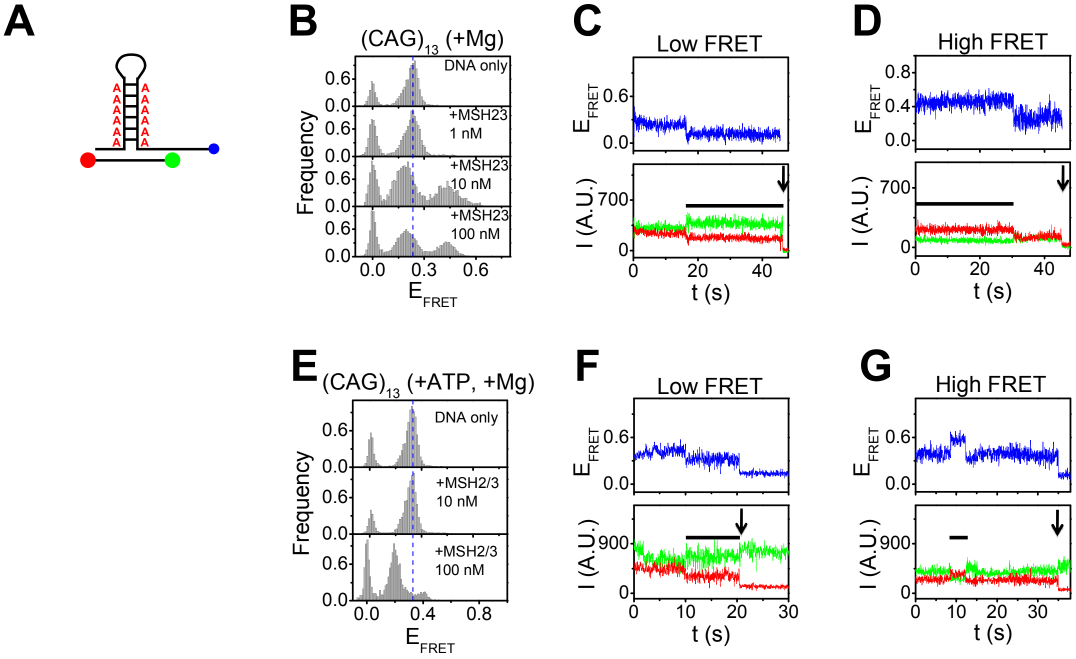


Figure 4.7

Repair-resistant $(CAG)_{13}$ hairpins trap nucleotide-bound MSH2/MSH3 in the low FRET state. (A) A loop of CAG repeats forms a hairpin with A/A mispaired bases every three nucleotides. In the schematic, the green ball is the Cy3 fluorophore, the red ball is the Cy5 fluorophore, and the blue ball is the biotin label. (B) The FRET efficiency histograms for MSH2/MSH3 binding to the $(CAG)_{13}$ hairpin in the absence of nucleotides show that MSH2/MSH3 binds this substrate in two discrete states—one with a higher FRET value and one with a lower FRET value than the DNA only conformation. Individual time traces of the emission intensities (donor, green; acceptor, red) and corresponding FRET values (blue) for the (C) low and (D) high FRET states. In (C) and (D), the black bars indicate the binding event, and the arrows indicate photobleaching of the donor dye. (E) Addition of ATP to (B) strongly reduces the relative abundance of the high FRET state compared to the low FRET state. The individual time traces indicate that (F) the low FRET state remains stable and (G) the time in the high FRET state is less in the presence of ATP than in the absence of ATP (D).

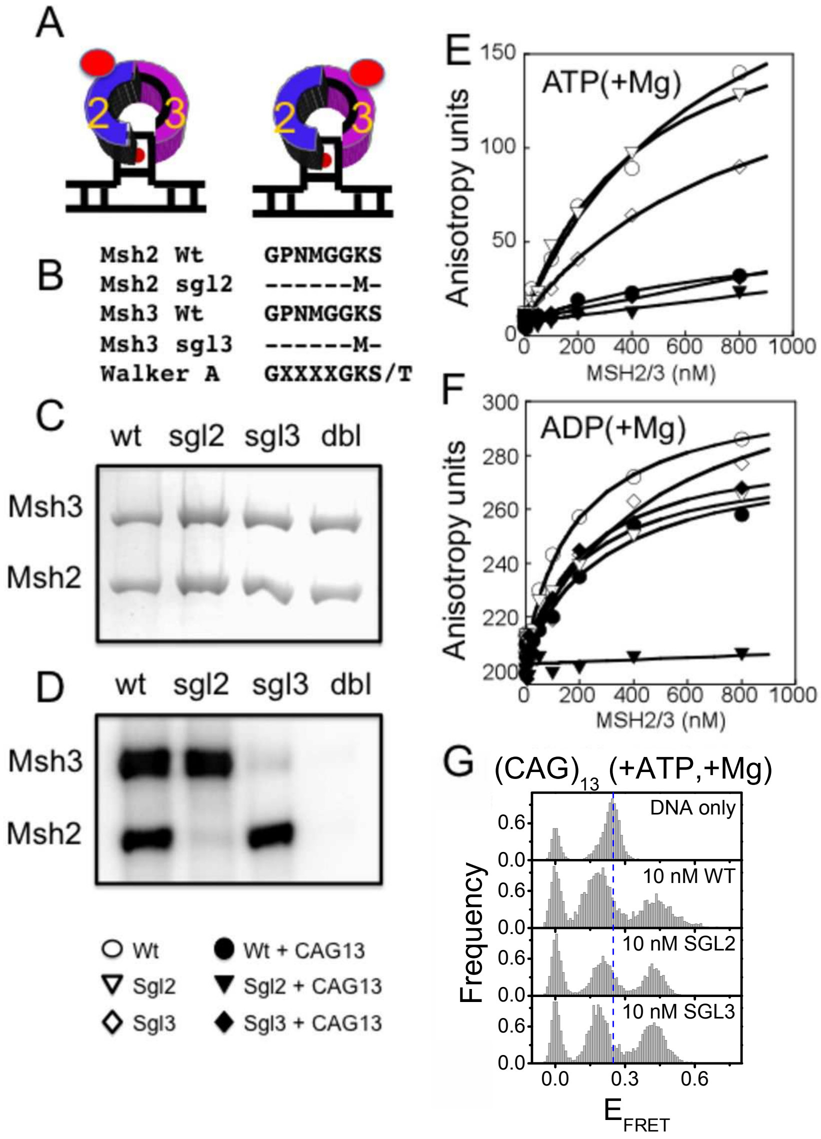
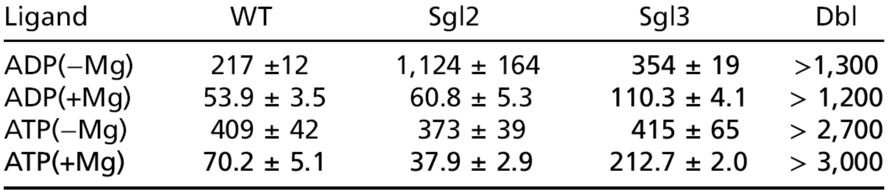


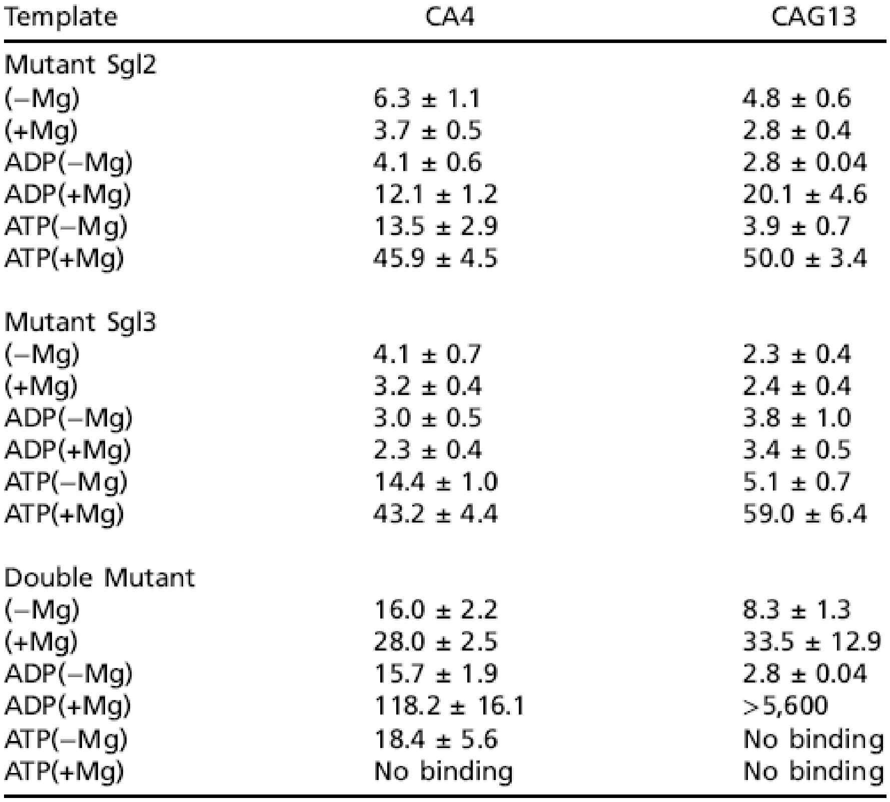
Figure 4.8

A $(CAG)_{13}$ hairpin binds only one nucleotide-bound MSH2/MSH3 complex but displays both high and low FRET states. (A) A schematic diagram of two possible MSH2/MSH3-$(CAG)_{13}$ hairpin complexes with nucleotide bound in either the MSH2 or MSH3 subunit is shown. (B) Sequences of the mutant MSH2 and MSH3 subunits aligned with the canonical Walker A box sequence motif of the wt MSH2/MSH3. The conserved lysine residue has been replaced with a methionine in the mutant proteins to destroy the ATP binding pocket. (C) Resolution of the wild-type and mutant MSH2/MSH3 on denaturing gels. The abbreviations are as follows: wt, wild-type MSH2/MSH3; sgl2, Walker A box mutations in the MSH2 subunit only; sgl3, Walker A box in the MSH3 subunit only; dbl, Walker A motif mutations in both subunits. (D) Binding of $[\alpha\text{-}^{32}P]$-ATP to wild-type and mutant MSH2/MSH3 proteins analyzed by UV-crosslinking followed by resolution on denaturing gel. Only intact nucleotide binding sites bind ATP efficiently. (E) Fluorescence anisotropy measurements of bodipy-labeled ATP binding to both wild-type and mutant MSH2/MSH3-$(CAG)_{13}$ hairpin complexes. (F) Mutation of the Walker A box motif in the MSH2 subunit only inhibits binding of bodipy-labeled ADP to a MSH2/MSH3-$(CAG)_{13}$ hairpin complex. (G) Association of nucleotide-bound wild-type and mutant MSH2/MSH3 to the $(CAG)_{13}$ templates result in high and low FRET states. ATP is retained poorly in the MSH3 subunit when bound to DNA. Thus, sgl3 and wild-type MSH2/MSH3 have similar binding activity in the presence of nucleotides. Sgl2 MSH2/MSH3 in the presence of nucleotides has the same binding activity as wt MSH2/MSH3 without bound nucleotide. ATP is 100 μM.



Ligand	WT	Sgl2	Sgl3	Dbl
ADP(−Mg)	217 ±12	1,124 ± 164	354 ± 19	>1,300
ADP(+Mg)	53.9 ± 3.5	60.8 ± 5.3	110.3 ± 4.1	> 1,200
ATP(−Mg)	409 ± 42	373 ± 39	415 ± 65	> 2,700
ATP(+Mg)	70.2 ± 5.1	37.9 ± 2.9	212.7 ± 2.0	> 3,000

Table 4.3
Nucleotide binding affinities of wild-type and mutant MSH2/MSH3 proteins determined by fluorescence anisotropy (K_D in nM).



Template	CA4	CAG13
Mutant Sgl2		
(−Mg)	6.3 ± 1.1	4.8 ± 0.6
(+Mg)	3.7 ± 0.5	2.8 ± 0.4
ADP(−Mg)	4.1 ± 0.6	2.8 ± 0.04
ADP(+Mg)	12.1 ± 1.2	20.1 ± 4.6
ATP(−Mg)	13.5 ± 2.9	3.9 ± 0.7
ATP(+Mg)	45.9 ± 4.5	50.0 ± 3.4
Mutant Sgl3		
(−Mg)	4.1 ± 0.7	2.3 ± 0.4
(+Mg)	3.2 ± 0.4	2.4 ± 0.4
ADP(−Mg)	3.0 ± 0.5	3.8 ± 1.0
ADP(+Mg)	2.3 ± 0.4	3.4 ± 0.5
ATP(−Mg)	14.4 ± 1.0	5.1 ± 0.7
ATP(+Mg)	43.2 ± 4.4	59.0 ± 6.4
Double Mutant		
(−Mg)	16.0 ± 2.2	8.3 ± 1.3
(+Mg)	28.0 ± 2.5	33.5 ± 12.9
ADP(−Mg)	15.7 ± 1.9	2.8 ± 0.04
ADP(+Mg)	118.2 ± 16.1	>5,600
ATP(−Mg)	18.4 ± 5.6	No binding
ATP(+Mg)	No binding	No binding

Table 4.4
DNA-binding affinities of mutant MSH2/MSH3 proteins determined by fluorescence anisotropy (K_D in nM).

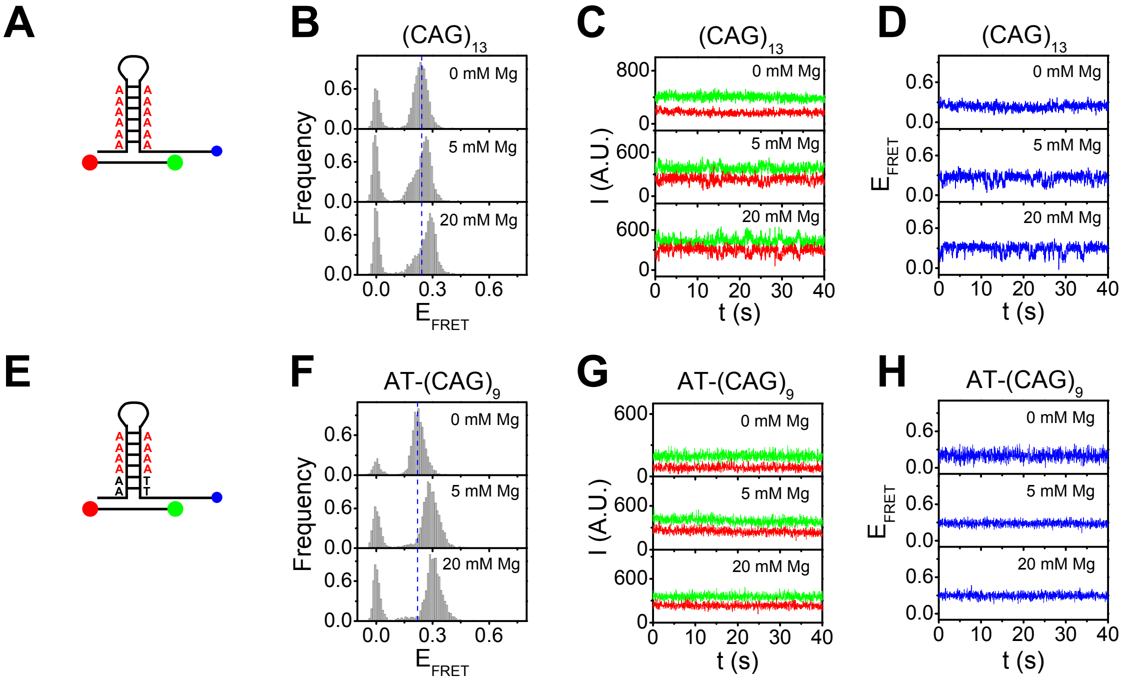


Figure 4.9
The (A) (CAG)$_{13}$ hairpin substrate intrinsically adopts two discrete conformational states in the absence of protein. (B) The FRET efficiency histograms of the (CAG)$_{13}$ substrate in the absence of protein are shown at different magnesium concentrations. Increasing the magnesium allows two discrete states to be resolved. (C) The single-molecule intensity time traces (donor, green; acceptor, red) show that few transitions are observed within the time resolution for concentrations of magnesium below 5 mM, and the transitions become more pronounced at 20 mM. (D) The corresponding FRET efficiency time traces (blue) for each magnesium concentration are shown. The (E) AT-(CAG)$_9$ hairpin substrate adopts a single conformational state in the absence of protein. (B) The FRET efficiency histograms of the AT-(CAG)$_9$ substrate in the absence of protein are shown at different magnesium concentrations. In the absence of magnesium, the FRET value of the AT-(CAG)$_9$ substrate is similar to the FRET value of the (CAG)$_{13}$ substrate in the same buffer condition. Adding magnesium increases the FRET value of the AT-(CAG)$_9$ hairpin substrate's conformational state, but there is still a single population of FRET values, and (G,H) the individual time traces show no intrinsic conformational dynamics for this substrate.

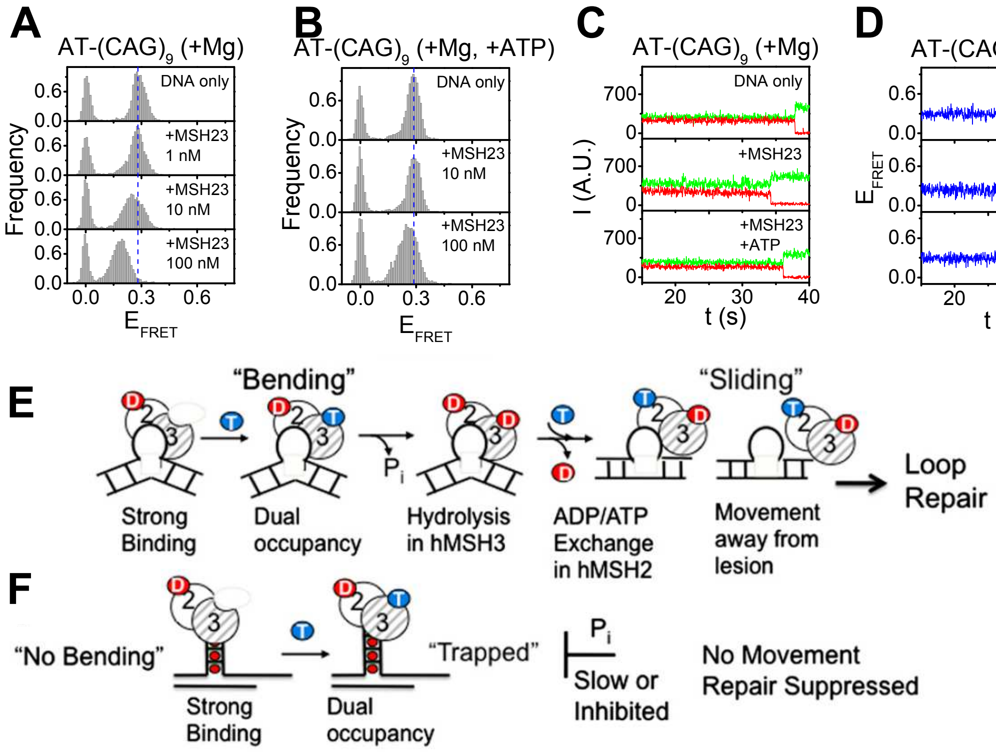


Figure 4.10
The junction of the AT-(CAG)$_9$ hairpin adopts only one stable three-way junction, from which MSH2/MSH3 does not dissociate. (A) The FRET efficiencies for binding of MSH2/MSH3 to the AT-(CAG)$_9$ hairpin substrate in the presence of magnesium results in a single low FRET population. (B) The FRET efficiency histograms for AT-(CAG)$_9$ in the presence of ATP under hydrolytic conditions are shown. The affinity of MSH2/MSH3 for the AT-(CAG)$_9$ substrate (+Mg) is reduced in the presence of ATP, but substrate binding results in the same low FRET state as observed in the absence of nucleotides. (C) Representative emission intensity time traces of the AT-(CAG)$_9$ substrate alone (+Mg, top), AT-(CAG)$_9$ bound to MSH2/MSH3 (-ATP, +Mg, middle), and AT-(CAG)$_9$ bound to MSH2/MSH3 in hydrolytic conditions (+ATP, +Mg, bottom) are shown. All traces are similar and display no dynamics (donor, green; acceptor, red). (D) The FRET efficiency time traces that correspond to the emission intensity time traces are shown. (E) The proposed model for conformational regulation of loop repair by MSH2/MSH3 at three-way DNA junctions is shown. In this model, the conformational flexibility of the substrate determines the possible binding modes of MSH2/MSH3. MSH2/MSH3 binds and bends a repair-competent loop, and downstream nucleotide hydrolysis and exchange causes MSH2/MSH3 to adopt a doubly bound form which verifies the mismatch and leaves the lesion to signal downstream repair by the MMR machinery. (F) Nucleotide-bound MSH2/MSH3 traps repair-resistant loops in a straightened conformation which is a non-productive complex, and MSH2/MSH3 cannot undergo the necessary biochemical steps to leave the lesion and initiate efficient repair by the MMR machinery. Successful DNA MMR couples DNA binding and ATP hydrolysis, and conformational trapping does not allow processing of ATP in the MSH3 subunit, preventing ADP/ATP exchange needed to leave the site. Circles with 2 and 3 represent the MSH2/MSH3 heterodimer. The red balls are ADP, and the blue balls are ATP.

