SINGLE MOLECULE FRET STUDY ON THE MECHANISM OF RECA MEDIATED STRAND EXCHANGE

BY

KAUSHIK MADABHUSI RAGUNATHAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics and Computational Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Taekjip Ha, Chair, Director of Research
Professor Satish Nair
Assistant Professor Maria Spies
Assistant Professor Yann Chemla

Abstract

RecA plays a critical role during double strand break repair via homologous recombination. During the strand exchange reaction, RecA forms a helical filament on single stranded (ss) DNA that searches for homology and exchanges complementary base pairs with a homologous double strand (ds) DNA to form a new heteroduplex. The study of strand exchange in ensemble assays is limited by the diffusion limited homology search process which masks the subsequent strand exchange reaction. We developed a single molecule fluorescence assay with a few basepair and milliseconds resolution which can separate initial docking from the subsequent propagation of joint molecule formation. Our data suggests that propagation occurs in 3 bp increments with destabilization of the incoming dsDNA and concomitant pairing with the reference ssDNA. Our model for strand exchange links structural models of RecA to its catalytic function.

Next, we investigated the mechanism of RecA mediated homology search. Using tools with high spatiotemporal resolution to observe the encounter complex between the RecA filament and dsDNA, we present evidence in support of the “sliding model” wherein a RecA filament diffuses on a dsDNA track. Our results suggest that the sliding of the dsDNA relative to the RecA filament can explain the rapid changes in FRET which we have observed upon the docking of non-homologous dsDNA to the RecA filament. We further show that homology can be identified during such sliding. Sliding is thermally driven and occurs in the absence of ATP hydrolysis. Furthermore, homology recognition and basepairing can involve as few as 6 bp of complementarity. Our observation presents an example of how a multi-protein complex bound to DNA can serve as a vehicle enabling homology search processes via 1-D sliding.

Finally, we demonstrate how an extension of the two color FRET assay to measure four colors simultaneously allows us to measure the correlation of reaction completion between the two ends of a single synaptic complex. We expect that this method will enable a multi dimensional analysis of independent reaction coordinates with broad applications in measuring the correlated dynamics of more complex biological systems

Acknowledgements

I owe immense gratitude to my adviser Taekjip Ha. He has been inspirational in terms of his hard work, dedication and sincerity to research and science. I can only hope to imbibe some of these ideals in my own life and career as a scientist. His keen insight and continued support truly helped me move mountains.

Chirlmin Joo represented a special friend and colleague. His insight, attention to detail and passion for science inspire me to this day. The finish line is visible thanks to his constant support, mentorship and encouragement.

I thank my family for their unwavering faith in each of my decisions. My mother has taught me the value of perseverance in the face of difficulty and the fact that half the battle is won in truly believing that anything is possible. From my father, I learned about the value of simplicity, honesty and strength of character. They served as my bedrock guiding me through each of my endeavors. My sister, Malvi, both literally and figuratively brought about music and liveliness to my life in Champaign.

In Preeti, I found a unique outlet that helped me grow to become the person I am today. She and Chandu blurred the line for me between friendship and family. Anshu showed me the value of tenacity and determination. I ran two marathons just by virtue of his ability to inspire even in the most ordinary of circumstances. Thanks to Chamanti and Madhav for constantly urging me to demand more out of myself. Pradeep and Harini taught me how to see the world beyond the missteps and daily frustrations of my PhD. Their ability to find joy in the most mundane incidents painted Champaign in a different light. It seems almost like I spent a lifetime with roommate Vivek considering he's the one person with whom I've lived the longest and I thank him for all his support. Milu, Dog, Vinod, Taathu, Aloo, Vidisha, JP, Gowri, Nav, Gayatri and Vinay made Friday nights something Iooked forward to every week. Their endearing qualities constituted the reason why I came to regard Champaign as a second home.

Urmi was my rock towards the end of my PhD. Everything in my life seemed so much easier with her around. In her, I found a passion and kindness that makes me strive to be a better person each day.