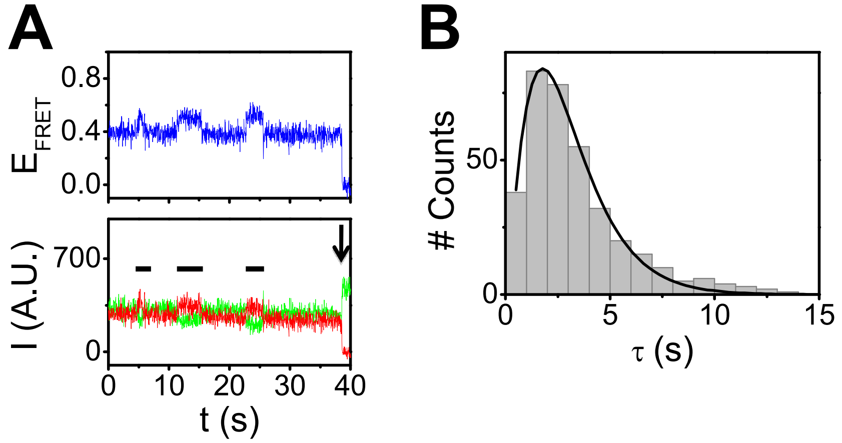


Figure 4.11
ATP increases dissociation of MSH2/MSH3 from a $(CA)_4$ loop substrate, and the dwell time of the high FRET states have non-exponential dwell-time distributions. (A) Time traces of representative donor fluorescence (green, Cy3) and acceptor fluorescence (red, Cy5) and the corresponding calculated FRET efficiency (blue) are shown. The black lines denote the dwell times of the high FRET states (MSH2/MSH3 binding). These traces are also discussed in **Figure 4.6**. (B) The distribution of dwell times (τ) shows a non-exponential distribution with a clear dip at short times (that are still much longer than the time resolution of the experiment, 32 ms). This data clearly indicates that the high FRET (or bound, bent) state of the $(CA)_4$ loop in the presence of ATP involves more than a single kinetic step (if there were a single kinetic step, we would observe an exponential distribution).

Chapter 5

The binding kinetics of E. coli MutS influence DNA mismatch selectivity[8]

Here, single-molecule FRET is used to observe binding and dissociation of *E. coli* MutS proteins to and from DNA substrates in real time. By comparing the binding kinetics of MutS in complex with the efficiently repaired G/T mismatch to the binding kinetics of MutS in complex with the less efficiently repaired C/T mismatch, we conclude that efficient repair requires a stable MutS-mismatch binding interaction in order to allow ATP uptake, and ATP binding but not hydrolysis allows MutS to form a sliding clamp that can interact with downstream repair proteins in the DNA mismatch repair pathway.

5.1 Introduction

The DNA mismatch repair system protects the genome from spontaneous mutations by recognizing and correcting a wide range of DNA mismatches, and much of what is known about the first step of the DNA mismatch recognition pathway, DNA mismatch recognition, has been derived from studies of *E. coli* MutS [17-19, 118]. *E. coli* MutS (hereafter referred to as MutS) functions to bind and signal the repair of base-base mismatches and 1-4 nucleotide loops [27, 28], but the biochemical activities of MutS and the mismatch repair efficiencies of each lesion are site-specific [27, 119]. The

---

[8] This work is preparation for publication:
J. E. Coats, Y. Lin, I. Rasnik. 'The binding kinetics of *E. coli* MutS influence DNA mismatch selectivity.'

DNA mismatch repair system of *E. coli* repairs G/T, A/C, G/G, and A/A mismatches with high efficiency *in vivo* [120-122] and *in vitro [27]*, and MutS has high affinity (K_D ~ 5-50 nM) for these mismatches. A/G, C/T, T/T, and C/C mismatches are repaired with lower efficiency than the previously mentioned mismatches, and MutS has intermediate affinity (K_D ~ 50-500 nM) for these mismatches. Since there is some correlation between the binding affinities of MutS for a mismatch and the corresponding efficiencies of mismatch repair, initial recognition by MutS proteins is thought to be a major player in the efficiency of the DNA mismatch repair pathway [27].

Crystallographic studies of MutS in complex with G/T, A/C, G/G, A/A and +T mismatches show MutS binds DNA as a homodimer and that in the DNA binding domain of one subunit, a conserved Phe-X-Glu motif near the N-terminal makes specific interactions with mismatches and bends the duplex (~60°) at the site of the lesion [20-22]. Crystallographic studies of the MutS homolog human MSH2/MSH6 (which also has the Phe-X-Glu motif) bound to a number of mismatches and chemically modified bases reveal that it also bends DNA at the site of a lesion [23]. Together, these studies have led to the ideas that protein-induced bending is an important component of DNA mismatch recognition and MutS proteins use a common binding mode to recognize a wide range of mismatches.

After binding a mismatch, MutS uses its ATPase activity to activate downstream repair proteins [17-19, 118]. Each subunit of the MutS dimer has a carboxyl-terminal ATP binding site that is required for the function of the protein, and upon binding ATP, MutS proteins undergo a conformational change to form a sliding clamp that is capable of moving along the helix [37, 40, 43-46]. Two models have been proposed in the literature to describe the role of ATP for the formation of the sliding clamp and movement of MutS along the DNA contour. (i) In the ATP-dependent translocation model, movement depends on ATP hydrolysis by the DNA-bound protein [34, 36, 37, 44, 46]. (ii) In the molecular switch model, ATP binding results in a conformational change to a sliding clamp that moves along the DNA backbone via diffusion [39, 43, 45, 47].

Since binding, bending, and dissociation by MutS proteins are thought to play a large role in the efficiency of the DNA mismatch repair pathway, we used single-molecule FRET to compare the conformational dynamics and binding kinetics of MutS in complex with a well-repaired G/T mismatch to the conformational dynamics and binding kinetics of MutS in complex with a C/T mismatch that is repaired less efficiently. We found that in the absence of added nucleotides, MutS binds the G/T mismatch stably, while MutS binds only transiently to the C/T mismatch. When ATP is present, MutS dissociates from the G/T mismatch with a non-exponential dwell-time distribution for the binding times which indicates the presence of more than one kinetic step. The presence of ATP also increases the dissociation rate of MutS from the C/T mismatch, but an exponential dwell-time distribution for the binding times indicates that dissociation from the C/T mismatch occurs in a single kinetic step. Our results support a model in which stable mismatch binding allows ATP uptake, allowing MutS to form the sliding clamp necessary for downstream repair.

5.2 Results

smFRET experiment design

The DNA substrates used in the smFRET experiments were 18 base pair duplexes that were constructed by annealing two oligonucleotide strands. One strand was labeled on the 5' end with a donor (Cy3) fluorophore, and the other strand was labeled on the 5' end with an acceptor (Cy5) fluorophore and on the 3' end with a biotin molecule for surface immobilization. In addition, the acceptor-labeled strand had a 15 nt polydT tail on the 3' end between the duplex and the biotin molecule to prevent potential interactions of the fluorophores with the streptavidin-coated surface (**Figure 5.1**). The heteroduplex DNA substrates had centrally located G/T or C/T mismatches (named GT18 and CT18, respectively), while a control homoduplex DNA substrate had a centrally located G/C base pair (named GC18). Besides the centrally located bases, the sequence of each of the duplexes was identical (**Figure 5.1**). The length of the duplexes was designed to be more than sufficient to accommodate MutS, as the reported footprint of MutS on mismatched DNA is 13 base pairs [20]. The design of the DNA substrates is such than the donor and acceptor fluorophores are on opposite ends of the duplexes, allowing detection of protein-induced bending that causes a decrease in the relative end-to-end distance between each FRET pair by a change in the efficiency of energy transfer between the fluorophores. Detecting FRET at the single-molecule level allows observation of the dynamics of individual molecules and permits analysis of individual populations of conformers to obtain quantitative kinetic information on the intermediate steps of molecular processes. Prior to the single-molecule FRET experiments, *E. coli* MutS was expressed and purified (see **Appendix B.1** for protocol), and ensemble DNA binding and ATPase experiments were performed to demonstrate that MutS has characteristic DNA binding and ATPase activity with the DNA substrates (see **Appendix B.2** for results).

MutS binds to a G/T mismatch in a stable, high FRET state

Since MutS binds G/T mismatches more tightly than any other mismatch and the DNA mismatch repair system repairs G/T mismatches more efficiently than any other mismatch, the majority of studies of mismatch binding by MutS have been done with G/T mismatched DNA. Thus, we first used single-molecule FRET to characterize the conformational dynamics of MutS in complex with the G/T mismatched DNA substrate, GT18.

The distribution of FRET efficiency (E_{FRET}) values and the dynamics of individual GT18 molecules were determined for hundreds of FRET pairs in the absence and presence of MutS. In the absence of protein, the GT18 substrate has a single narrow peak at E_{FRET} ~0.17 (**Figure 5.2A**, DNA only; the peak at E_{FRET} ~0 represents substrates with an inactive acceptor). The dynamics of each individual FRET pair were observed by plotting the time traces of the donor (Cy3) and acceptor (Cy5) emission intensities and the corresponding time traces of the FRET efficiencies (**Figures 5.2B**), and there were no observable FRET transitions for any of the GT18 molecules when no protein was present.

Addition of MutS to the GT18 substrate led to the appearance of a new FRET efficiency peak, centered at E_{FRET} ~0.29 (**Figure 5.2A**, +0.5, 1, and 20 nM MutS). The increase in FRET efficiency upon protein addition indicates that the protein binds and bends the DNA by a single 55° angle, bringing the ends of the DNA duplex closer together. This measured angle is in agreement with crystallographic studies that report that bacterial MutS bends a G/T mismatch with an angle of ~60° [20-22]. The MutS bound population (at E_{FRET} ~0.29) increased with protein concentration until the entire population of GT18 molecules were bound. Consistent with high affinity binding (the K_D of MutS in complex with a G/T mismatched DNA substrate is in the low nanomolar range [27, 28]), the high FRET efficiency GT18 state was saturated at 20 nM MutS.

The time trajectories for the emission intensities and corresponding FRET efficiency values show that MutS binds the GT18 substrate stably in the high FRET efficiency state (**Figure 5.2C**). The time trajectories shown in **Figure 5.2C** show that MutS stays bound until acceptor photobleaching. Since the average lifetime of the MutS-GT18 complex was longer than the observation time of the measurement (~80 seconds), few molecules showed dissociation to the DNA only state. The molecules that do show MutS dissociation before fluorophore photobleaching show a single-step transition from the bound state (E_{FRET} ~0.29) to the DNA only state (E_{FRET} ~0.17) (**Figure 5.3A**). Likewise, since the MutS-GT18 experiments had stable binding that lasted longer than the observation time of the measurement, few MutS-GT18 association events were seen in the steady-state experiments (**Figure 5.2**). Flowing MutS into a sample chamber containing only GT18, however, demonstrates that MutS binds the GT18 substrate in a single-step transition (**Figure 5.3B**). Furthermore, the MutS-GT18 association times for the flow experiments can be fit to a single-exponential function, consistent with a single kinetic step for the MutS-GT18 binding/bending interaction (**Figure 5.3C**), meaning bending occurs at the instant of the MutS-GT18 complex formation.

As a control, we evaluated the interaction of MutS with a homoduplex DNA substrate. Previous studies have reported that MutS has weak affinity and rapidly dissociates from homoduplex DNA. Consistent with these studies, no high FRET state was observed for the homoduplex DNA (GC18) in the presence of high MutS concentrations (**Figure 5.2D-F**)

MutS binds to a C/T mismatch transiently in a high FRET state

Next, we studied the conformational dynamics of MutS in complex with the C/T mismatched DNA substrate, CT18. As with the GT18 DNA experiments, the distribution of FRET efficiency (E_{FRET}) values and the dynamics of individual CT18 molecules were determined for hundreds of FRET pairs in the absence and presence of MutS, and in the absence of protein, the CT18 substrate has a single narrow peak at E_{FRET} ~0.17 (**Figure 5.4A**, DNA only). The dynamics of each individual FRET pair were observed by plotting the time traces of the donor and acceptor emission intensities and the corresponding time traces of the FRET efficiencies (**Figure 5.4B**), and there were no observable FRET efficiency transitions for any of the CT18 molecules when no protein was present.

Addition of MutS to the CT18 substrate led to the appearance of a new FRET efficiency peak, centered at E_{FRET} ~ 0.27 (**Figure 5.4A**, + 10, 50, and 100 nM MutS), so MutS bends the C/T mismatch with an angle of ~50°, which is similar to the angle that it bends the G/T mismatch. The bound, bent population (at E_{FRET} ~ 0.27) increased with increasing protein concentration until the entire population of GT18 molecules were bound. Consistent with ensemble experiments which have shown that MutS binds a C/T mismatch with lower affinity than it binds a G/T mismatch, more MutS was required to saturate binding of the C/T mismatch (**Figure 5.4A**) than was required to saturate binding for the G/T mismatch (**Figure 5.2A**).

The single-molecule time traces show that MutS binds the CT18 substrate in the high FRET efficiency state (E_{FRET} ~ 0.27) transiently, as the average dwell time of the bound state (t_{off}) was ~3.0 seconds, irrespective of MutS concentration (**Figure 5.4C**).

ATP accelerates dissociation of MutS from a G/T mismatch, and the dissociation kinetics are non-exponential

Since MutS proteins couple DNA binding and ATP binding/hydrolysis to initiate downstream repair, we tested the effects of ATP on the conformational dynamics of the MutS-bound GT18 substrate. Addition of MutS to the GT18 substrate in hydrolytic conditions (+ATP, +Mg), decreased the affinity of MutS for the GT18 DNA (**Figure 5.5A**, +20 nM MutS, +ATP) relative to adding MutS to the GT18 substrate in the absence of ATP (-ATP, +Mg) (**Figure 5.5A**, +20 nM MutS, -ATP). More specifically, the population of molecules in the high FRET state decreases with increasing ATP at a given MutS concentration (**Figure 5.5A**, +20 nM MutS, +50-500 μM ATP).

The addition of ATP results in a striking alteration in the dynamics of the MutS-GT18 complex. The single-molecule emission intensity and FRET efficiency traces show multiple transitions between high FRET states and low FRET states—consistent with MutS binding and dissociation, respectively (**Figure 5.5B**). Hidden Markov analysis [3] of FRET efficiency traces at each ATP concentration tested (50, 100, and 500 μM) reveals that while the low FRET dwell times (t_{on}) are not affected by the concentration of ATP present (**Figure 5.5C**), the high FRET dwell times (t_{off}) decreased with increasing ATP concentration (**Figure 5.5D-E**). The dwell time distributions of the low FRET states (**Figure 5.5C**) were fit to a single-exponential function, indicating that the MutS binds and bends the mismatch in a single kinetic step. The dwell-time distributions of the high FRET states were all non-exponential (**Figure 5.5D**), demonstrating the presence of multiple kinetic steps for ATP-dependent dissociation.

ATP accelerates dissociation of MutS from a C/T mismatch, but the dissociation kinetics are exponential

Like for the MutS-GT18 complex, the kinetics of the MutS-CT18 complex were also altered by the presence of ATP. The addition of ATP decreased the affinity of MutS for the CT18 DNA relative to adding MutS to the CT18 DNA substrate in the absence of ATP (**Figure 5.6A**). Like in the MutS-GT18 interaction, the population of molecules in the high FRET state decreases with increasing ATP at a given MutS concentration (**Figure 5.6A**, +50 nM MutS, +50-500 μM ATP).

Hidden Markov analysis [3] of traces at each ATP concentration tested (50, 100, and 500 μM) reveals that while the binding time (t_{on}) for the MutS-CT18 interaction is not affected by the concentration of ATP present (**Figure 5.6C**), the dissociation time time (t_{off}) for MutS-CT18 dissociation is decreased with increasing ATP concentration (**Figure 5.6D-E**).

Hidden Markov analysis [3] of FRET efficiency traces at each ATP concentration tested (50, 100, and 500 μM) reveals that while the low FRET dwell times (t_{on}) are not affected by the concentration of ATP present (**Figure 5.5C**), the high FRET dwell times (t_{off}) decreased with increasing ATP concentration (**Figure 5.5D-E**). Both the dwell time distributions of the low FRET states (**Figure 5.5C**) and the high FRET states (**Figure 5.5D**) were fit to a single-exponential function, indicating that the MutS binds and bends the mismatch in a single kinetic step and then dissociates from the mismatch in a single kinetic step. Thus, the dissociation kinetics of the MutS-GT18 and the MutS-CT18 complexes were distinct.

ATP binding, but not hydrolysis, leads to accelerated dissociation of MutS from a G/T mismatch

Although it is well established that ATP accelerates the dissociation of MutS from mismatched DNA, the role of ATP in the mechanism of MutS dissociation has been disputed in the literature. Evidence exists for both an ATP-dependent translocation model [34, 36, 37, 44, 46] that requires ATP hydrolysis to move away from this mismatch and a molecular switch model [39, 43, 45, 47] in which ATP binding induces a conformational change in the protein, allowing the protein to leave the mismatch via diffusion. To determine the role of ATP in accelerating the dissociation of the MutS-GT18 complex, we created a mutant of MutS that can bind but not hydrolyze ATP [33]. The mutant, named MutS-E694A, was created by using site-directed mutagenesis to change the glutamic acid-694 of MutS's ATP binding site to an alanine. MutS-E694A, like wild-type MutS, binds GT18 DNA with low nanomolar affinity, and it bends DNA in a stable kinked state (with $E_{FRET} \sim 0.29$), but it has no steady-state ATPase activity (see **Appendix B.3** for characterization of MutS-E694A).

Since MutS-E694A is not able to turn over ATP that it binds, experiments with MutS-E694A were done in non-steady state conditions. Wild-type MutS or MutS-E694A was pre-bound to GT18 DNA, and a flow system was used to flow buffer (with or without ATP) into the sample chamber. When buffer was flowed in the absence of ATP, MutS and MutS-E694A proteins did not dissociate from the GT18 DNA during the observation time of the experiment (~60 s) (**Figure 5.7A**). However, when buffer containing 100 μM ATP was flowed into a sample chamber containing either pre-bound wild-type MutS or MutS-E694A, the protein dissociated within seconds (average dissociation times: $t_{\text{off,MutS}}$ ~4.4 s, $t_{\text{off,MutS-E694A}}$ ~4.7 s) (**Figure 5.7B-C**). Since the dissociation times were similar for wild-type MutS and the hydrolysis-deficient MutS-E694A, ATP binding, but not dissociation is required for the ATP-dependent dissociation observed for the MutS-GT18 complex.

5.3 Discussion

MutS proteins initiate the DNA mismatch repair pathway by locating mismatches and communicating the presence of mismatches to other mismatch repair proteins, but many details of the mismatch recognition process are not well understood [17-19, 118]. In particular, the mechanism by which mismatch binding is connected to downstream repair remains enigmatic [17-19, 118]. In this study, we used single-molecule FRET to compare the binding kinetics of MutS in complex with a G/T mismatch, which is repaired efficiently, and a C/T mismatch, which is repaired with lower efficiency, in order to determine molecular details underlying mismatch recognition that leads to efficient repair.

From a conformational perspective, we show that MutS binds the efficiently-repaired G/T mismatch and the less efficiently-repaired C/T mismatch in similar conformations, kinking the mismatched substrates with angles of 55° and 50° for the G/T and C/T mismatches, respectively. Since the MutS-GT18 and MutS-CT18 complexes are conformationally similar, the conformational flexibilities of these MutS-mismatch interactions cannot explain the differences in their mismatch repair efficiencies.

From a dynamical viewpoint, MutS interacts with the efficiently-repaired G/T mismatch and the less efficiently-repaired C/T mismatch very differently. In the absence of nucleotides, the typical MutS-GT18 interaction lasts longer than the observation time of the experiment (**Figure 5.2C**), so MutS binds stably to the GT18 DNA substrate. In contrast, MutS interacts only transiently with the CT18 DNA (t_{off} ~ 3.0 s). Thus, even though MutS bends the G/T and C/T mismatches with similar angles, the binding stability is specific for each mismatch. The presence of ATP decreases the duration of the MutS-mismatch interaction for both mismatched substrates, but the effect of ATP on the dissociation kinetics was unique. ATP-bound MutS dissociates from the G/T mismatch with non-exponential dwell-time kinetics (**Figure 5.5D**), while ATP-bound MutS dissociates from the C/T mismatch with exponential dwell-time kinetics (**Figure 5.6D**). Thus, ATP-dependent dissociation of the MutS-GT18 complex occurs in more than one kinetic step, while MutS dissociates from the CT18 DNA in a single kinetic step. In addition, the experiments with MutS-E694A, which binds but does not hydrolyze ATP, demonstrate that ATP binding, but not hydrolysis, is needed for ATP-catalyzed dissociation of the MutS-GT18 complex.

Together, the results of this study support a model in which stable MutS-mismatch binding is required to allow ATP binding by MutS, and ATP binding induces a conformational change in MutS, allowing it to move away from the mismatch and interact with other mismatch repair proteins. Unstable MutS-mismatch complexes have a high probability of dissociating before ATP binding occurs, so many non-productive binding interactions may occur before MutS binds ATP and forms a sliding clamp capable of initiating downstream repair.

5.4 Experimental Procedures[9]

Cloning, expression, and purification of E. coli MutS

The *E. coli mutS* gene was amplified by PCR from *E. coli* K-12 MG1655 (Yale Coli Genetic Stock Center) genomic DNA. The PCR product was cloned into a TA cloning vector (pGEMT, Promega), amplified in JM109 recombinant cells, and inserted into the gene expression vector (pET28a, Novagen) which was modified with the replacement of the kanamycin resistance gene with an ampicillin resistance gene. The mutS insert was sequenced *(*Agencourt Bioscience Corp.) to ensure the integrity of the DNA sequence. pET28a/*mutS* was transformed into *E. coli* HMS174 (λDE3) cells by the heat shock method and spread on LB plates containing 100 μg/ml ampicillin. After an overnight incubation, a single colony was inoculated into LB broth supplemented with 100 μg/ml ampicillin. The cells were incubated with vigorous shaking at 37 °C until the absorbance of the cells at 600 nm reached 0.6, and gene expression was induced by addition of 1 mM IPTG. Cells were harvested after 3 hours by centrifugation at 5,000 × g. The cell pellets were resuspended in 35 ml/liter of culture of lysis buffer (20 mM HEPES [pH 7.8], 100 mM NaCl, 1 mM EDTA, 10% glycerol). The cells were lysed using a French press, and the cell debris was removed by centrifugation (10,000 × g for 20 min at 4°C). The supernatant was applied to a lysis buffer-equilibrated metal affinity resin (TALON cobalt affinity resin; Clontech). The resin was washed extensively with lysis buffer (~10 column volumes) and then washed with lysis buffer containing 10 mM imidazole (~3 mL). The bound protein was eluted with lysis buffer containing 300 mM imidazole, and the eluted protein was dialyzed against buffer A (20 mM HEPES pH 7.8 and 250 mM NaCl). For further purification, the dialyzed protein was applied to an anion-exchange column (HiTrap Q HP; GE Healthcare) equilibrated with buffer A, and the column was washed with five column volumes of buffer A. The bound proteins were eluted with buffer B (20 mM HEPES pH 7.8 and 400 mM NaCl). Protein fractions judged to be 95% pure or greater by Coomassie staining after SDS-gel electrophoresis were dialyzed against MutS dialysis buffer (20 mM HEPES [pH 7.8], 100 mM NaCl, 1 mM EDTA, 30% glycerol, and 1 mM dithiothreitol [DTT]). The protein concentration (in dimers) was determined by the Bradford protein assay kit (Bio-Rad) using bovine serum albumin as the standard.

---

[9] Cloning was done by Y.L. and J.E.C. Protein purification, protein characterization by ensemble assays, and single-molecule FRET experiments were done by J.E.C.

Site-directed Mutagenesis

The E694A mutant was constructed by changing the GAG codon (Glu 694) to GCG (Ala). Mutations were introduced to pET28*/mutS* by site-directed mutagenesis using a QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Stratagene). The 33 nt primer used in the PCR reaction had the following sequence: 5' CTG GTG CTG ATG GAT **GCG** ATT GGG CGC GGA ACG 3'.

Oligonucleotides

Oligonucleotides were purchased from Integrated DNA Technologies (IDT). Complementary strands were annealed in a 1:1.2 ratio by heating the strands at 80°C for three minutes and cooling them slowly to room temperature. Substrates used for ensemble fluorescence studies were purified using PAGE. The sequences are as follows:

GT18
5’ Cy3 TGG CGA CGG TAG CGA GGC 3’
5’ Cy5 GCC TCG CTG CCG TCG CCA TTT TTT TTT TTT TTT biotin 3’
GC18
5’ Cy3 TGG CGA CGG CAG CGA GGC 3’
5’ Cy5 GCC TCG CTG CCG TCG CCA TTT TTT TTT TTT TTT biotin 3’
CT18
5’ Cy3 TGG CGA CGG TAG CGA GGC 3’
5’ Cy5 GCC TCG CTC CCG TCG CCA TTT TTT TTT TTT TTT biotin 3’

Single-molecule FRET Experiments

For single molecule measurements, clean quartz slides and glass coverslips were coated with polyethylene glycol (PEG) and assembled with double-sided tape and epoxy to create sample chambers ~5 mm wide, ~20 mm long, and ~0.1 mm tall, and buffer exchange was accomplished through holes on each end of the sample chamber. Chamber surfaces were prepared for DNA immobilization by adding 0.25 mg/ml streptavidin (Molecular Probes) to the slide chamber. Subsequently, 20-50 pM biotinylated DNA was added to the streptavidin-coated chamber. Samples were imaged in MutS single molecule imaging buffer (20 mM HEPES [pH 7.8], 50 mM NaCl, 5 mM MgCl2, 0.1 mg/ml BSA, 10% glycerol, 0.4% (w/v) glucose, 1 mg/ml glucose oxidase, 0.04 mg/ml catalase, and saturated Trolox®), and MutS and nucleotides were included as indicated in the text. The experiments were performed using a prism-type total internal reflection (TIR) microscope that was built on an Olympus IX71 inverted microscope and allows for the simultaneous observation of 200-300 molecules at a time. Surface-immobilized donor (Cy3) fluorophores were excited with a 532 nm Nd:YAG laser (CrysalLaser), and the fluorescence emission of the donor and acceptor dyes was collected with a 60x water objective (1.2 NA, Olympus). The donor and acceptor emission intensities were separated using a dichroic mirror and recorded with a CCD camera (iXon, Andor Technology) with 32 ms time resolution. FRET efficiency for each individual pair was calculated as $I_A/(I_A+I_D)$ (where I_A is the acceptor emission and I_D is the donor emission) and corrected for crosstalk. For the flow experiments, a 0.5 mL tube was glued to the inlet hole of the slide and a syringe needle was glued to the outlet hole of the slide, and buffer was pulled through using a syringe pump controller (Harvard Apparatus).

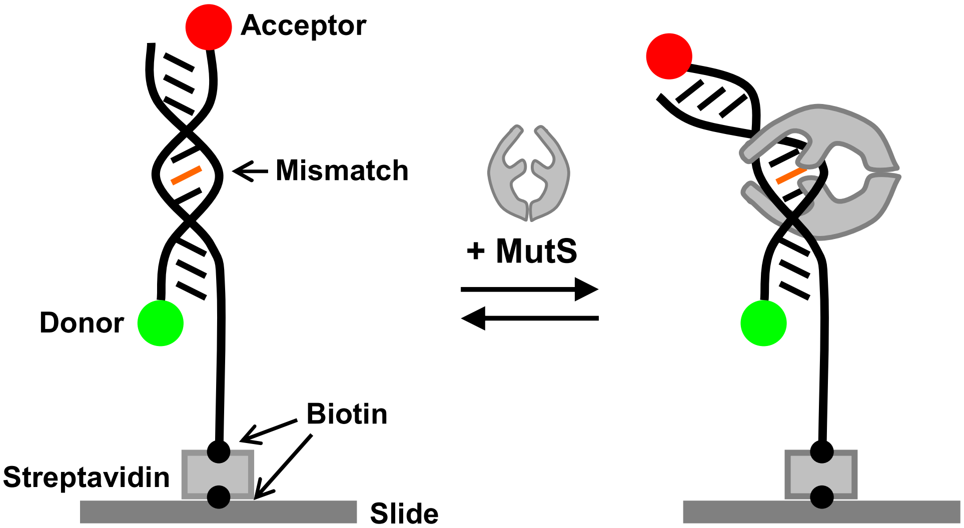

Acceptor
Mismatch
Donor
+ MutS
Biotin
Streptavidin
Slide

Figure 5.1
Schematic representations of mismatched DNA immobilized to a quartz slide (dark grey line) (left) without MutS and (right) with MutS are shown. The green dot is the donor (Cy3) fluorophore and the red dot is the acceptor (Cy5) fluorophore. The black dots represent the biotin molecules. Upon binding a mismatch, MutS bends the DNA, bringing the fluorophores closer together. Thus, MutS binding results in an increase in the FRET efficiency of the DNA molecule.

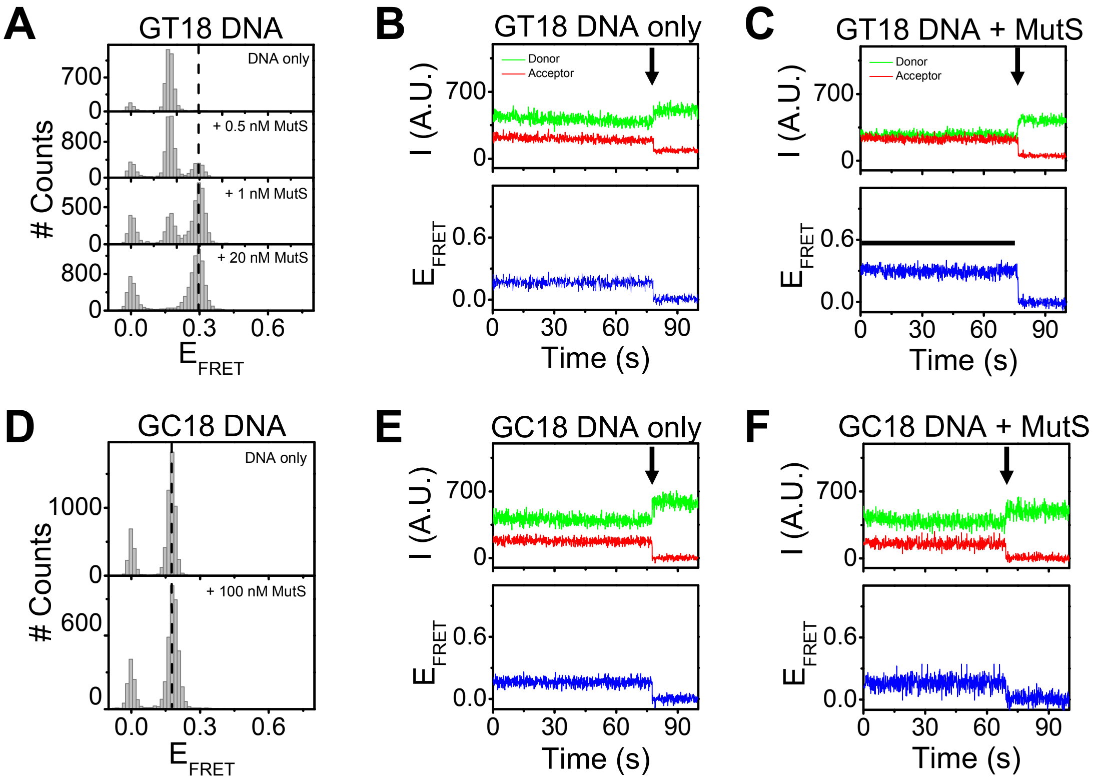


Figure 5.2

(A) FRET efficiency (E_{FRET}) histograms for G/T mismatched DNA (GT18) in the absence and presence of MutS are shown. Representative intensity (I) and corresponding FRET efficiency (E_{FRET}) time traces for GT18 in the (B) absence of protein and (C) presence of protein are shown. The black line highlights the MutS bound state. (D) FRET efficiency (E_{FRET}) histograms for homoduplex DNA (GC18) in the absence and presence of MutS are shown. Representative intensity (I) and corresponding FRET efficiency (E_{FRET}) time traces for GC18 in the (E) absence of protein and (F) presence of protein are shown. In all intensity traces, the black arrows denote the instant of acceptor photobleaching.

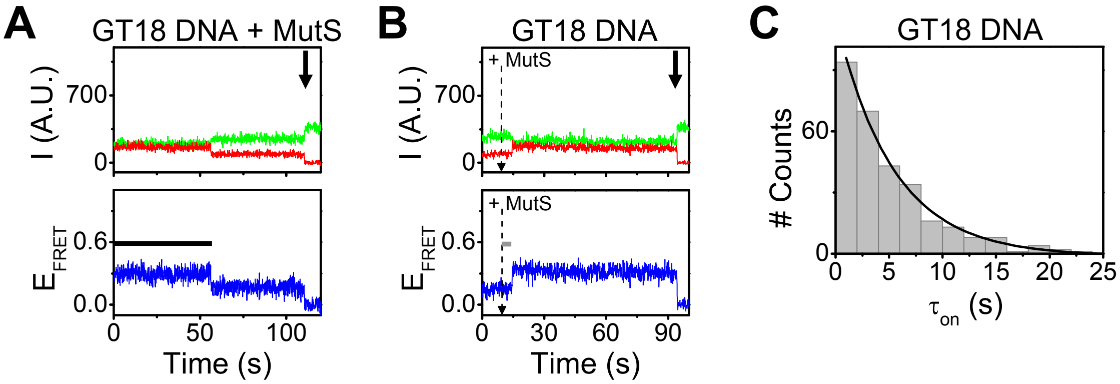


Figure 5.3
(A) The intensity (I) and corresponding FRET efficiency (E_{FRET}) time trace shows that MutS dissociates directly from the bound state (E_{FRET} ~0.29) to the DNA only state (E_{FRET} ~0.17) in a single step. The black line highlights the MutS bound state. (B) The intensity (I) and corresponding FRET efficiency (E_{FRET}) time trace shows that when 50 nM MutS is flowed into the sample chamber (at time = 10 s, dotted line) containing GT18, MutS binds the DNA within seconds in a single step from the DNA only state (E_{FRET} ~0.17) to the MutS bound state (E_{FRET} ~0.29). The grey line highlights the time it took MutS to bind the GT18 DNA, and (C) the corresponding binding time distribution for MutS binding to GT18 shows that the binding times fit a single exponential decay, indicating one kinetic step in the DNA binding reaction (the average binding time <τon> ~5.3 s). In both (A) and (B), the black arrows denote the instant of acceptor photobleaching.

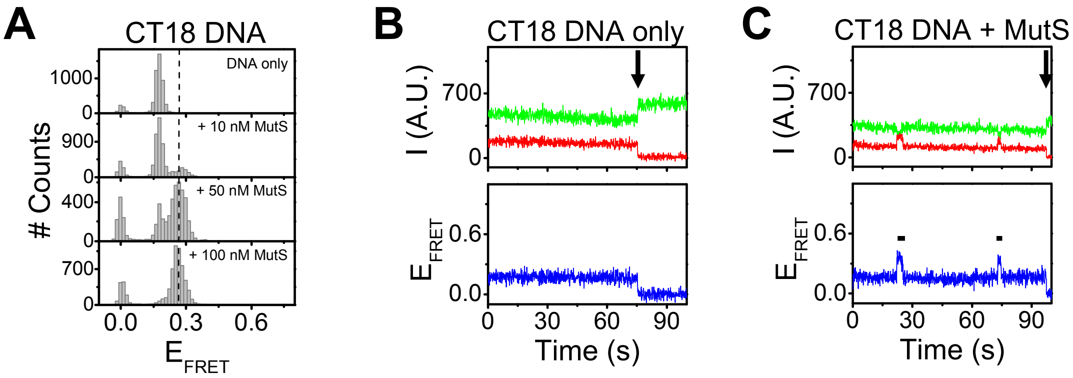


Figure 5.4
(A) FRET efficiency (E_{FRET}) histograms for C/T mismatched DNA (CT18) in the absence and presence of MutS are shown. Representative intensity (I) and corresponding FRET efficiency (E_{FRET}) time traces for CT18 in the (B) absence of protein and (C) presence of protein are shown. The black lines highlight the MutS bound state. The black arrows denote the instant of acceptor photobleaching.

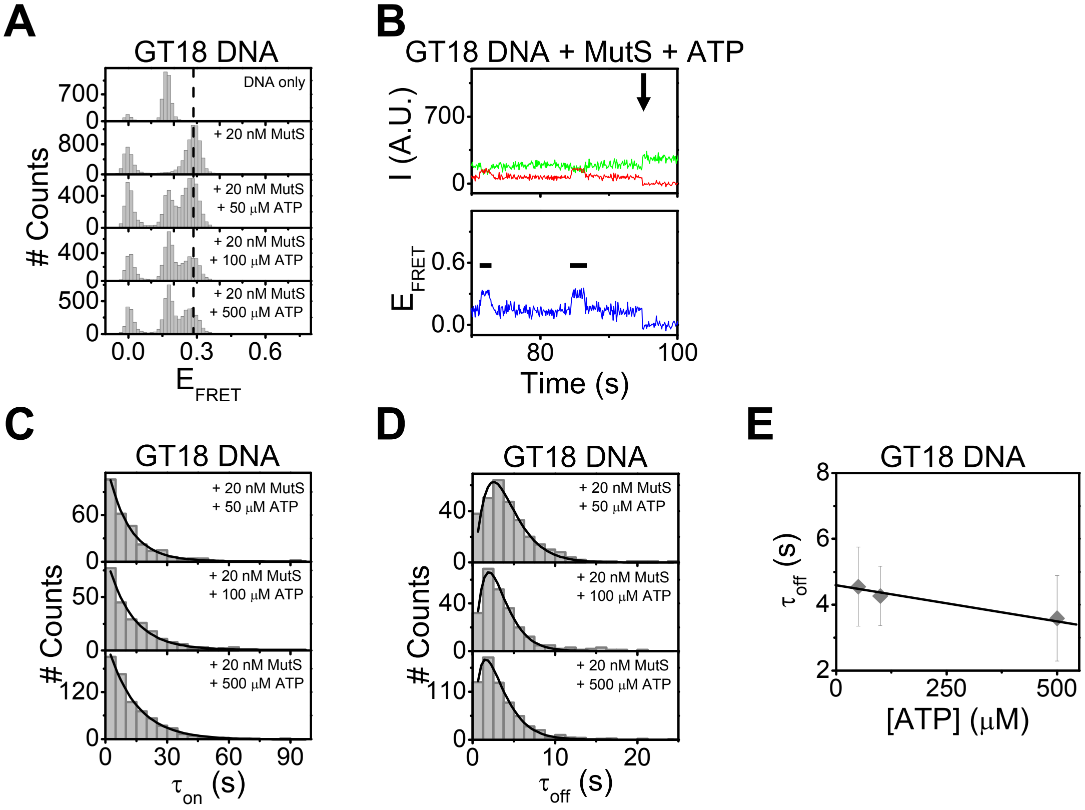


Figure 5.5

(A) FRET efficiency (E_{FRET}) histograms for the G/T mismatched DNA (GT18) in the absence of MutS and in the presence of 20 nM MutS and 0-500 μM ATP are shown. (B) A representative single-molecule intensity time trace and the corresponding FRET efficiency time trace for GT18 DNA in the presence of 20 nM MutS and 100 μM ATP is shown. The black lines highlight MutS binding events. The black arrow denotes the instant of acceptor photobleaching. (C) The distributions of times (τ_{on}) between MutS binding events (to GT18) at 50-500 μM ATP are shown. (D) The distributions of MutS-GT18 binding times (τ_{off}) for 20 nM MutS at 50-500 μM ATP are shown. (E) The average MutS binding times (τ_{off}) from the histograms shown in (D) are plotted as a function of ATP concentration.