My cousins Sriram, Sridhar and Sindhu were simply extraordinary in how understanding they were to the many times when my eccentricity and aloofness mingled with each other. My thanks to Meera perima and Murali perepa who got me interested in biology at an early age. Chitra perima, Raju perepa and Maya in many ways were my second family. The summers I spent in their home impacted me positively in subtle ways. Ammamma and Bobby deserve a huge chunk of my gratitude for giving me a home during my formative years.

I thank Ankur for giving me the space to debate and explore ideas. I surely registered a measurable intellectual development thanks to my conversations with him. Sultan and Hye Ran always showed tremendous patience and sage advice in the most trying situations. Sua, Peter, Reza, Ruobo, Yuji, Ben, Kyungsuk, Rahul, Jaya, Sinan, Jungmin, Salman and Prakrit were wonderful colleagues who not only challenged my way of thinking but also provided me with a supportive environment to carry out my work. Sungchul Hohng, Jinwoo and Hajin helped chart out new directions for my research agenda.

Cindy Dodds was my graduate secretary of Biophysics extraordinaire. Her attention to detail and her ability to quell the most apocalyptic scenarios I could conjure up always kept me on track.

My undergraduate mentors V. Raghunathan (RRI), S. Mahadevan (IISc), M.K. Mathew (NCBS) and V. Ramamurthy (PSG Tech) deserve thanks for providing me with valuable insight and advice all of which prompted me to embark on a fascinating journey of sorts.

And of course, thanks to a place called Champaign-Urbana. I never anticipated that these 6 years in Champaign doing a PhD would turn out to be the most beautiful years of my life.

Table of Contents

Chapter 1

Overview of DNA repair in *E.coli*

I later came to realize that DNA is so precious that probably many distinct repair mechanisms would exist
Francis Crick (1974)

1.1 Introduction

The ubiquitous presence of DNA damaging agents poses a constant threat to the integrity of the genome in all organisms. Cells have different pathways and mechanisms that are engaged in the detection, surveillance and repair of damaged DNA. In humans, malfunctions in DNA repair could result in the accumulation of mutations that could lead to cancer (Halazonetis et al., 2008). In bacteria, DNA damage and repair could cause mutations which may result in the development of altered traits such as antibiotic resistance that pose serious problems to human health (Cirz et al., 2005). Bacteria, archaea and eukaryotes encounter similar types of DNA lesions and have homologous pathways that restore DNA to its nascent state following the damage event. Hence, understanding DNA repair in *E.coli* lends insights that can be extrapolated to homologous repair proteins in humans.

1.2 DNA repair pathways in *E.coli*

Rather than a 'one size fits all' approach, *E.coli* employs specific DNA repair pathways to target the different types of lesions encountered in DNA. Here, I summarize some key features of DNA repair pathways (Davidsen and Tonjum, 2006) (Figure 1.1).

a. Base excision repair - Deamination or alkylation of bases invokes results in the formation of an abasic site. DNA glycosylase detects and removes the damaged nucleotide. For example, the formation of 8- oxoguanine via oxidative damage is detected and repaired by a DNA glycosylase, OGG1.

b. Mismatch repair – Misincorporation of a base during DNA replication engages the mismatch repair pathway consisting primarily of three proteins- MutS, MutH and MutL. The MutHSL system of proteins is involved in the detection of the site of damage and repair of the newly synthesized DNA based on differential methylation between the daughter and parental DNA strands.

c. Nucleotide excision repair – Thymine dimers formed as a consequence of UV damage are repaired by the UvrABC system of proteins that excise the damaged region and subsequently recruit polymerases that reinitiate DNA synthesis in the gap region.

d. Reversal of damage- This is a mechanism where a damaged nucleotide is directly converted to its undamaged state without the need for base excision mechanisms. Proteins such as photolyases mediate conversions of cyclobutane pyrimidine dimer (CPD) formed by crosslinking between adjacent thymine bases via UV exposure. In addition, *E.coli* possesses proteins that ensure that DNA damaging agents are eliminated from the cell. For example, superoxide dismutase (SOD) quenches free radicals which can cause breaks in DNA.

e. Recombinational repair – In contrast to the DNA repair pathways where damage at a specific nucleotide position is interrogated, recombinational repair involves large scale rearrangements of DNA across several kilobase lengths allowing bypass of the site of damage and restart of replication at a downstream position (Cox, 1999). This repair process is crucial for addressing double strand breaks and lesions generating single strand DNA gaps.

f. Translesional synthesis- While the repair processes previously described are non-mutagenic and restore DNA to its nascent state, the presence of extensive DNA damage can invoke translesion DNA synthesis activity wherein a low fidelity DNA polymerase (PolV in *E.coli*) inserts a base opposite to a lesion containing site thus leading to rescue of stalled replication (Goodman, 2000).