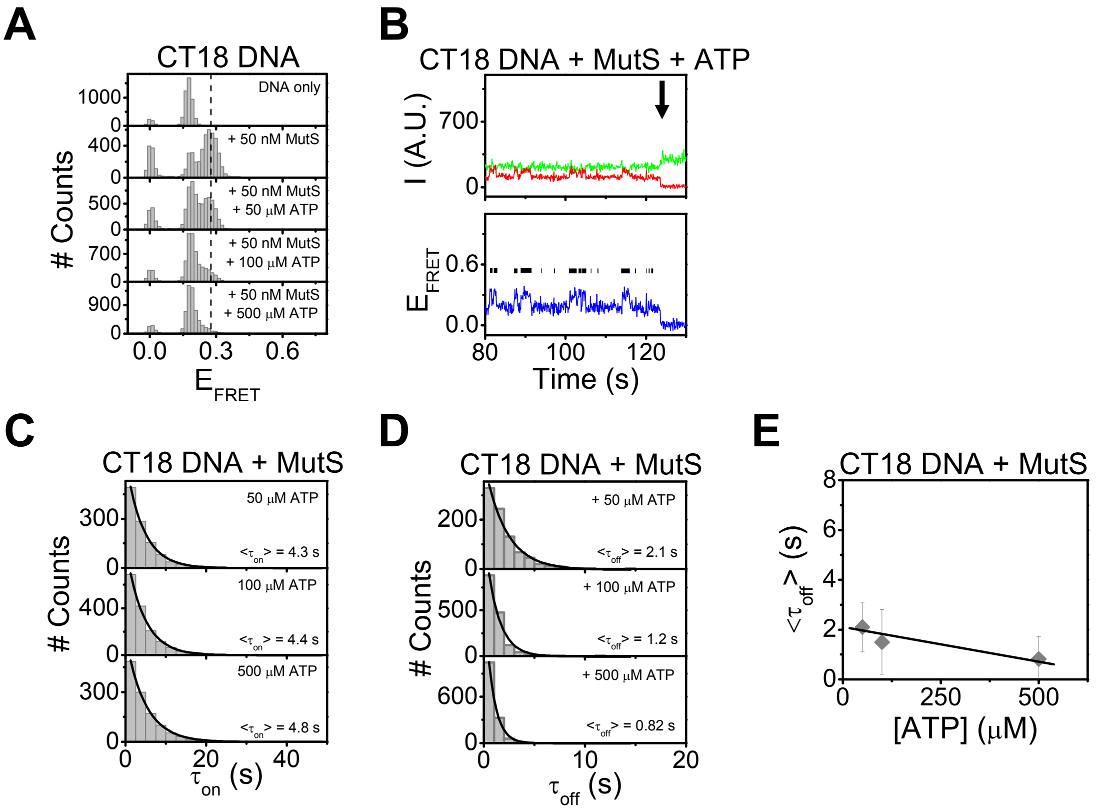


Figure 5.6
(A) FRET efficiency (EFRET) histograms for the C/T mismatched DNA (CT18) in the absence of MutS and in the presence of 50 nM MutS and 0-500 μM ATP are shown. (B) A representative single-molecule intensity time trace and the corresponding FRET efficiency time trace for CT18 DNA in the presence of 50 nM MuS and 100 μM ATP is shown. The black lines highlight MutS binding. The black arrow denotes the instant of acceptor photobleaching. (C) The distributions of times (τ_{on}) between MutS binding events (to CT18) at 50-500 μM ATP are shown. (D) The distributions of times of MutS-CT18 binding times (τ_{off}) for 50 nM MutS at 50-500 μM ATP are shown. (E) The average MutS binding times (τ_{off}) from the histograms shown in (D) are plotted as a function of ATP concentration.

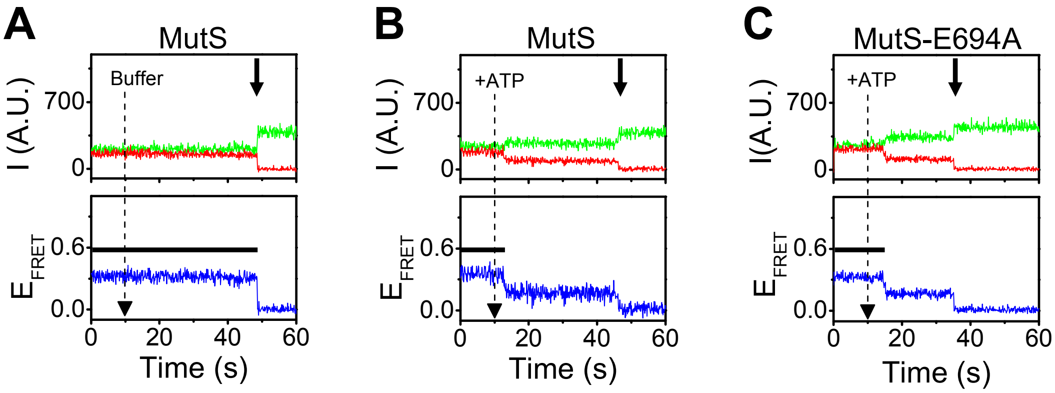


Figure 5.7
Flow experiments with GT18 DNA. (A) 50 nM MutS was pre-bound to GT18 DNA, and buffer was flowed into the sample chamber at time = 10 s. (B) 50 nM MutS was pre-bound to GT18 DNA, and buffer containing 100 µM ATP was flowed into the sample chamber at time = 10 s. (C) 50 nM MutS-E694A was pre-bound to GT18 DNA, and buffer containing 100 µM ATP was flowed into the sample chamber at time = 10 s.

Chapter 6

Conclusions

Accurate copying of a cell's DNA is essential for normal cell growth and viability, and the DNA mismatch repair system works to find, excise, and correct errors that occur during DNA replication. If DNA synthesis errors are not properly repaired, harmful mutations can accumulate which can lead to genetic instability and disease. In the studies presented in this dissertation, single-molecule FRET was used to investigate the molecular mechanisms of the first step of the DNA mismatch repair pathway, DNA mismatch recognition. Single-molecule FRET is a useful technique for measuring conformations and conformational changes in biomolecules for distances ranging from 20-80 Å. In this work, we end-labeled DNA molecules with a donor/acceptor FRET pair so that the relative end-to-end distances of the DNA molecules could be observed via single-molecule FRET in the absence and presence of DNA mismatch recognition proteins. Since DNA mismatch recognition proteins bend DNA substrates upon specific mismatch binding, this method allowed direct observation of the binding/dissociation kinetics of mismatch recognition proteins in a variety of buffer conditions. Below, the main conclusions, relevance of the findings, and suggested future directions for each study are discussed.

6.1 Chapter 3: Discrete conformational dynamics in three-way DNA junctions

The purpose of this study was to systematically characterize the conformational dynamics of three-way DNA junctions that can be formed when partially palindromic triplet repeat sequences (such as CAG and CTG repeats) slip out to form an extra-helical hairpin. Through the analysis of the conformations and conformational dynamics of various three-way DNA junctions, it was determined that DNA molecules with hairpins formed from CAG and CTG repeats have intrinsic conformational fluctuations between discrete states, and the conformational states of each conformer and the transition rates each conformer are dependent on the base-base mispairing on the hairpin arm, the presence of the hairpin loop, the sequence about the branch point, NaCl and $MgCl_2$ concentrations, and the temperature. The conclusion of this study is that three-way DNA junctions are DNA structures that can have highly heterogeneous conformational dynamics between discrete states, and the conformational dynamics of these structures originate from base pairing rearrangements at the branch point that are driven by DNA melting and re-annealing.

Three-way DNA junctions formed by self-pairing of triplet repeats are believed to be important intermediates in triplet repeat expansion, so understanding the dynamic properties of these non-canonical nucleic acids may have important implications for proteins that recognize the structures. In particular, previous studies of three-way DNA junctions using ensemble techniques have noted that these DNA structures are conformationally flexible—meaning that these types of junctions have a wide energy landscape. However, in this study, discrete, narrow conformational populations of were resolved with single-molecule FRET—demonstrating that the energy landscape of these molecules consists of narrow potential wells. For protein-DNA interactions, the distinction between a conformationally flexible energy landscape and a conformationally discrete energy landscape is a subtle, but important point. Proteins that bind to a conformationally flexible DNA junction would have nonspecific conformational interactions, since the protein-DNA interactions would occur at a large range of DNA conformations. However, proteins that bind a DNA molecule may have specific interactions with each particular conformer that may result in conformationally specific biochemical outcomes. Finally, understanding the dynamics of three-way DNA junctions is important for a general understanding of nucleic acid folding. For example, three-way junctions are common motifs in RNA, so the knowledge gained from studying three-way DNA junctions will be useful for understanding three-way RNA structures.

The study presented here is the first systematic single-molecule FRET study of the conformational dynamics of three-way DNA junctions, and it sets the stage for future work on these dynamic nucleic acid structures. Future studies with multiple vector labeling and three-color FRET spectroscopy will allow a more detailed analysis of the specific geometries of each conformer.

6.2 Chapter 4: Conformational trapping of Human Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops

The objective of this study was to determine how human MSH2/MSH3—a protein whose function is to find extra-helical loops and initiate repair of the lesion—can contribute to the CAG repeat expansion, which is the lethal mutation in Huntington's disease. Single-molecule FRET was used to observe binding and dissociation of

MSH2/MSH3 proteins to and from repair-competent $(CA)_4$ loops and repair-resistant $(CAG)_{13}$ DNA loops in real time. MSH2/MSH3 binds/bends all extra-helical loops in a concerted manner with low nanomolar affinity. After binding, ATP-bound MSH2/MSH3 dissociates quickly from the repair-competent loops, while the repair-resistant loops trap nucleotide-bound MSH2/MSH3, inhibiting its dissociation from the DNA. Since ATP-stimulated dissociation from mismatches is proposed to be an important mechanism for downstream signaling, these results are consistent with a model in which specific interactions between MSH2/MSH3 and the conformation of the extra-helical loop—not just the flexibility of the mismatched DNA itself—governs whether MSH2/MSH3 can bind/hydrolyze ATP in order to move away from the mismatch and signal repair or whether the enzymatic activity of MSH2/MSH3 becomes stalled and the protein stabilizes the extra-helical loop which becomes a precursor for mutation.

The model proposed in this study has important implications for the expansion of triplet repeat sequences. When bound to $(CAG)_{13}$ hairpins, MSH2/MSH3 fails to effectively couple DNA binding with ATP hydrolysis, so MSH2/MSH3 remains bound to the lesion, stabilizing the loop. This result implies that the role of MSH2/MSH3 in triplet repeat expansion (or type two mutations as described in the introduction) is to bind and stabilize loops, which ironically protects the extra-helical loops from repair.

In addition, this study has significant implications for protein-DNA interactions in general, as this study clearly demonstrates that the enzymatic activity of MSH2/MSH3 is sensitive to the conformation of the junction. Thus, the shape of the DNA binding site is a player in the activity of the enzyme.

The study presented here is the first single-molecule FRET study of human MSH2/MSH3, and this work sets the stage for future studies of MSH2/MSH3 in complex with other loops. It will be interesting to investigate the conformational dynamics of MSH2/MSH3 in complex with other well-repaired and poorly repaired loops to see how MSH2/MSH3 processes other extra-helical loops. In addition, it will be instructive to use other fluorescent labeling schemes to observe more details about the structural changes in three-way DNA junctions upon MSH2/MSH3 binding.

6.3 Chapter 5: The binding kinetics of E. coli MutS influence DNA mismatch selectivity

The purpose of this study was to compare the conformational dynamics of MutS in complex with the efficiently repaired G/T mismatch to MutS in complex with the less efficiently repaired C/T mismatch in order deduce what aspects of DNA mismatch recognition may contribute to the efficiency of the DNA mismatch repair pathway. In the absence of ATP, MutS binds/bends both G/T and C/T mismatched DNA substrates in a concerted manner, but the binding stability of MutS in complex with the G/T mismatch is more than 30× than the binding stability of MutS in complex with the C/T mismatch. In the presence of ATP, the dwell times of MutS in complex with both substrates is similar (the binding interaction lasts a few seconds), but the dwell time distributions of MutS in complex with the substrates is distinct. The dwell time distributions of MutS in complex with the G/T mismatch show that multiple kinetic steps are required for protein dissociation, while the dwell time distributions of MutS in complex with the C/T mismatch show that only a single kinetic step is required for protein dissociation. These results support a model in which stable protein-DNA interactions are required for ATP uptake and formation of a clamp that can then move away from the mismatch to signal repair. Unstable MutS-mismatch complexes have a high probability of dissociating before ATP binding occurs, so many non-productive binding interactions may occur before MutS binds ATP and forms a sliding clamp that is capable of initiating downstream repair. An interesting aspect of this study from the viewpoint of understanding fundamental characteristics of protein-DNA interactions is the observation that MutS binds/bends the efficiently repaired G/T mismatch and the less efficiently repaired C/T mismatch with similar angles, so conformational flexibility of the substrate does not confer efficient mismatch recognition intrinsically. Thus, mismatch selectivity is regulated not at the level of binding a mismatch but rather at the level of nucleotide binding. These results are relevant for understanding type one mutations (described in the introduction), since mutations in MutS that may decrease the stability of mismatch binding may prevent proper ATP uptake and may lead to spontaneous mutations and disease.

The study presented here is the first single-molecule FRET study of *E. coli* MutS, and this work sets the stage for future studies of MutS in complex with MutL in order to understand how MutS recruits MutL and how MutS-MutL interactions activate strand excision.

6.4 Final comments

Understanding the details of individual protein-DNA interactions is important for a full understanding of molecular pathways. Traditionally, studies of protein-DNA interactions have been accomplished using bulk biochemical methods. Such methods are useful for determining a protein's DNA site specificity, DNA binding affinity, oligomeric state, and (if applicable) ATPase activities. Details on the kinetics of protein-DNA interactions, however, are difficult to derive from bulk biochemical studies, since the observables of bulk experiments are an ensemble average of all the protein-DNA interactions. Structural studies of protein-DNA interactions can provide more detailed information on specific protein-DNA interactions, but these studies do not reveal the dynamic nature of such interactions. Single-molecule studies, however, provide information on individual protein-DNA interactions that is difficult or even impossible to obtain in traditional biochemical and structural studies. In the studies presented in this dissertation, single-molecule FRET was used to detect the conformational dynamics of fluorescently labeled non-canonical DNA molecules alone and in the presence of DNA mismatch recognition proteins, and these studies have shed light on many details of DNA mismatch recognition that were not previously known.

Appendix A

The following protocols outline the experimental details needed to successfully carry out single-molecule FRET experiments.

A.1 Construction of a prism-type TIR microscope

This section describes how to setup a prism-type TIR microscope.

(1) Excitation

Equipment:

Brass stage plate and prism holder

Laser (green, 532 nm [Nd:YAG], 50 mW; Coherent, LaserQuantum, Newport, CrystaLaser)

Lens (BK7 plano-convex, ½'' diameter, 50-mm focal length, Thorlabs)

Micrometers (3; 1'' travel translation stage, Thorlabs)

Mirrors (broadband dielectric, 1'' diameter, Thorlabs)

Objective lens (1.2 NA water immersion, 60×, UPlanApo, Olympus)

Prism (Pellin-Broca [fused silica, 11.0×20.0×6.4], EKSPLA)

Methods:

1. Put the prism in the brass prism holder and secure it on top of a quartz slide with a thin layer of immersion oil in between the prism and the slide to match the index of refraction.
2. Mount the lens on the micrometers, and put the construct on the microscope stage (the lens should be ~50 mm from the prism).

3. Use a system of mirrors in concert with the lens to focus the laser beam onto the side of the prism. Control the location and size of the excitation area by the position of the lens.

(2) Emission

Equipment:

Dichroic mirror (long-pass, 630-645 nm cutoff wavelength, Chroma Technology)

Filter (long-pass, 550 nm cutoff wavelength, Chroma Technology)

Lens (visible achromat doublet, 2’’ diameter, 100 and 150 mm focal length, Thorlabs)

Mirrors (2, broadband dielectric, 2’’ diameter, Thorlabs)

Mechanical slit (Thorlabs)

Methods:

1. Place a fluorescent bead sample on the microscope and send the fluorescence signal to a side port.
2. Place an archromat (L1, focal length = 100 mm) 100 mm away from the image plane to collimate the beam.
3. Place a mirror (M1) where the beam becomes the tightest.
4. Place the other acromat (L2, focal length = 150 mm) about 5’’ away from M1. Reflect the beam with the other mirror (M2), and install the camera 150 mm away from L2. Make sure that the camera is in the proper focal plane by comparing the images from an eyepiece and from the CCD screen. If the distance between L2 and the camera is far from 150 mm, the beam is not well collimated by L1. (Note, the emission beam should travel through the centers of all the lenses; otherwise,

geometric aberration deforms the image of single molecules. Therefore, it is necessary that the side port, the camera, and the lenses all be the same height.)

5. Prepare a vertical slit about 2.5 mm in width and place it in the image plane.
6. Adjust the angle of M2 so that the image is on the left side of the CCD screen.
7. To split the two colors, mount a dichroic mirror (DM) on a filter holder, and place the DM a few mm away from M1. While adjusting the position of the DM, watch the CCD screen and determine the proper position/angle of the DM. After immobilizing the mirror mount, adjust the relative position of the donor and acceptor images by changing the angle of either M1 or DM. The positions of both images can be simultaneously adjusted by M2.

A.2 Imaging buffer

The imaging buffer is composed of the components of the reaction buffer for the system under study, 0.1 mg/mL glucose oxidase, 0.4% (w/v) D-glucose, 0.02 mg/mL catalase, and saturated Trolox ®. To prepare an imaging buffer:

1. Prepare a 5X reaction buffer:
    a. 5X buffers are composed of a 5X reaction buffer for the system + 2% (w/v) D-glucose (Note: The 5X buffer is stored in the refrigerator and can be used for a year).
2. Prepare a 1X reaction buffer:
    a. Make a stock solution of saturated TROLOX ®:

        i. Get the TROLOX ® (Chemical name: 6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid, 97%) from the refrigerator and let it warm up to room temperature.
        ii. Put ~63 mg in a 50 mL tube and fill the tube to the 50 mL line with Millipore-filtered DI water (TROLOX ® is saturated at 2 mM, but it is hard to get it to dissolve in water, so we add enough for ~5 mM solution so that the concentration of TROLOX ® is as high as possible).
        iii. Let the TROLOX ® solution sit in the dark for at least one hour so that the powder will dissolve as much as possible.
        iv. Get out a sterile 30 mL syringe and a 0.20 µm filter. Pre-wash the syringe and filter by flowing Millipore-filtered DI water through it slowly (one drop at a time). Pour the TROLOX ® solution into the syringe and filter out the TROLOX ® solution into a new 50 mL tube.
        v. Store the filtered saturated TROLOX ® solution in the refrigerator for up to four weeks.
    b. Mix 2ml 5X buffer + 8ml saturated TROLOX ® (This buffer should be prepared with fresh TROLOX ® every 4 weeks).
3. Before imaging single molecules:
    a. Make a gloxy solution. Gloxy is a 100x stock solution containing 1 mg/ml glucose oxidase and 0.2 mg/ml catalase in T50 buffer (T50 buffer contains 10 mM Tris-HCl pH 8.0 and 50 mM NaCl, and the solution is filtered

prior to use with a 0.2 μm filter). Gloxy is added to a 1x imaging buffer right before imaging single molecules because the gloxy + glucose reacts together in the oxygen scavenging reaction and generates glutonic acid which will cause the solution to become very acidic over a long time period. By adding gloxy right before imaging single molecules, the oxygen is removed from solution, and the buffer's acidity is mediated. To make a gloxy solution:

i. Weigh out 10 mg of glucose oxidase (glucose oxidase is stored at -20°C, so let it warm up to room temperature before opening it), and put it in a 1.5 mL tube.
ii. Add 48.6 μL catalase (stored in the refrigerator; invert the bottle gently several times to mix it before using it) to the tube.
iii. Add 71.4 μL T50 buffer to the tube.
iv. Mix the solution by tapping (no aggressive vortexing—this will deactivate the enzymes).
v. Centrifuge the solution for one minute.
vi. Store the solution in the refrigerator. The gloxy supernatant can be used for several months.

4. Add gloxy to the 1X buffer (at the time of the experiment)
    a. Add 1% gloxy to the 1X buffer. Let the mixture sit for 5 minutes to get rid of the oxygen. Add the solution to the slide. Image single-molecules.

A.3 Surface preparation

This section describes the slide cleaning and assembly protocols for the preparation of sample chambers with either BSA- or PEG-coated surfaces to which single molecules can be tethered for use in smFRET experiments.

Cleaning, assembling, and preparing slides for immobilization via BSA-biotin coated surfaces

Quartz slides are coated with BSA-biotin by (1) cleaning the slides and coverslips, (2) assembling the sample chamber, and (3) adding BSA-biotin to the sample chamber.

(1) Cleaning Slides

1. Let the used slides sit in H_20 for a few hours so the epoxy will begin to dissolve.
2. Put the slides in the microwave for 3-4 minutes to help the epoxy get soft.
3. Use a razor to take off the cover slips (put the coverslips in the glass waste box).
4. Put slides in the slide glassware.
5. Rinse slides with Millipore-filtered DI water.
6. Put the slides in a solution of 20% detergent and 80% Millipore-filtered DI water; sonicate for 15 minutes.
7. Rinse slides with Millipore-filtered DI water.
8. Fill the glassware up with Millipore-filtered DI water and sonicate for another 5 minutes.
9. Rinse with acetone. Dump acetone into the acetone waste container. Fill glassware with acetone and sonicate for 10 minutes. Dump acetone into the acetone waste container.
10. Rinse with Millipore-filtered DI water. Fill glassware with Millipore-filtered DI water and sonicate for 5 minutes.

11. Rinse with Millipore-filtered DI water. Fill glassware with KOH. Get a separate piece of glassware and put coverslips (make sure they are labeled 24x40 with 1.5 thickness) in the container and fill with KOH. Sonicate both for 20 minutes. The KOH waste can be dumped into the sink and flushed with tap water.
12. Rinse both slides and coverslips well with Millipore-filtered DI water.

(2) Assembling Sample Chamber

Do these steps immediately after cleaning slides.

13. Dry slides and coverslips with N_2. Hold the slides/coverslips with the tweezers and only blow them sideways so they will not break. Use less air when drying the coverslips—they are more fragile!
14. Put double-sided tape on the slide and put the coverslip on top of the slide and tape. Use a pipette tip to press down and seal the coverslip to the tape. Then, cut the excess tape with a razor blade.
15. Store slides in a covered box in a drawer for a few weeks.

(3) Creating a BSA-coated surface

Do these steps immediately before a single-molecule FRET experiment.

16. Get a clean, assembled BSA side and wash it with 100 μL T50 buffer.
17. Epoxy the edges of the slide and wait ~10 minutes for the epoxy to dry.
18. Wash the slide again with 100 μL T50 buffer to make sure the slide is sealed properly.
19. Add 60 μL of 1 mg/mL biotinylated BSA in T50 buffer. Incubate the slide for 10 minutes.
20. Wash the slide with 100 μL T50 buffer.

21. The slide is now ready for immobilization of DNA by biotin-streptavidin bonds (see **Appendix A.4**).

Cleaning, assembling, and preparing slides for immobilization via PEG coated surfaces

Quartz slides are coated with PEG by (1) cleaning the slides and coverslips using the same method given for BSA slides, (2) coating the slides with aminosilane, (3) conjugation with the NHS-ester form of PEG, and (4) assembling the sample chamber.

(1) Cleaning Slides

Set the temperature of the sonicator to 50° (for later). Pour methanol into the labeled glassware (labeled slide container, labeled coverslip container, and labeled small Erlenmeyer flask).

1. Let the used slides sit in H_20 for a few hours so the epoxy will begin to dissolve.
2. Put the slides in the microwave for 3-4 minutes to help the epoxy get soft.
3. Use a razor to take off the cover slips (put the coverslips in the glass waste box).
4. Put slides in the slide glassware.
5. Rinse slides with Millipore filtered DI water.
6. Put the slides in a solution of 20% detergent and 80% Millipore filtered DI water; sonicate for 15 minutes.
7. Rinse slides with Millipore filtered DI water.
8. Fill the glassware up with Millipore filtered DI water and sonicate for another 5 minutes.

9. Rinse with acetone. Dump acetone into the acetone waste container. Fill glassware with acetone and sonicate for 10 minutes. Dump acetone into the acetone waste container.
10. Rinse with Millipore filtered DI water. Fill glassware with Millipore filtered DI water and sonicate for 5 minutes.
11. Rinse with Millipore filtered DI water. Fill glassware with KOH. Get a separate piece of glassware and put coverslips (make sure they are labeled 24x40 with 1.5 thickness) in the container and fill with KOH. Sonicate both for 20 minutes. Dump the KOH into the KOH waste container.
12. Rinse both slides and coverslips well with Millipore filtered DI water.
13. Get the bottle of aminosilane (3-Aminopropyltrimethoxysilaine) out of the freezer so it can warm up to room temperature.
14. Rinse both the slides and the coverslips very well with pure water. Fill the slide and coverslip glassware up with pure water and cover with the lid. This is a good place to stop if one does not intend to carry out the entire procedure on a given day. If equipped to proceed, continue on.

(2) Coating the slides with aminosilane

Organosilanes are compounds containing silicon to carbon bond. Organosilanes can be used to modify a surface. For instance, aminosilane—a specific type of organosilane—can be attached to the surface of a negatively charged quartz slide, and this reaction will result in a ~2nm coating of positively charged, reactive amine groups that can be used for downstream reactions.

15. Pour the methanol out of the labeled slide and coverslip glassware (leave the methanol in the flask for now). Shake out any remaining methanol from the bottom of the containers. Always collect methanol waste in the appropriately labeled waste container—do not pour it down the drain.
16. Take slides/coverslips over to the drying table. Also, take the set of labeled glassware and a set of unlabeled, dry glassware to the drying table.
17. Dry coverslips and slides:
    a. Start with coverslips:
        i. Rinse each coverslip well with pure water.
        ii. Dry each coverslip with N_2.
        iii. Put each dried coverslip in unlabeled glassware, and put the lid on the container.
    b. Proceed to slides:
        i. Rinse each slide well with pure water.
        ii. Dry each slide with N_2.
        iii. Put each dried slide in unlabeled glassware, and put the lid on the container.
18. Burn coverslips and slides (to remove any residual fluorescent impurities):
    a. Start with coverslips:
        i. Swipe each coverslip across the flame four times (2 swipes for each side of the coverslip). If you leave the glass coverslip over the flame too long, it will break.

        ii. Place the coverslips in the labeled glassware and put a lid on the glassware.

    b. Proceed to slides:

        i. Burn each side of every slide with the flame. Make sure you do this very well because this surface must be free of any impurities.

        ii. Place the slides in the labeled glassware and put a lid on the glassware.

19. Dump the methanol out of the labeled flask.

    a. Put 100 mL of methanol into the flask.

    b. Put 5 mL of Millipore filtered DI water into the flask.

    c. Put 1 mL of acetic acid (from the flammable cabinet) into the flask. Make sure you use a glass tip over the plastic tip when you are getting the acetic acid.

       This mixture of acid, water, and methanol is your solvent.

    d. Put 2 mL of amino-silaine in the flask (use a glass pipet tip for this). This step must be done quickly so that it doesn’t oxidize.

    e. Seal the aminosilane bottle.

20. Make sure that the mixture in the labeled flask is mixed well by swirling the container.

21. Pour the solution into the labeled glassware containing the slides and coverslips. Put the lids on the glassware. Immediately, put the labeled flask in the sink and rinse it well with Millipore filtered DI water.

22. Put the glassware in the hot bath for 10 minutes.

   a. Put the jar of aminosilane in the vacuum chamber.
      i. Leave the aminosilane bottle in the vacuum chamber for 10 minutes. Make sure the lid on the bottle is loose.
      ii. Put N_2 knozzle in the vacuum and fill the chamber with N_2.
      iii. Open the chamber quickly and tighten the lid on the aminosilane.
      iv. Seal the lid with parafilm.

23. Sonicate the glassware for 1 minute.
24. Leave the glassware in the hot bath for another 10 minutes.
25. Remove the flasks from the sonicator. At this point the aminosilane solution should be attached to the surfaces.
26. For both the coverslips and the slides:
    a. Dump out the solution.
    b. Rinse two times with methanol.
    c. Rinse many times with Millipore filtered DI water.
    d. Fill the flasks with Millipore filtered DI water.

(3) PEG conjugation

27. Make a sodium bicarbonate buffer ($NaHCO_3$)
    a. In the shared lab, measure of ~84 mg of the $NaHCO_3$ power with a clean spatula.
    b. Put the power in a 15 mL plastic container.
    c. Fill the container up to the 10 mL line with Millipore filtered DI water.
    d. Label the container.

   e. *Note* You must used the sodium bicarbonate buffer the same day you make it.

28. From the environment-controlled container in the freezer, get the PEG with biotin and without biotin. Let the vials warm up to room temperature.
    a. Weigh out 80 mg of the PEG in the shared lab and put it in a 1.5 ml tube.
    b. Weigh out ~3 mg of the PEG with biotin and put it in the 1.5 ml tube.
    c. Put 320 μL of $NaHCO_3$ buffer into the test tube. Flick it. Centrifuge it for one minute. The solution is highly viscous, so we must spin it to eliminate bubbles. After one minute of spinning, you should have a clear solution.
    d. Put the PEG vials in the vacuum for 30 minutes. Break the vacuum with N_2 and quickly close the lids and seal the vials with parafilm.
29. Dry the slides (Do the remaining steps relatively quickly; the quicker the slides and coverslips are prepared, the cleaner they will be.):

    Make sure the symmetry of the human body is exploited to keep track of where the PEG surface is (be very careful to keep track of which side of the slides/coverslips the PEG is on).
    a. Start with the coverslips:
       i. Rinse coverslips well with Millipore filtered DI water from a squirt bottle.
       ii. Dry the coverslips with N_2.
       iii. Put the coverslips in dry coverslip glassware and cover with a lid.
    b. Proceed to the slides:

   i. Rinse the slides well with Millipore filtered DI water from a spray bottle.
   ii. Dry the slides with N_2.
   iii. Lay the slides in a pipette container and close the lid.

30. Put 70 µL of the PEG solution on top of each slide and put a coverslip over each slide. By this method, both internal surfaces are getting coated by the PEG.
31. Put water in the boxes to keep the slides from drying out and put the lids on the boxes. Put the boxes in a dark place (a drawer) so the polymer will not get hydrolyzed by the light.
32. Throw out the $NaHCO_3$ buffer. It can only be used the day it is made.
33. Wait 3 hours.

(4) Assembling the sample chamber

34. Rinse slides and coverslips well with nanopure water. Dry slides and coverslips with filtered air (or N_2). Hold the slides/coverslips with the tweezers and only blow them sideways so they will not break. Use less air when drying the coverslips—they are more fragile!
35. Put double-sided tape on the slide and put the coverslip on top of the slide and tape, as shown in the figure below. Use a pipette tip to press down and seal the coverslip to the tape. Then, cut the excess tape with a razor blade.
36. Store slides in a dark desicator at room temperature for up to 2 weeks.

Do the next steps immediately before a single-molecule FRET experiment.

37. Get a clean, assembled PEG slide and wash it with 100 µL T50 buffer.
38. Epoxy the edges of the slide and wait for the epoxy to dry.

39. Wash the slide with 100 μL T50 buffer to make sure the slide is sealed properly.

A.4 Surface immobilization

1. Get an epoxyed BSA-biotin or PEG coated slide.
2. Add imaging buffer and look at the slide in multiple areas with the microscope to see if there are any fluorescence impurities on the slide surface.
3. (Optional) To check for non-specific binding to the slide surface (and thus ensure that the BSA-biotin/PEG is fully covering the slide surface), add 0.1-1 nM Cy3-labeled DNA to the slide, wait 5 minutes, wash the slide with 100 μL T50 buffer, add imaging buffer, and look at the slide under the microscope. No fluorescent spots should be visible. Wash out the slide with 100 μL T50 buffer.
4. Add 0.25 mg/mL streptavidin. Wait 10 minutes. Wash out the slide with 100 μL T50 buffer.
5. Add 20-50 pM DNA (labeled with donor, acceptor, and biotin) to the slide in T50 buffer containing 1 mg/mL BSA. Wait 10 minutes. Wash out the slide with 100 μL T50 buffer.
6. Add imaging buffer, and the slide is ready for single-molecule FRET experiments.

A.5 Data acquisition

Total internal reflection (TIR) microscopy is a wide-field microscopy technique which allows the simultaneous observation of ~200-300 individual fluorophores. In this technique, total internal reflection of the incident laser beam occurs at the surface of a slide, and the fluorophores near the surface of the slide are excited by the resultant evanescent field. Because the excitation light is totally internally reflected, this technique has low background, an essential component of single molecule imaging.

Typically, we use prism-type TIR because this set-up allows us to achieve a very high signal-to-noise ratio and is necessary to perform flow experiments reliably. The limitations to prism-type TIR are that it can be difficult to find the excitation area, the top side of the slide is covered with a prism (which is inconvenient and can cause the imaging area to go out of focus), and it requires use of relatively brittle and expensive quartz slides to reduce scattered light at the slide surface.

As with all microscope set-ups, extreme care must be taken when using the TIR microscope. Do not lean on the optical table or move/ bump any optical components, as these events could spoil the alignment of the system.

Instructions for using the TIR microscope for smFRET

1. Turn on the laser (first, turn on the power button, wait ~45 seconds; then, turn the key).
2. Turn on the shutter.
3. Plug in the CCD camera.
4. Log into the computer.
5. Open the TIR program (double-click the ‘Single’ icon on the deskop—from the menu bar select Module—from the drop-down menu pick TIR).
6. Wait for the camera temperature to reach -90°C (check the temperature of the camera by clicking ‘Get Temperature’). NEVER open the camera before it reaches -90°C.
7. Turn off the default autoscale function.

8. Calibrate the microscope with a fluorescent bead sample (see **Appendix A6** for instructions for making the fluorescent bead sample). For this procedure, the Single program settings should read:

   Exposure time: 0.032

   Data scaler : 600

   Background sub: 390

   Gain: 1

   To calibrate the machine:

   a. Put a drop of water on the objective lens. Carefully, put the slide on the stage, epoxy side down; secure the slide with the clips. DO NOT press down on the clips. Put a drop of oil on top of the slide. Make sure that there is no oil on the prism (wipe it with a kimwipe), and position it on top of the slide. Secure the prism in place (push it back and to the left) with the screw, but DO NOT screw it tightly.
   b. Find the excitation area on the sample surface. To do this, click the shutter button on the Single program to open the shutter in order to allow the excitation light through to the bead sample. Turn the knob on the front of the microscope up, to allow the image to the eyepiece, and remove the filter by rotating the filter holder from the ‘1’ position to the ‘2’ position. In general, locating the excitation beam requires (a) correct placement of the objective lens and (b) proper positioning of the focusing lens.
      i. The objective lens can be moved up and down by the knobs

located on the base of the microscope. The larger of the two knobs is the coarse adjustment and the smaller one is the fine adjustment. Clockwise movements of these knobs lower the objective, and counterclockwise movements of these knobs raise the objective. Always exercise caution in moving the objective counterclockwise (up), as careless raising of the objective can damage it. NOTE: IF THE MICROSCOPE HAS BEEN USED RECENTLY FOR smFRET, THE COURSE FOCUS PROBABLY DOES NOT NEED TO BE ADJUSTED. When you are using this microscope (or any microscope system) you must exercise caution when focusing the system so that you do not crash the objective into the slide. Objective lenses are very expensive and should be treated with great care.

    ii. The focusing lens can be positioned by turning the knobs located on the top of the stage. When adjusting the focusing lens, always write down the starting numbers for each knob. Adjust the intensity of excitation at the sample by changing the size of the spot and/or rotating the filter in front of the excitation laser.

c. Once the excitation spot is in the right position, take a short (~20 second) movie of the fluorescent beads.

d. Go to E:\user\tir, and change the names of the two new files (they will be called hel1.pma and hel1.txt) to rough.pma and rough.txt.

e. Open IDL. From the menu bar select macros, and from the drop-down menu select initialize. Next, select macros from the menu bar, and select color initialize from the drop-down menu.

In the command line type:

```
.run maketiff > enter
rough > enter
```

A .tiff file ‘rough.tiff’ was created from the .pma file ‘rough.pma’.

In the command line type:

```
.run calc_mapping2 > enter
rough > enter
```

Use the mouse to point to a spot in the upper left-hand corner of the donor image, and use the keys DRFC to center the circle around the spot. Next, use the keys GYHB to center the circle on the acceptor image around the corresponding spot. Then, hit the ‘S’ key to save the circle. Repeat this process two more times: Once for a spot located in the right center of the images and once for a spot in the lower left corner of the images.

Then, in the command line type:

```
.run nxgn1_cm > enter
rough > enter
```

If the donor/acceptor channels are properly calibrated, white circles will appear around the dots.

In the IDL toolbar, click the open folder, then open the file P_nxgn1_ffp.pro. Scroll down to line 191 of the program and change the code to reflect the current date: E:\user\tir\Month##_##\rough.map. Save the changes. Select 'Run' from the menu bar and choose compile P_nxgn1_ffp.pro.

Go to E:\user\tir, make a new folder 'Month##_##', and move all the rough files to the new folder. The machine is now ready to do single-molecule FRET experiments.

Carefully, remove the prism from the microscope stage, wiping off the oil with a kimwipe. Next, unclip and remove the slide from the stage. Using a clean kimwipe, first remove the water from the objective facing side of the slide, and then wipe the oil from the prism facing side of the slide. Do not get oil on the side of the slide that touches the water objective lens.