1.3 Structure and function of RecA

1.3.1 Role of RecA during DNA repair

DNA damage is a catastrophic event that is frequently encountered during replication

of DNA, and is repaired via the homologous recombination pathway (Cox et al., 2000; Kowalczykowski, 2000; Kowalczykowski et al., 1994; Spies and Kowalczykowski, 2005). During the initial stages of homologous recombination, one strand of a blunt duplex end is processed to generate a long stretch of single stranded DNA (ssDNA). Then, in a process known as strand exchange, the ssDNA finds a homologous double stranded DNA (dsDNA) partner in the cell and exchanges complementary base pairs to form a new heteroduplex product. Strand exchange reaction is catalyzed by RecA in *E.coli*. Homologs of RecA such as Rad51 and Dmc1 in eukaryotes and RadA in archaea perform similar functions during DNA repair via homologous recombination (Bianco et al., 1998). Strand exchange is followed by the formation of Holliday junction intermediates (Potter and Dressler, 1976) which are eventually resolved by branch migration proteins to complete homologous recombination. In addition to the aforementioned functions during strand exchange, the activation of RecA by the cell upon sensing a double strand break underpins the mechanisms leading to initiation of the SOS response pathway (Sutton et al., 2000).

1.3.2 SOS response – a regulated response to DNA damage

In bacteria, different repair processes can be integrated into a single pathway which directs the response of the cell to DNA damage. This pathway is called the SOS response pathway and its activation signals a high activity state of DNA damage and repair within the cell (Figure 1.2). It is known that the presence of extensive DNA damage turns on the SOS response pathway leading to the induction of over 40 different genes (Sutton et al., 2000). The SOS response results in the induction of genes that are involved in different aspects of DNA repair such as nucleotide excision repair (*uvrA, uvrB, uvrD*), homologous recombination (*recA and recA loading proteins)* cell division inhibitors which stall cytokinesis (*sfiA*) and a mutagenic translesional DNA polymerase PolV (*umuC* and *umuD*). Interestingly, the expression of these genes displays temporal regulation with nucleotide excision repair genes being expressed first and translesion polymerases being induced towards the end of the SOS response nearly 40 minutes after the initial damage has occurred (Michel, 2005). An important aspect of RecA function is its ability to trigger a dramatic transformation of cellular activity in response to DNA damage. RecA participates in the initiation of the SOS response by controlling the activity of a second protein

called LexA. LexA is a transcriptional regulator inhibiting the expression of the genes expressed during the SOS pathway (including recA). The interaction of LexA with an activated RecA filament (RecA bound to ssDNA) initiates the autocatalytic cleavage of LexA leading to its inactivation and the subsequent expression of genes involved in the SOS response. The change in transcription and protein levels is dramatic leading to a 50 fold change in RecA expression levels from a basal level of ~1200 molecules/ cell due to the loss of LexA repression (Sutton et al., 2000). Hence, loss of RecA function in bacterial cells leads to the inability of cells to mount an SOS response in addition to them being defective in homologous recombination activity.

1.3.3 RecA mediated strand exchange reaction

The coordination between different proteins leading to the assembly of the RecA filament sets the stage for a reaction involving the rearrangement of DNA strands that lies at the heart of the homologous recombination pathway (Wyman, 2011). In prokaryotes, the homologous recombination pathway is important for replication and repair of DNA but in eukaryotes it also serves an additional function of creating genetic diversity during meiosis. Chapter 3 will provide a more detailed description of RecA function during strand exchange.