9. Take data with a DNA sample. For DNA samples, the Single program settings should be changed to:

Exposure time: 0.032
Data scaler: 6000
Background sub: 390
Gain: 240

To view a DNA sample:

a. Put a drop of water on the objective lens. Carefully, put the slide on the stage, epoxy side down; secure the slide with the clips. Put a drop of oil on top of the slide. Make sure that there is no residual oil on the prism (wipe it with a kimwipe), and position it on top of the slide. Secure the prism in place (push it back and to the left) with the screw, but DO NOT screw it tightly.

b. Find the excitation area on the sample surface. To do this, click the shutter button on the Single program to open the shutter in order to allow the excitation light through to the bead sample. Turn the knob on the front of the microscope up, to allow the image to the eyepiece, and remove the filter by rotating the filter holder from the '1' position to the '2' position. In general, locating the excitation beam requires (a) correct placement of the objective lens and (b) proper positioning of the focusing lens.
    i. The objective lens can be moved up and down by the knobs located on the base of the microscope. The larger of the two knobs is the coarse adjustment and the smaller one is the fine adjustment. Clockwise movements of these knobs lower the objective and counterclockwise movements of these knobs raise the objective. Always exercise caution in moving the objective counterclockwise (up), as careless raising of the objective can damage it. NOTE: IF THE MICROSCOPE HAS BEEN USED RECENTLY FOR smFRET, THE COARSE FOCUS PROBABLY DOES NOT NEED TO BE ADJUSTED. When you are using this microscope (or any microscope system) you must exercise caution when focusing the system so that you do not crash the objective into the

slide. Objective lenses are very expensive and should be treated with great care.

   ii. The focusing lens can be positioned by turning the knobs located on the top of the stage. When adjusting the focusing lens, always write down the starting numbers for each knob. Adjust the intensity of excitation at the sample by changing the size of the spot and/or rotating the filter in front of the excitation laser.

c. Take fifteen short (~20 frame) movies and three long (~2000 frame) movies. As a rule, it is necessary to take data until most of the acceptor molecules have photobleached (even if that is longer than 2000 frames) because we need for the acceptor fluorophores to photobleach so that we can determine the background level. If it takes more than ~3000 frames to achieve acceptor photobleaching, turn up the excitation intensity. In contrast, if the fluorophores are bleaching too fast (for example, the majority of the fluorophores should not bleach in ~1000 frames or ~30 seconds), the imaging conditions need to be examined.

d. Create a new folder inside the folder labeled Month##_## that is descriptive of the sample and experimental conditions and move the files (they are written to E:\user\tir) to the new folder. It is VERY important that your folder names are very descriptive so that anyone who may at your data in the future will know precisely what materials and conditions you used in your experiment. Move (cut/paste) the data you just took to this folder.

e. To analyze the data:

Go to the opened IDL window and type (in the command line):

```
.run ana_all > enter
Month##_### > enter
```

And wait for the program to run.

Open MATLAB, and in the command line type:

```
tjbatchfret('__enter path to folder here__') > enter
```

Specify low/high cutoffs > enter

Open Origin. In the upper left-hand corner of a new worksheet, there is a white rectangle. Right click the white rectangle and select ‘Import ASCII’. Select the FRETresult.dat file from the appropriate folder and click ‘Open’. Right click the A(X) column and select Set As > Y. Right click the A(Y) column and select Plot > Statistical graphs > Histogram. Double-click on the histogram, click the Data tab, and enter the following values: Bin Size: 0.01, Begin: -0.1, End: 1. Click Ok, and click the rescale button. Label the histogram. Right click on the A(Y) column and select Set column values by typing in ‘(Col(A) – 0.1±0.0X/0.9-/+0.0X)’, adjusting the amount of correction by changing X.

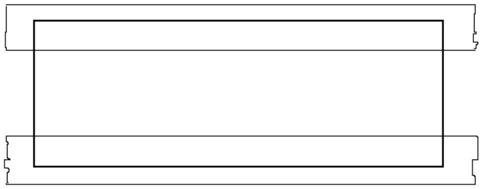
Repeat step 9 with the DNA sample for all conditions necessary.

A.6 Fluorescent Bead Alignment/Calibration Slide

The fluorescent bead sample is used for aligning and calibrating the TIR microscope for smFRET experiments, and it is made by diluting fluorescent beads which will attach to the imaging surface made of untreated glass. It is not necessary to use quartz slides for the bead sample because the light scattering from the glass is overcome by the intensity of the bead emission. Once a bead sample is constructed, it can be used for years.

To Prepare a Bead Sample:

1. Get a glass slide.
2. Get a glass cover slip.
3. Put tape on the edges of the slide, as shown in the figure below:



4. From the refrigerator, get the fluorescent beads (FluoSpheres carboxylate-modified microspheres, 0.2 μm, crimson fluorescent [625/645], Invitrogen).
5. Get 10 μL of the beads in the tip of a pipet.
6. Put the small volume on the surface in the center of the slide.
7. Spread them around with a pipet tip so there will be a uniform density.
8. Wait for the buffer to evaporate.
9. In the same container from the refrigerator, find the mounting medium.
10. Put a couple of drops of the mounting medium on the center of the slide. It is very viscous. Use a different tip for each drop so there is no cross contamination.
11. Put the coverslip on the slide. If there is air in the slide, the slide will not work right because the refractive index will be off.
12. Use a razor to cut off the excess tape.

13. Look at the slide under the microscope. If it works, seal it with epoxy and it will work for years.

A.7 Commonly used materials for smFRET experiments

T50 Buffer

T50 buffer is a buffer solution commonly used for washing slide chambers in single-molecule FRET experiments. T50 buffer is composed of 10 mM Tris-HCl pH 8.0 and 50 mM NaCl, and the solution is filtered prior to use with a 0.2 μm filter.

Streptavidin

DNA molecules are immobilized to slide chamber surfaces using the strong binding of biotin and streptavidin. Streptavidin is bought commercially (Sigma, product #) in dehydrated form. Upon receiving streptavidin, it should be re-hydrated with T50 buffer to 5 mg/mL and stored at 4 °C.

BSA

BSA (Albumin bovine serum, Sigma, product #) is a protein that is commonly added to buffers to minimize DNA or proteins from sticking to surfaces. It is stored in dehydrated from at 4 °C. To make 10 mg/mL stock solution:

1. Measure out 10 mg of BSA and put it in a 1.5 mL tube.
2. Add 1 mL T50 to the 1.5 mL tube.
3. Flick to mix, and store solution in the refrigerator.

BSA-biotin

BSA-biotin (Albumin, biotin labeled, bovine, Sigma, product #) is used to coat the surface of quartz slides for the immobilization of nucleic acids via biotin-streptavidin bonds. It is received in dehydrated form. After receiving BSA-biotin, stocks are made in the following way:

1. In the bottle that the BSA-biotin powder comes in: Add 1 mL of T50 to a 10 mg bottle of BSA-biotin upon opening a new bottle of BSA-biotin, and label the bottle '10 mg/mL'.
2. Make a personal aliquot of 1 mg/mL BSA-biotin (50 µL 10 mg/mL BSA-biotin stock + 450 µL T50).

ATP stock

1. To make 5 mL of 100 mM ATP stock: Put 0.27555 g ATP (FW = 551.1) in a 15 mL tube and add ~2.5 mL T50 to the tube. Check the pH with the pH meter and adjust the pH of the solution to ~7. Once the pH is at 7, fill the tube to the 5 mL line with T50.
2. Put the ATP in 20 µL aliquots and put the aliquots in a box in the freezer. Use one aliquot/experiment and then throw the tube away (do not put a used aliquot back in the freezer).
3. Every couple of weeks, put a few drops of ATP on a pH strip to make sure that the pH is stable.

ADP stock

1. To make 5mL of 100 mM ADP stock: Put 0.2136 g of ADP (FW = 427.2) in a 15ml tube and add ~2.5 mL T50 to the tube. Check the pH with the pH meter and adjust the pH of the solution to ~7. Once the pH is at 7, fill the tube to the 5 mL line with T50.

2. Put the ADP in 20 μL aliquots and put the aliquots in a box in the freezer. Use one aliquot/experiment and then throw the tube away (do not put a used aliquot back in the freezer).
3. Every couple of weeks, put a few drops of ADP on a pH strip to make sure that the pH is stable.

A.8 DNA protocols

Oligonucleotides are purchased from Integrated DNA Technologies (IDT) and are received in dehydrated form. Upon receiving a new oligo, the ssDNA must be rehydrated (to 100 μM), numbered and added to the 'oligonucleotides' folder and database computer file. Creating dsDNA substrates requires annealing two (or more) oligos together. The procedure for receiving and annealing oligos:

1. Upon receiving a new oligo:
    a. Spin the oligo in the microcentrifuge to make sure all the powder is at the bottom.
    b. Add Millipore filtered DI water to the tube so that the concentration of DNA is 100 μM (the amount of water that needs to be added will depend on the amount of DNA that is in the tube; for example, 5.2 nmol DNA requires 52 μL water).
    c. Spin the hydrated oligo once more.
    d. Number the oligo and store it in the oligo box in the freezer.
2. To anneal two oligos (to create linear dsDNA):

a. Determine which oligos you want to anneal and triple-check that they are complementary. To estimate the stability of the duplex, MFOLD can be used.

b. Turn the heat block to 80°C.

c. Combine two strands of DNA in a 1:1.2 ratio (2 µL strand 1 which is biotin-labeled, 2.4 µL strand 2, 4 µL annealing buffer, 11.6 µL sterilized nanopure water). The annealing buffer contains 100 mM Tris (pH 8.0) and 2500 mM NaCl. Flick the tube. Spin the tube in the microcentrifuge. Cover the tube with foil.

d. Put the tube on the heating block. After three minutes, turn the heating block off. Leave the heating block/tube alone for 3 hours to cool down to room temperature.

e. After 3 hours, spin the DNA on the microcentrifuge and put it on ice.

f. The concentration of the DNA is 10 µM. Dilute the DNA to 100 nM (1 µL of 10 µM stock + 1 µL of 10 mg/mL BSA + 9 8µL T50). Then, dilute the DNA to 1 nM (1 µL of 100 nM stock + 1 µL of 10 mg/ml BSA + 98 µL T50). Make many 1nM aliquots from the 1nM dilution (you will use the 1nM dilutions and then throw them away). Store all aliquots in your DNA box in the freezer.

3. To anneal three oligos (to create a three-way junction):

   a. Determine which oligos you want to anneal and triple-check that they are complementary. To estimate the stability of each arm, MFOLD can be used.

b. Turn the heat block to 80°C.

c. Combine two of the strands of DNA in a 1:1.6 ratio (2 μL strand 1 which is biotin-labeled, 3.2 μL strand 2, 2 μL annealing buffer, 2.8 μL sterilized nanopure water). The annealing buffer contains 100mM Tris (pH 8.0) and 2500 mM NaCl. Flick the tube. Spin the tube in the microcentrifuge. Cover the tube with foil.

d. Put the tube on the heating block. After three minutes, turn the heating block off. Leave the heating block/tube alone for 3 hours to cool down to room temperature.

e. In the meantime, get another tube and add strand 3 to it (3 μL strand 3, 2 μL annealing buffer, 5 μL sterilized nanopure water). Flick the tube. Spin the tube in the microcentrifuge. Cover the tube with foil and put it on ice.

f. After 3 hours, spin tube 1 on the microcentrifuge and add the contents of tube 2 to it. Cover the tube with foil and leave it at room temperature for 3 hours.

g. After the second 3 hour incubation, spin the DNA on the microcentrifuge and put it on ice.

h. The concentration of the DNA is 10 μM. Dilute the DNA to 100 nM (1 μL of 10 μM stock + 1 μL of 10 mg/mL BSA + 98 μL T50). Then, dilute the DNA to 1nM (1 μL of 100 nM stock + 1 μL of 10 mg/mL BSA + 98 μL T50). Make many 1nM aliquots from the 1nM dilution (you will use the 1nM dilutions and then throw them away). Store all aliquots in your DNA box in the freezer.

Appendix B

B.1 Protocol for expression and purification of *E. coli* MutS

The following protocol outlines the experimental details needed to successfully express and purify E. coli MutS. This protocol has been adapted from previously published protocols [42, 123].

Things you need to make sure you have before you start:

LB + ampicillin plates

eMutS/pET28 plasmid (or mutant plasmid)

0.1g/10mL ampicillin stock (filtered)[10]

LB medium (concentration is 25 g of LB broth powder in 1 L of purified water)

Gloves (wear gloves during all steps)

(1) MutS expression

Day 1 (mid-day/evening)

1. Do quick transformation:

   Get out an aliquot (10 μL) of HMS174 (Novagen) competent cells (aliquots are stored in the -80°C freezer), and put it on ice. Get out the plasmid eMutS/pET28 plasmic (-20°C freezer, in 'clone' box), and melt it. Add 0.5 μL eMutS/pET28 to the 10 μL aliquot of cells, and leave it on ice for 30 minutes (do not disrupt the cells). Also, get a LB + ampicillin plate from the refrigerator so that it can warm up to room temperature. After 30 minutes, induce heat shock by putting the sample in a water bath (42°C) for 30 seconds. Then, put the sample on ice again for 2 minutes. Finally, put the cells on the plate and spread them out with a loop and put them in an incubator (37°C) upside-down (LB agar side up) overnight.

---

[10] To make ampicillin stock, weigh out 0.5 g of ampicillin, put it in a 50 mL tube, and fill it to the 50 mL line with purified water. Filter it and transfer the solution to four 15 mL tubes labeled: 0.1/10mL ampicillin.

Day 2 (morning)

2. Small-scale inoculation:

   Pick a single colony off the plate and put it in a 15 mL tube that has 5 mL LB medium + 50 μL ampicillin (0.1g/10mL) and let it grow (37°C, 225 rpm) for 5-7 hours.

Day 2 (afternoon)

3. Large-scale inoculation:

   Dump the 5 mL from step 2 into one or more flasks containing 500 mL LB medium + 5 mL ampicillin (0.1 g/10mL) and let it grow (37°C, 225 rpm) until it reaches OD 0.5 (~3-3.5 hours for MutS/pET28). Add 500 μL of IPTG (100 mM stock), and incubate (37°C, 250 rpm) the culture for another three hours. After incubation, chill culture on ice for 5 minutes and then centrifuge the cells (5000 rpm rotor #10) for 10 minutes. The cell pellets will be at the bottom and you can dump the supernatant. When you finish spinning the cells, resuspend them and spin them in a 50 mL tube (this will make subsequent steps easier and will result in less cell loss). Dispose of the supernatant (you can dispose of the waste by adding Clorox to it or autoclaving the flasks) and store the pellets in the -20°C freezer until you are ready to do protein purification.

(2) MutS purification

Day 3 (morning)

Because of the plasmid we use (pET28), eMutS has a histidine tag on the N terminal, we can use an ion affinity column for protein purification. We use a cobalt affinity resin (TALON) (stored at 4°C).

Keep the cells/protein cold (on ice) during each step of protein purification.

1. Get cell pellets out of the -20°C freezer and put them on ice to thaw.
2. Re-suspend the cells (add ~25-30 mL chilled lysis buffer/2L culture to the cells and use a pipette to re-suspend the cells).

    Lysis buffer (from Schofield et al. [42])

    20 mM HEPES pH 7.8 -- 40 mL of 0.5M HEPES, pH 7.8 stock (4°C)

    500 mM NaCl -- 29.2 g NaCl

    10% glycerol -- 100 mL of 100% glycerol

    1 mM βME -- 70 µL βME (stock βME is ~14M)

    *Fill beaker/flask to 1000 mL with nanopure water

3. French press the re-suspended cells. Do not use the French press unless you have been trained to use it because the French press is dangerous if you do not use it properly and you can break it if you do not use it properly. The French press procedure is as follows:
    a. Get the French press apparatus and clean it well with 70% ethanol and DI water. Chill it on ice if it is not already cold to make it cold—having the apparatus cold will help keep the cells cold.
    b. Assemble the pieces and apply the cells to the chamber. Run the cells through the machine 2x to make sure that they are fully lysed.
    c. Put the lysed cells in a 50 mL centrifuge tube.

   d. Centrifuge the lysed cells for 20 minutes (5000 rpm [we can not go faster than this because the top will not fit on top of the centrifuge tubes, and it is dangerous to centrifuge faster than this without the top on], 4°C). Collect the supernatant. Filter the supernatant (into a 50 mL tube) to get rid of the large particles in solution (the filtered supernatant is called the 'crude'). Save the cell pellets (on ice) in case we do not get any protein because we may have to run the cells through the French press again if they were not lysed well. *Note that there should be less pellet volume now compared to before because the stuff inside of the cells should now be out in the solution.*

4. *During centrifugation of the lysed cells:* Add TALON cobalt affinity resin (stored in 4°C) to a plastic column (~2 mL resin for 2 liter cell culture). Wash the cobalt resin with water and then lysis buffer (5-10x resin volume) to get rid of the ethanol that the resin is stored in.
5. Cap the end of the column. Add the crude (filtered supernatant) to the resin. Cap the other end of the column, and mix the resin and the crude together on the rotor at 4°C for 1 hour.
6. Hang the column in the refrigerator, uncap the ends of the column, and collect the flow through (in case the protein did not bind to the resin).
7. With the column still hanging in the refrigerator, wash the protein-bound cobalt resin with lysis buffer (5-10x resin volume). You do not have to collect the flow through here.

8. Elute the protein with elution buffer 1. (You can do this on the lab bench, but work quickly so the protein does not warm up much. Eluting the protein with cold elution buffer will help keep the protein cold.) Collect the fractions that contain MutS protein (test for the presence of protein using the Bradford protein assay [Biorad]). You do not have to collect the first X mL (where X = resin volume in mL – 1 mL) of the eluted protein because nothing will be there. Start collecting protein in 0.75 mL fractions after 4 mL has been eluted. Not much will be in the first mL. Fractions 2-4 will be pretty dirty, but there will be a lot of protein. Fractions 5-? should be more pure. Fractions 2-10 will be run on a SDS PAGE gel to check that the protein is there and is not degraded. Next, the protein will be dialyzed and passed through an anion exchange column for the second step of protein purification.

   Elution buffer 1: Lysis buffer + 300 mM imidazole
9. Cobalt affinity column (TALON): After eluting the protein, cap the ends of the column and put it in the refrigerator. We will recharge the resin later (instructions for recharging the resin are included at the end of this section).
10. Dialyze the protein that you will pass through the anion affinity column against anion exchange wash buffer (dialysis buffer A) for 1 hour (do not dialyze the protein longer or else it will precipitate):

    Anion Exchange Wash Buffer (dialysis buffer A):

    20 mM HEPES, pH 7.8

    250 mM NaCl

11. Prepare the anion exchange column for protein purification. Note: The maximum flow rate for this column is 3 mL/minute, so DO NOT exceed this flow rate or you will damage the column. You should set up the Harvard apparatus pump to help you flow buffer through the column at the appropriate rate. Take the column labeled 'E. coli MutS' from the refrigerator and:
    a. Wash with 5-10 × column volume nanopure water.
    b. Wash with 5-10 × column volume **dialysis buffer A**
    c. Add the dialyzed protein to the column. Collect the protein flow through in case the protein does not bind to the column.
    d. Wash with 5-10 × column volume **dialysis buffer A**. Collect 'wash' flow through.
    e. Elute the protein with anion exchange elution buffer and collect the protein in 0.5-0.75 mL fractions.

       Anion exchange elution buffer:

       20 mM HEPES, pH 7.8

       400 mM NaCl
12. Clean the anion exchange column by washing it with 5x column volume of (1) elution buffer, (2) nanopure water, and (3) 20% ethanol (for storage).
13. Run the fractions on an SDS page gel to determine which fractions are pure enough (>95% pure) to dialyze. Then, you need to go ahead and do dialysis because the protein is not in its holding buffer after elution, so it will precipitate if it stays in the elution buffer too long. Always take a picture of the SDS PAGE gel

after you get it (you can destain overnight and take the picture the next morning so it will look nice).

14. Dialyze appropriate fractions in dialysis buffer B overnight at 4°C. The purpose of dialysis is to get rid of the salt used to elute the protein and to put the protein in the holding buffer that is appropriate for it.

Dialysis buffer B (Peggy Hsieh's dialysis buffer):

20 mM HEPES, pH 7.8 -- 40 mL of 0.5 M HEPES pH 7.8 stock
100 mM NaCl -- 5.84 g NaCl
1 mM DTT -- 0.15 g DTT
1 mM EDTA -- 5 mL of 0.2 M EDTA stock
30% glycerol -- 300 mL glycerol
*Fill beaker to the 1000 mL line with nanopure water

Dialysis procedure:

i. Put on gloves and get out dialysis tubing from the refrigerator.
ii. Cut off a section of the tubing with scissors and pre-wet the tubing by letting it sit in the dialysis buffer (which should be sitting on ice) for awhile and then making sure that the inside of the tubing is wet.
iii. Clamp off one end of the tubing and check for leakage by adding a couple mL of holding buffer to the tubing.
iv. Dump out the buffer, add MutS to the tubing, and clamp off the other end of the tubing.

   v. Add a stir bar to the buffer, put the flask on a stir plate in the refrigerator, and leave the tubing/protein in the holding buffer overnight.
   vi. After dialysis, each protein fraction should be put in 1.5 mL tubes and put on ice in the refrigerator until it is time to determine their concentrations.

Day 4 (morning)

15. Next morning: Get the dialyzed protein and filter it with a 0.2 μM filter (use the small ones so you loose less protein). It is important to filter the protein so that no bacteria will grow in the solution over time. Measure the protein concentration using the Biorad protein assay using BSA as the standard.

    General notes on protein purification:

    As a reducing agent during protein purification, people use βME instead of DTT (even if DTT is in the final holding buffer) because DTT destroys colbalt and nickel affinity resins (it will not be able to be recharged). EDTA should not be included in any washing buffers because the EDTA will strip the resin (we use EDTA to strip the resin completely before we recharge it).

Protocol for recharging the cobalt resin (cobalt resin washing, reuse, and storage):

All details for recharging the cobalt resin (TALON) can be found in the TALON metal affinity resins user manual (Clontech).

1. Pack all used resin into a large glass column.

2. Mix resin with four resin volumes of 6M guanidinium (pH 5.0) + 1% nonionic detergent, and stir for 10 minutes

    Add 57.318 g Guanidinium + 1 mL Tween20 (cut off the pipette tip to get this out because it is very sticky) to a container and fill the container up to the 100 mL line. Guanidinium is harmful so be careful and, as always, wear gloves.

3. Apply the solution to the glass column. Use water to help you get all the resin into the column by washing the beaker well.
4. Wash the resin with 5-10 column volumes of distilled water.

    Wash the resin with 5-10 column volumes of 20 mM MES buffer (pH 5.0) containing 0.1M NaCl (20 mM MES buffer, pH 5.0 + 0.1 M NaCl: MES 3.904 g/L, NaCl 5.844 g/L)

5. Wash the resin with 5-10 column volumes of distilled water.
6. Strip the resin of cobalt ions with 5 bed volumes of 0.2 M EDTA (pH 8.0) (0.2 M EDTA: 74.44 g/L)
7. Wash excess EDTA from the resin with an additional 10 bed volumes of distilled water.
8. Charge the resin with 10 column volumes of 50 mM $CoCl_2$ solution ($CoCl_2$: 11.9 g/L)
9. Wash the resin with 7 bed volumes of distilled water followed by 3 bed volumes of 300 mM NaCl and then 3 bed volumes of distilled water to remove excess cobalt metal ions. (To make 300 mM NaCl stock, add 17.532 g NaCl to 1L water.)

10. Store the resin at 4°C in 20% ethanol.

B.2 Ensemble characterization of *E. coli* MutS

E. coli MutS was purified by a two step purification procedure (described in **Appendix B.1**) and was resolved by PAGE in a single band according to the expected molecular mass, ~97 kDa (**Figure B.2.1**).

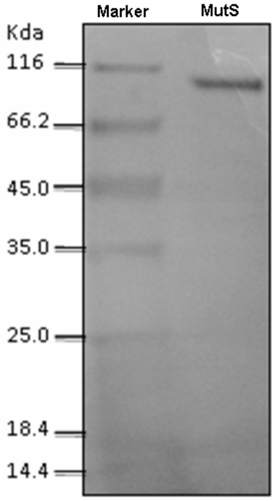


Figure B.2.1
The SDS page gel shows that the two-step purification procedure results in pure histidine-tagged MutS (~97 kDa) that has no degradation.

Since *E. coli* MutS is known to have high affinity for G/T mismatched DNA and 10-1000x lower affinity for homoduplex DNA [18], we used fluorescence anisotropy to characterize MutS's mismatch binding affinity and specificity for a G/T mismatch. The sequences of the 18 bp duplexes in the anisotropy assays were identical to those in the smFRET assays, and a TAMRA dye was attached to one end of the duplex to serve as the fluorescent reporter. MutS binds with high affinity to an 18 base pair duplex with a centrally positioned G/T mismatch (GT18) (K_D = 7.0 ± 1.4 nM) and to the same G/T mismatched duplex with a single 15 nucleotide polydT 3' tail (GT10-ext) (K_D = 6.6 ± 1.1

nM) (**Figure B.2.2**). These dissociation constants are in agreement with the K_D values that have been reported for *E. coli* MutS and its homologs [27, 42, 124-126], and these results demonstrate that the binding affinity for a G/T mismatch is not significantly affected by the presence of a polythymine tail. MutS binds with intermediate affinity to a C/T mismatched substrate (CT18) (K_D = 75 nM ± 3.0), in agreement with previous studies (ref). The mismatch binding activities for an 18 bp G/C control duplex (GC18) and an 18 bp control duplex with a 15 base polythymine 3' tail (GC18-ext) are too low to be measured in this assay. The low affinity of MutS for homoduplex DNA demonstrates that MutS binds specifically mismatched duplexes, and MutS has a low affinity for the duplex ends and the single-stranded polythymine tail.

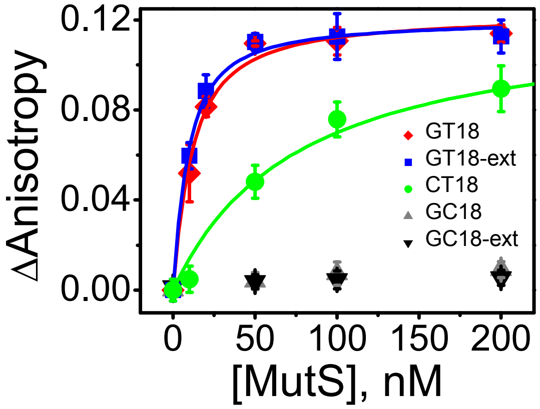


Figure B.2.2
Fluorescence anisotropy was used to characterize MutS's mismatch binding affinity and specificity for various substrates. The change in anisotropy (ΔAnisotropy) is plotted against MutS concentration (in nM) and fit with a nonlinear regression 1:1 binding equation to determine the dissociation constants (K_D's) reported in the text.

To test the ability of the purified MutS to bind and hydrolyze ATP, the steady-state ATPase activity of MutS was determined using a spectrophotometric assay. The ATPase activity was measured in the absence of DNA (k_{cat}=0.32±0.5min^{-1}), in the presence of 18 bp homoduplex DNA (GC18) (k_{cat}=1.9±0.5min^{-1}), in the presence of circular form duplex DNA pUC19 (k_{cat}=2.0±0.4min^{-1}), in the presence of 18 bp G/T mismatched DNA (GT18) (k_{cat}=3.9 ±0.5min^{-1}), and in the presence of 18 bp C/T mismatched DNA (k_{cat}=4.2±0.6min^{-1}) (**Figure B.2.3**). The sequences of the GC18, GT18, and CT18 substrates are the same as the sequences used in the anisotropy and smFRET experiments. The results of the ATPase experiments show that MutS has the expected mismatch-stimulated ATPase activity, as the as the ATPase rate of GT18 is ~2.0x that of GC18. In addition, since the ATPase rate of GC18 is the same as pUC19, the ends of the DNA are not responsible for stimulating ATPase activity. Finally, the G/T and C/T mismatches stimulate the ATPase activity of MutS equivalently, within the experimental error. These results are all in agreement with previously reported steady-state ATPase experiments (references).

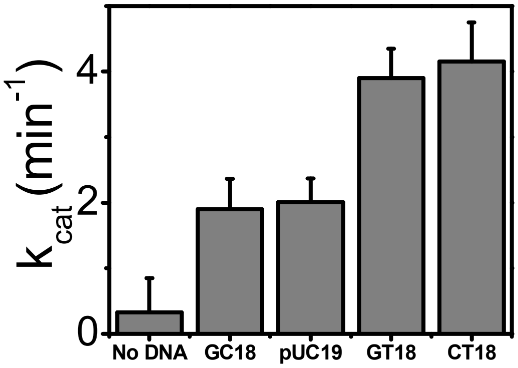


Figure B.2.3
A spectrophotometric assay was used to determine the k_{cat} for MutS with no DNA, with GC18 (18 bp homoduplex) DNA, with pUC19 (circular form homoduplex) DNA, with GT18 (18 bp G/T mismatched) DNA, and with CT18 (18 bp C/T mismatched) DNA.

Together, these biochemical assays demonstrate that the MutS protein is pure, active, and that the MutS-DNA system used in this study has characteristic mismatch binding and ATPase behavior.

The experimental procedures for the ensemble measurements are as follows:

Ensemble fluorescence measurements

Ensemble measurements were performed in MutS binding buffer (20 mM HEPES [pH 7.8], 50 mM NaCl, 1 mM DTT, 5 mM MgCl2, and 0.1 mg/ml BSA) using a FluoroMax-4 fluorometer (Jobin-Yvon, Horiba). All experiments were carried out at room temperature (~22°C) in a quartz fluorometer cell (Starna Cells, Inc.). For all fluorescence anisotropy experiments, TAMRA-labeled dsDNA was excited at 540 nm, and the fluorescence emission was detected at 583 nm. In the binding experiments, 0-400 nM MutS (with or without nucleotides) was incubated for 5 minutes with 5 nM TAMRA-labeled dsDNA and the anisotropy value was measured for 10 minutes. Relative anisotropy verses MutS concentration was plotted and fit with a nonlinear regression 1:1 binding equation to determine the dissociation constant KD. The equation is as follows:

$$Y_{MutS} = \left(\frac{K_D + [MutS] + [DNA] - \sqrt{\left(K_D + [MutS] + [DNA]\right)^2 - 4[MutS][DNA]}}{2[DNA]} \right) Y_{\max}$$

where [MutS] is the concentration of the MutS (in dimers), [DNA] is the concentration of TAMRA-labeled dsDNA, $Y_{\max}$ is the maximum value of MutS at saturation, and Y_{MutS} is the MutS value at a given [MutS]. The anisotropy measurements were performed three times at each concentration, and the values were averaged.

ATPase Assay

The hydrolysis of ATP by MutS was measured spectrophotometrically using a malachite green assay. MutS (100 nM) was incubated with increasing concentrations of ATP (5, 10, 20, 50, 100, 200 and 500 μM) in a 80 μl reaction mixture containing 25 mM HEPES (pH 7.5), 10 mM MgCl2, 140 mM NaCl, 15% glycerol, and 1 mM DTT at 37°C for 30 min. The DNA dependency of the ATPase was investigated using various forms of DNA (250 nM of HD18, HD40, GT18 and GT40). After incubation, the reactions were stopped by the addition of 80 µl of 25 nM EDTA. Then, the reactions were added with 40 µl of Malachite Green reagent (Biochain). The quenched reaction mixtures were incubated for 30 min at room temperature (25 °C), and the absorbance of the samples was measured at 620 nm.

Oligonucleotides

Oligonucleotides were purchased from Integrated DNA Technologies (IDT). Complementary strands were annealed in a 1:1.2 ratio by heating the strands at 80°C for three minutes and cooling them slowly to room temperature. Substrates used for ensemble fluorescence studies were purified using PAGE. The sequences are as follows:

Fluorescence Anisotropy substrates:

TAMRA-GT18

5’ TGG CGA CGG TAG CGA GGC TAMRA 3’

5’ GCC TCG CTG CCG TCG CCA 3’

TAMRA-GT18ext

5’ TGG CGA CGG TAG CGA GGC TAMRA 3’

5’ GCC TCG CTG CCG TCG CCA TTT TTT TTT TTT TTT 3’

TAMRA-GC18

5’ TGG CGA CGG CAG CGA GGC TAMRA 3’

5’ GCC TCG CTG CCG TCG CCA 3’

TAMRA-GC18ext

5’ TGG CGA CGG CAG CGA GGC TAMRA 3’

5’ GCC TCG CTG CCG TCG CCA TTT TTT TTT TTT TTT 3’

TAMRA-CT18

5’ TGG CGA CGG TAG CGA GGC TAMRA 3’

5’ GCC TCG CTC CCG TCG CCA 3’

ATPase substrates:

Unlabeled-GT18

5’ TGG CGA CGG TAG CGA GGC 3’

5’ GCC TCG CTG CCG TCG CCA 3’

Unlabeled-GC18

5’ TGG CGA CGG CAG CGA GGC 3’

5’ GCC TCG CTG CCG TCG CCA 3’

Unlabeled-CT18

5’ TGG CGA CGG TAG CGA GGC 3’

5’ GCC TCG CTC CCG TCG CCA 3’

B.3 Ensemble and single-molecule characterization of *E. coli* MutS-E694A

The mutant MutS that was used in this study, MutS-E694A, was created using site-directed mutagenesis to change the glutamic acid-694 to an alanine. This protein has been shown before to have the same mismatch binding activity and ATP binding activity as wild-type MutS, but it cannot hydrolyze ATP [33]. To demonstrate that the MutS-

E694A used in this study had these properties, we determined the mutant's mismatch binding activity and steady state ATPase activity. First, TAMRA-labeled GT18 DNA was titrated with increasing amounts of MutS-E694A, and the K_D was determined to be 6.3 ± 1.7 nM (**Figure B.3.1A**), which is the same (within error) as the reported wild-type MutS in complex with GT18 DNA ($K_{D,wt\text{-}MutS} = 7.0 \pm 1.4$ nM). Next, the steady-state ATP hydrolysis activity was measured using a spectrophotometric assay. No ATP hydrolysis activity was measured for MutS-E694A only, for MutS-E694A in the presence of homoduplex DNA (GC18), or for MutS-E694A in the presence of G/T mismatched DNA (GT18) (**Figure B.3.1B**).

Then, a single-molecule FRET experiment was done to ensure that MutS-E694A has similar bending activity as wild-type MutS in complex with the GT18 substrate. The distribution of FRET efficiency values and the dynamics of individual GT18 molecules were determined for hundreds of FRET pairs in the absence and presence of MutS-E694A. In the absence of protein, the GT18 substrate has a single narrow peak at $E_{FRET} \sim 0.17$ (**Figure B.3.1C**, DNA only), and MutS-E694A bends the DNA, increasing the FRET efficiency peak to $E_{FRET} \sim 0.29$ (**Figure B.3.1C**, +20 nM MutS).

Thus, the MutS-E694A used in the studies presented here have the same DNA binding activity as wild-type MutS, have no ATP hydrolysis activity, and bend the DNA in the same manner as wild-type MutS.

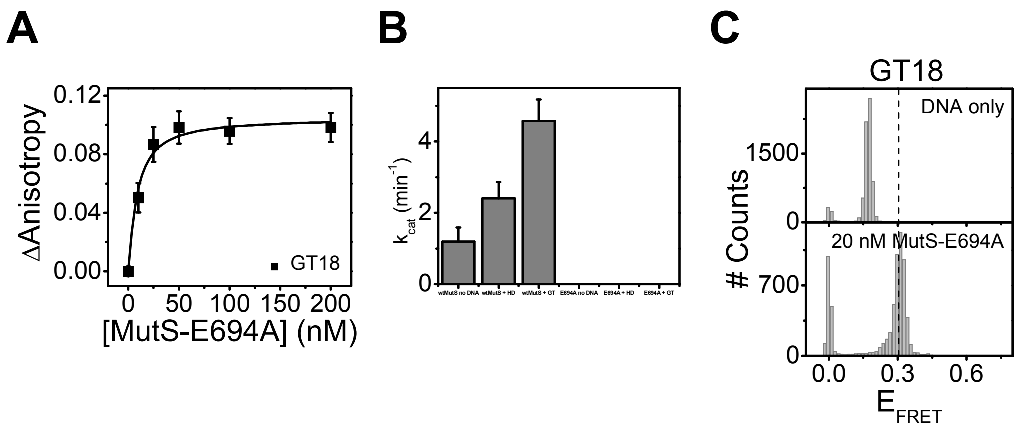


Figure B.3.1

(A) Fluorescence anisotropy was used to characterize MutS-E694A's mismatch binding affinity and specificity for various substrates. The change in anisotropy (ΔAnisotropy) is plotted against MutS-E694A concentration (in nM) and fit with a nonlinear regression 1:1 binding equation to determine the dissociation constants (K_D's) reported in the text. (B) A spectrophotometric assay was used to determine the k_{cat} for MutS-E694A with no DNA, with GC18 (18 bp homoduplex) DNA, and with GT18 (18 bp G/T mismatched) DNA. (C) FRET efficiency (E_{FRET}) histograms for the G/T mismatched DNA (GT18) in the absence of MutS-E694A and in the presence of 20 nM MutS-E694A are shown.

References

1. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science, 1999. **283**(5408): p. 1676-1683.
2. Bates, M., T.R. Blosser, and X.W. Zhuang, *Short-range spectroscopic ruler based on a single-molecule optical switch.* Physical Review Letters, 2005. **94**(10).
3. McKinney, S.A., C. Joo, and T. Ha, *Analysis of single-molecule FRET trajectories using hidden Markov modeling.* Biophysical Journal, 2006. **91**(5): p. 1941-1951.
4. Joo, C. and H. T., *Single-molecule FRET with Total Internal Reflection Microscopy*. Single-Molecule Techniques: A Laboratory Manual, ed. P.R. Selvin and T. Ha. 2008: Cold Spring Harbor Laboratory Press.
5. Ha, T., *Single-molecule fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 78-86.
6. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nature Methods, 2008. **5**(6): p. 507-516.
7. Joo, C., et al., *Exploring rare conformational species and ionic effects in DNA Holliday junctions using single-molecule spectroscopy.* Journal of Molecular Biology, 2004. **341**(3): p. 739-751.
8. McKinney, S.A., et al., *Structural dynamics of individual Holliday junctions.* Nature Structural Biology, 2003. **10**(2): p. 93-97.
9. McKinney, S.A., et al., *Observing spontaneous branch migration of Holliday junctions one step at a time.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(16): p. 5715-5720.
10. Ha, T., et al., *Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase.* Nature, 2002. **419**(6907): p. 638-641.
11. Myong, S., et al., *Repetitive shuttling of a motor protein on DNA.* Nature, 2005. **437**(7063): p. 1321-1325.
12. Myong, S., et al., *Blockade-induced repetition of ssDNA translocation by a DNA motor.* Biophysical Journal, 2005. **88**(1): p. 209a-209a.
13. Rasnik, I., et al., *Branch migration enzyme as a Brownian ratchet.* Embo Journal, 2008. **27**(12): p. 1727-1735.
14. Rasnik, I., et al., *Single molecule assays for DNA binding and unwinding mechanisms of helicases.* Biophysical Journal, 2002. **82**(1): p. 49a-49a.
15. Rasnik, I., et al., *Site-specific fluorescent labeling of E-coli Rep helicase for single molecule studies.* Biophysical Journal, 2003. **84**(2): p. 301a-301a.
16. Rasnik, I., et al., *DNA-binding orientation and domain conformation of the E-coli Rep helicase monomer bound to a partial duplex junction: Single-molecule studies of fluorescently labeled enzymes.* Journal of Molecular Biology, 2004. **336**(2): p. 395-408.
17. Hsieh, P., *Molecular mechanisms of DNA mismatch repair.* Mutation Research-DNA Repair, 2001. **486**(2): p. 71-87.
18. Kunkel, T.A. and D.A. Erie, *DNA mismatch repair.* Annual Review of Biochemistry, 2005. **74**: p. 681-710.

19. Lyer, R.R., et al., *DNA mismatch repair: Functions and mechanisms.* Chemical Reviews, 2006. **106**(2): p. 302-323.
20. Lamers, M.H., et al., *The crystal structure of DNA mismatch repair protein MutS binding to a G center dot T mismatch.* Nature, 2000. **407**(6805): p. 711-717.
21. Obmolova, G., et al., *Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA.* Nature, 2000. **407**(6805): p. 703-710.
22. Natrajan, G., et al., *Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates.* Nucleic Acids Research, 2003. **31**(16): p. 4814-4821.
23. Warren, J.J., et al., *Structure of the human MutS alpha DNA lesion recognition complex.* Molecular Cell, 2007. **26**(4): p. 579-592.
24. Gupta, S., M. Gellert, and W. Yang, *Mechanism of mismatch recognition revealed by human MutS beta bound to unpaired DNA loops.* Nature Structural & Molecular Biology, 2012. **19**(1): p. 72-U91.
25. Yang, W., *Structure and mechanism for DNA lesion recognition.* Cell Research, 2008. **18**(1): p. 184-197.
26. Parker, B.O. and M.G. Marinus, *Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia Coli.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(5): p. 1730-1734.
27. Su, S.S., et al., *Mismatch specificity of methyl-directed DNA mismatch correction in vitro.* Journal of Biological Chemistry, 1988. **263**(14): p. 6829-6835.
28. Su, S.S. and P. Modrich, *Escherichia Coli MutS-encoded protein binds to mismatched DNA base pairs.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(14): p. 5057-5061.
29. Owen, B.A.L., et al., *(CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition.* Nature Structural & Molecular Biology, 2005. **12**(8): p. 663-670.
30. Owen, B.A.L., et al., *(CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition (vol 12, pg 663, 2005).* Nature Structural & Molecular Biology, 2005. **12**(9): p. 824-824.
31. Acharya, S., et al., *The coordinated functions of the E-coli MutS and MutL proteins in mismatch repair.* Molecular Cell, 2003. **12**(1): p. 233-246.
32. Antony, E. and M.M. Hingorani, *Asymmetric ATP binding and hydrolysis activity of the Thermus aquaticus MutS dimer is key to modulation of its interactions with mismatched DNA.* Biochemistry, 2004. **43**(41): p. 13115-13128.
33. Baitinger, C., V. Burdett, and P. Modrich, *Hydrolytically deficient MutS E694A is defective in the MutL-dependent activation of MutH and in the mismatch-dependent assembly of the MutS center dot MutL center dot heteroduplex complex.* Journal of Biological Chemistry, 2003. **278**(49): p. 49505-49511.
34. Bjornson, K.P., D.J. Allen, and P. Modrich, *Modulation of MutS ATP hydrolysis by DNA cofactors.* Biochemistry, 2000. **39**(11): p. 3176-3183.
35. Bjornson, K.P. and P. Modrich, *Differential and simultaneous adenosine di- and triphosphate binding by MutS.* Journal of Biological Chemistry, 2003. **278**(20): p. 18557-18562.

36. Blackwell, L.J., et al., *Distinct MutS DNA-binding modes that are differentially modulated by ATP binding and hydrolysis.* Journal of Biological Chemistry, 2001. **276**(36): p. 34339-34347.
37. Blackwell, L.J., et al., *Nucleotide-promoted release of hMutS alpha from heteroduplex DNA is consistent with an ATP-dependent translocation Mechanism.* Journal of Biological Chemistry, 1998. **273**(48): p. 32055-32062.
38. Jacobs-Palmer, E. and M.M. Hingorani, *The effects of nucleotides on MutS-DNA binding kinetics clarify the role of MutS ATPase activity in mismatch repair.* Journal of Molecular Biology, 2007. **366**(4): p. 1087-1098.
39. Jeong, C., et al., *MutS switches between two fundamentally distinct clamps during mismatch repair.* Nature Structural & Molecular Biology, 2011. **18**(3): p. 379-U174.
40. Joshi, A., S. Sen, and B.J. Rao, *ATP-hydrolysis-dependent conformational switch modulates the stability of MutS-mismatch complexes.* Nucleic Acids Research, 2000. **28**(4): p. 853-861.
41. Lamers, M.H., et al., *ATP increases the affinity between MutS ATPase domains - Implications for atp hydrolysis and conformational changes.* Journal of Biological Chemistry, 2004. **279**(42): p. 43879-43885.
42. Schofield, M.J., et al., *Interaction of Escherichia coli MutS and MutL at a DNA mismatch.* Journal of Biological Chemistry, 2001. **276**(30): p. 28291-28299.
43. Gradia, S., et al., *hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA.* Molecular Cell, 1999. **3**(2): p. 255-261.
44. Allen, D.J., et al., *MutS mediates heteroduplex loop formation by a translocation mechanism.* Embo Journal, 1997. **16**(14): p. 4467-4476.
45. Gradia, S., S. Acharya, and R. Fishel, *The role of mismatched nucleotides in activating the hMSH2-hMSH6 Molecular Switch.* Journal of Biological Chemistry, 2000. **275**(6): p. 3922-3930.
46. Iaccarino, I., et al., *Mutation in the magnesium binding site of hMSH6 disables the hMutS alpha sliding clamp from translocating along DNA.* Journal of Biological Chemistry, 2000. **275**(3): p. 2080-2086.
47. Mazur, D.J., M.L. Mendillo, and R.D. Kolodner, *Inhibition of Msh6 ATPase activity by mispaired DNA induces a Msh2(ATP)-Msh6(ATP) state capable of hydrolysis-independent movement along DNA.* Molecular Cell, 2006. **22**(1): p. 39-49.
48. Junop, M.S., et al., *Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair.* Molecular Cell, 2001. **7**(1): p. 1-12.
49. Lakowicz , J., *Principles of Fluorescence Spectroscopy*. 4th ed. 2006: Springer.
50. Berland, K.M., *Basics of fluorescence*. Methods in Cellular Imaging, ed. A. Periasamy. 2001.
51. Clegg, R.M., *Fluorescence resonance energy transfer and nucleic acids.* Methods in Enzymology, 1992. **211**: p. 353-388.
52. Murphy, M.C., et al., *Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy.* Biophysical Journal, 2004. **86**(4): p. 2530-2537.
53. Ha, T., et al., *Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor.*

Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(13): p. 6264-6268.

54. Dave, R., et al., *Mitigating Unwanted Photophysical Processes for Improved Single-Molecule Fluorescence Imaging.* Biophysical Journal, 2009. **96**(6): p. 2371-2381.
55. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and longlasting single-molecule fluorescence imaging.* Nature Methods, 2006. **3**(11): p. 891-893.
56. Rasnik, I., S.A. Mckinney, and T. Ha, *Surfaces and orientations: Much to FRET about?* Accounts of Chemical Research, 2005. **38**(7): p. 542-548.
57. Yildiz, A., et al., *Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization.* Science, 2003. **300**(5628): p. 2061-2065.
58. Yildiz, A., et al., *Kinesin walks hand-over-hand.* Science, 2004. **303**(5658): p. 676-678.
59. Smith, S.B., Y.J. Cui, and C. Bustamante, *Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules.* Science, 1996. **271**(5250): p. 795-799.
60. Strick, T.R., et al., *The elasticity of a single supercoiled DNA molecule.* Science, 1996. **271**(5257): p. 1835-1837.
61. Funatsu, T., et al., *Imaging of single fluorescent molecules and individual ATP turnovers by single miosin molecules in aqueous solution.* Nature, 1995. **374**(6522): p. 555-559.
62. Sass, L.E., et al., *Single-Molecule FRET TACKLE Reveals Highly Dynamic Mismatched DNA-MutS Complexes.* Biochemistry, 2010. **49**(14): p. 3174-3190.
63. Vamosi, G., C. Gohlke, and R.M. Clegg, *Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: Multiple conformers of single-stranded and double-stranded dye-DNA complexes.* Biophysical Journal, 1996. **71**(2): p. 972-994.
64. Iqbal, A., et al., *Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nuclelic acids.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(32): p. 11176-11181.
65. Norman, D.G., et al., *Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies.* Biochemistry, 2000. **39**(21): p. 6317-6324.
66. Benesch, R.E. and R. Benesch, *Enzymatic removal of oxygen for polargraphy and related methods.* Science, 1953. **118**(3068): p. 447-448.
67. Deniz, A.A., et al., *Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Forster distance dependence and subpopulations.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(7): p. 3670-3675.
68. Lilley, D.M.J., *Structures of helical junctions in nucleic acids.* Quarterly Reviews of Biophysics, 2000. **33**(2): p. 109-159.
69. Oussatcheva, E.A., et al., *Structure of branched DNA molecules: Gel retardation and atomic force microscopy studies.* Journal of Molecular Biology, 1999. **292**(1): p. 75-86.

70. Pearson, C.E. and R.R. Sinden, *Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci.* Biochemistry, 1996. **35**(15): p. 5041-5053.
71. Pearson, C.E. and R.R. Sinden, *Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA.* Current Opinion in Structural Biology, 1998. **8**(3): p. 321-330.
72. Mitas, M., *Trinucleotide repeats associated with human disease.* Nucleic Acids Research, 1997. **25**(12): p. 2245-2253.
73. Wells, R.D., et al., *Small slipped register genetic instabilities in Escherichia coli in triplet repeat sequences associated with hereditary neurological diseases.* Journal of Biological Chemistry, 1998. **273**(31): p. 19532-19541.
74. Duckett, D.R. and D.M.J. Lilley, *The three-way DNA junction is a Y-shaped molecule in which there is no helix helix stacking.* Embo Journal, 1990. **9**(5): p. 1659-1664.
75. Guo, Q., et al., *Asymmetric Structure of a Three-arm DNA Junction.* Biochemistry, 1990. **29**(49): p. 10927-10934.
76. Lu, M., Q. Guo, and N.R. Kallenbach, *Effect of sequence on the structure of 3-arm DNA junctions.* Biochemistry, 1991. **30**(24): p. 5815-5820.
77. Stuhmeier, F., D.M.J. Lilley, and R.M. Clegg, *Effect of additional unpaired bases on the stability of three-way DNA junctions studied by fluorescence techniques.* Biochemistry, 1997. **36**(44): p. 13539-13551.
78. Shlyakhtenko, L.S., et al., *Structure and dynamics of three-way DNA junctions: atomic force microscopy studies.* Nucleic Acids Research, 2000. **28**(18): p. 3472-3477.
79. Duckett, D.R., et al., *The Structure of the Holliday Junctions, and Its Resolution.* Cell, 1988. **55**(1): p. 79-89.
80. Leontis, N.B., W. Kwok, and J.S. Newman, *Stability and structure of three-way DNA junctions containing unpaired nucleotides.* Nucleic Acids Research, 1991. **19**(4): p. 759-766.
81. Leontis, N.B., et al., *Helical Stacking in DNA Three-Way Junctions Containing Two Unpaired Pyrimidines: Proton NMR Studies.* Biophysical Journal, 1995. **68**(1): p. 251-265.
82. Overmars, F.J.J., et al., *NMR studies of DNA three-way junctions containing two unpaired thymidine bases: The influence of the sequence at the junction on the stability of the stacking conformers.* Journal of Molecular Biology, 1996. **255**(5): p. 702-713.
83. Rosen, M.A. and D.J. Patel, *Structural Features of a Three-Stranded DNA Junction Containing a C-C Junctional Bulge.* Biochemistry, 1993. **32**(26): p. 6576-6587.
84. Stuhmeier, F., et al., *Global structure of three-way DNA junctions with and without additional unpaired bases: A fluorescence resonance energy transfer analysis.* Biochemistry, 1997. **36**(44): p. 13530-13538.
85. van Buuren, B.N.M., et al., *Solution structure of a DNA three-way junction containing two unpaired thymidine bases. Identification of sequence features that decide conformer selection.* Journal of Molecular Biology, 2000. **304**(3): p. 371-383.

86. Welch, J.B., D.R. Duckett, and D.M.J. Lilley, *Structures of bulged three-way DNA junctions.* Nucleic Acids Research, 1993. **21**(19): p. 4548-4555.
87. Welch, J.B., F. Walter, and D.M.J. Lilley, *Two Inequivalent Folding Isomers of the Three-way DNA Junction with Unpaired Bases: Sequence-dependence of the Folded Conformation.* Journal of Molecular Biology, 1995. **251**(4): p. 507-519.
88. Yang, M.S. and D.P. Millar, *Conformational flexibility of three-way DNA junctions containing unpaired nucleotides.* Biochemistry, 1996. **35**(24): p. 7959-7967.
89. Zhong, M., et al., *Effects of Unpaired Bases on the Conformation and Stability of Three-Arm DNA Junctions.* Biochemistry, 1994. **33**(12): p. 3660-3667.
90. Ouporov, I.V. and N.B. Leontis, *Refinement of the solution structure of a branched DNA 3-way junction.* Biophysical Journal, 1995. **68**(1): p. 266-274.
91. Rosen, M.A. and D.J. Patel, *Conformational Differences between Bulged Pyrimidines (C-C) and Purines (A-A, I-I) at the Branch Point of Three-Stranded DNA Junctions.* Biochemistry, 1993. **32**(26): p. 6563-6575.
92. Pearson, C.E., et al., *Slipped-strand DNAs formed by long (CAG).(CTG) repeats: slipped-out repeats and slip-out junctions.* Nucleic Acids Research, 2002. **30**(20): p. 4534-4547.
93. Ma, R.I., et al., *Three-arm nucleic acid junctions are flexible.* Nucleic Acids Res, 1986. **14**(24): p. 9745-53.
94. Person, B., et al., *Correlated Movement and Bending of Nucleic Acid Structures Visualized by Multicolor Single-Molecule Spectroscopy.* Chemphyschem, 2009. **10**(9-10): p. 1455-1460.
95. Ross, J., et al., *Multicolor single-molecule spectroscopy with alternating laser excitation for the investigation of interactions and dynamics.* Journal of Physical Chemistry B, 2007. **111**(2): p. 321-326.
96. McKinney, S.A., et al., *Spontaneous and protein-induced structural dynamics of individual Holliday junctions.* Biophysical Journal, 2003. **84**(2): p. 303a-304a.
97. Markham, N.R. and M. Zuker, *DINAMelt web server for nucleic acid melting prediction.* Nucleic Acids Research, 2005. **33**: p. W577-W581.
98. Fishel, R., *The selection for mismatch repair defects in hereditary nonpolyposis colorectal cancer: Revising the mutator hypothesis.* Cancer Research, 2001. **61**(20): p. 7369-7374.
99. Peltomaki, P., *Deficient DNA mismatch repair: a common etiologic factor for colon cancer.* Human Molecular Genetics, 2001. **10**(7): p. 735-740.
100. Wei, K., R. Kucherlapati, and W. Edelmann, *Mouse models for human DNA mismatch-repair gene defects.* Trends in Molecular Medicine, 2002. **8**(7): p. 346-353.
101. Acharya, S., et al., *hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(24): p. 13629-13634.
102. Harrington, J.A. and R.D. Kolodner, *Saccharomyces cerevisiae Msh2-Msh3 acts in repair of base-base mispairs.* Molecular and Cellular Biology, 2007. **27**(18): p. 6546-6554.
103. Palombo, F., et al., *hMutS beta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion leaps in DNA.* Current Biology, 1996. **6**(9): p. 1181-1184.

104. Schofield, M.J. and P. Hsieh, *DNA mismatch repair: Molecular mechanisms and biological function.* Annual Review of Microbiology, 2003. **57**: p. 579-608.
105. Wilson, T., S. Guerrette, and R. Fishel, *Dissociation of mismatch recognition and ATPase activity by hMSH2-hMSH3.* Journal of Biological Chemistry, 1999. **274**(31): p. 21659-21664.
106. Owen, B.A.L., W.H. Lang, and C.T. McMurray, *The nucleotide binding dynamics of human MSH2-MSH3 are lesion dependent.* Nature Structural & Molecular Biology, 2009. **16**(5): p. 550-557.
107. Bill, C.A., et al., *Repair bias of large loop mismatches during recombination in mammalian cells depends on loop length and structure.* Mutation Research-DNA Repair, 2001. **485**(3): p. 255-265.
108. Genschel, J., et al., *Isolation of MutS beta from human cells and comparison of the mismatch repair specificities of MutS beta and MutS alpha.* Journal of Biological Chemistry, 1998. **273**(31): p. 19895-19901.
109. Marsischky, G.T., et al., *Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair.* Genes & Development, 1996. **10**(4): p. 407-420.
110. McMurray, C.T., *Mechanisms of trinucleotide repeat instability during human development.* Nature Reviews Genetics, 2010. **11**(11): p. 786-799.
111. Miret, J.J., L. PessoaBrandao, and R.S. Lahue, *Instability of CAG and CTG trinucleotide repeats in Saccharomyces cerevisiae.* Molecular and Cellular Biology, 1997. **17**(6): p. 3382-3387.
112. Brouwer, J.R., R. Willemsen, and B.A. Oostra, *Microsatellite repeat instability and neurological disease.* Bioessays, 2009. **31**(1): p. 71-83.
113. McMurray, C.T., *Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease.* DNA Repair, 2008. **7**(7): p. 1121-1134.
114. Tome, S., et al., *MSH2 ATPase Domain Mutation Affects CTG.CAG Repeat Instability in Transgenic Mice.* Plos Genetics, 2009. **5**(5).
115. Kim, H.M., et al., *Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair.* Embo Journal, 2008. **27**(21): p. 2896-2906.
116. Tian, L., et al., *Mismatch Recognition Protein MutS beta Does Not Hijack (CAG)(n) Hairpin Repair in Vitro.* Journal of Biological Chemistry, 2009. **284**(31): p. 20452-20456.
117. Antony, E. and M.M. Hingorani, *Mismatch recognitioh-coupled stabilization of Msh2-Msh6 in an ATP-bound state at the initiation of DNA repair.* Biochemistry, 2003. **42**(25): p. 7682-7693.
118. Kunkel, T.A. and R. Bebenek, *DNA replication fidelity.* Annual Review of Biochemistry, 2000. **69**: p. 497-529.
119. Jiricny, J., et al., *Mismatch-containing oligonucleotide duplexes bound by the Escherichia Coli MutS-encoded protein.* Nucleic Acids Research, 1988. **16**(16): p. 7843-7853.
120. Dohet, C., R. Wagner, and M. Radman, *Repair of defined single base-pair mismatches in Escherichia Coli.* Proceedings of the National Academy of Sciences of the United States of America, 1985. **82**(2): p. 503-505.

121. Jones, M., R. Wagner, and M. Radman, *Repair of a mismatch is influenced by the base composition of the surrounding nucleotide sequence.* Genetics, 1987. **115**(4): p. 605-610.
122. Kramer, B., W. Kramer, and H.J. Fritz, *Different base-base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch repair system of Escherichia Coli.* Cell, 1984. **38**(3): p. 879-887.
123. Feng, G. and M.E. Winkler, *Single-step Purifications of His(6)-MutH, His(6)-MutL and His(6)-MutS Repair Proteins of Escherichia coli K-12.* Biotechniques, 1995. **19**(6): p. 956-965.
124. Schofield, M.J., et al., *The Phe-X-Glu DNA binding motif of MutS - The role of hydrogen bonding in mismatch recognition.* Journal of Biological Chemistry, 2001. **276**(49): p. 45505-45508.
125. Yang, Y., et al., *Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions.* Nucleic Acids Research, 2005. **33**(13): p. 4322-4334.
126. Zhai, J. and M.M. Hingorani, *Saccharomyces cerevisiae Msh2-Msh6 DNA binding kinetics reveal a mechanism of targeting sites for DNA mismatch repair.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(2): p. 680-685.

Author's Biography

Julie E. Coats was born March 4, 1984 in Little Rock, Arkansas to Dr. Bobby and Eleanor Coats. She was raised in Conway, Arkansas where she attended high school at Conway High School and graduated with highest honors in 2002. She attended Hendrix College (Conway, Arkansas) where she majored in physics, did undergraduate research, and graduated *cum laude* with a B.A. in physics in 2006. In August 2006, she came to Emory University (Atlanta, Georgia) for graduate school in physics, and she joined the experimental molecular biophysics laboratory of Dr. Ivan Rasnik in October 2008.