The events which lead to the loading of RecA protein on ssDNA and subsequent strand exchange may occur via two different pathways (Spies and Kowalczykowski, 2005) 1) RecBCD pathway where the encounter of a double strand break leads to unwinding and resection of duplex DNA to generate a ssDNA substrate(Figure 1.3A) 2) RecF pathway involving RecFOR proteins which are involved in the repair of ssDNA gaps generated during the replication of lesion containing DNA and consist of proteins that facilitate RecA loading on ssDNA (Figure 1.3B)

The first step of the strand exchange reaction involves the binding of RecA proteins on ssDNA during a step referred to as pre-synapsis. In the next step called synapsis, the active RecA filament mediates the search for a homologous sequence and catalyzes basepair exchange to form a new heteroduplex product. Interestingly, both pre-synapsis and synapsis require ATP as a co-factor but ATP hydrolysis itself is not mandatory for reaction progression. RecA dissociation from ssDNA and the final heteroduplex requires ATP hydrolysis though it is not well understood how filament disassembly facilitates strand exchange (Cox, 2007).

1.3.4 Mechanism of RecA filament formation

Pre-synapsis involves the assembly of a RecA filament along ssDNA by a nucleation event where 4-5 RecA monomers bind simultaneously (Galletto et al., 2006; Joo et al., 2006) followed by rapid extension of the RecA filament via monomer addition (Joo et al., 2006). The kinetics of RecA filament assembly is asymmetric with one end growing faster (3' end) than the other (5' end) and the molecular basis for such an asymmetry arises from differences in association rates at the two ends (Joo et al., 2006) (Figure 1.4A). Given that nucleation acts as a critical rate limiting step in the assembly of the RecA filament it also potentially serves as a point of regulation. Several accessory proteins such as RecF, RecO and RecR have been hypothesized to play an important role in acting as a nucleation center for RecA filament assembly *in vivo*. RecA filament assembly also encounters other kinetic barriers such as the presence of secondary structure in ssDNA emphasizing a role for proteins such as SSB which remove secondary structure and facilitate filament extension (Roy et al., 2009) (Figure 1.4B).

1.3.5 RecA filament structure

RecA forms a filament which stretches the ssDNA to a length of 1.5 times the length of B-form DNA (Dunn et al., 1982; Stasiak et al., 1981; Stasiak and Egelman, 1986). The RecA filament consists of 6.2 monomers per helical turn with a binding stoichiometry of three nucleotides per RecA monomer (Figure 1.5B). The average rise between nucleotides in the filament is 5.1 Å. The first crystal structure of a RecA filament in 1992 consisted of RecA monomers in a helical arrangement without the bound ssDNA (Story and Steitz, 1992; Story et al., 1992). The filament structure was consistent with the collapsed state of the RecA filament (ADP bound state) and provided atomic resolution details about several aspects of RecA filament structure.

In 2008, Pavletich and his colleagues obtained a crystal structure of the RecA filament with ssDNA in its central axis (Chen et al., 2008) (Figure 1.5A). This structure dramatically altered our view of RecA interactions with ssDNA. Most importantly, the stretching of DNA by the RecA filament was found to be asymmetric (Figure 1.5C). The filament holds three successive nucleotides of ssDNA in B-form configuration with an average separation of 3.4 Å. The RecA mediated stretching occurs between successive nucleotide triplets with a rise of ~ 7.1 Å between them.

The crystal structure also revealed an allosteric interaction between the protein, ATP and the single strand DNA thus clarifying the mandatory requirement of a co-factor for the formation of functional RecA filaments. Since DNA is locally held in B- form configuration, it is tempting to propose a mechanism in which both homology recognition and basepair propagation might involve pairing of successive triplets.

A second intriguing aspect is that the crystal structure with RecA bound to double strand DNA shows the second strand held in place predominantly by base pairing interactions with the primary ssDNA. The second DNA strand makes only few contacts with the RecA protein lending credence to the hypothesis that RecA is more of a scaffolding protein rather than being actively involved in discriminating between complementary and non-complementary base interactions.

1.4 Overview of the thesis

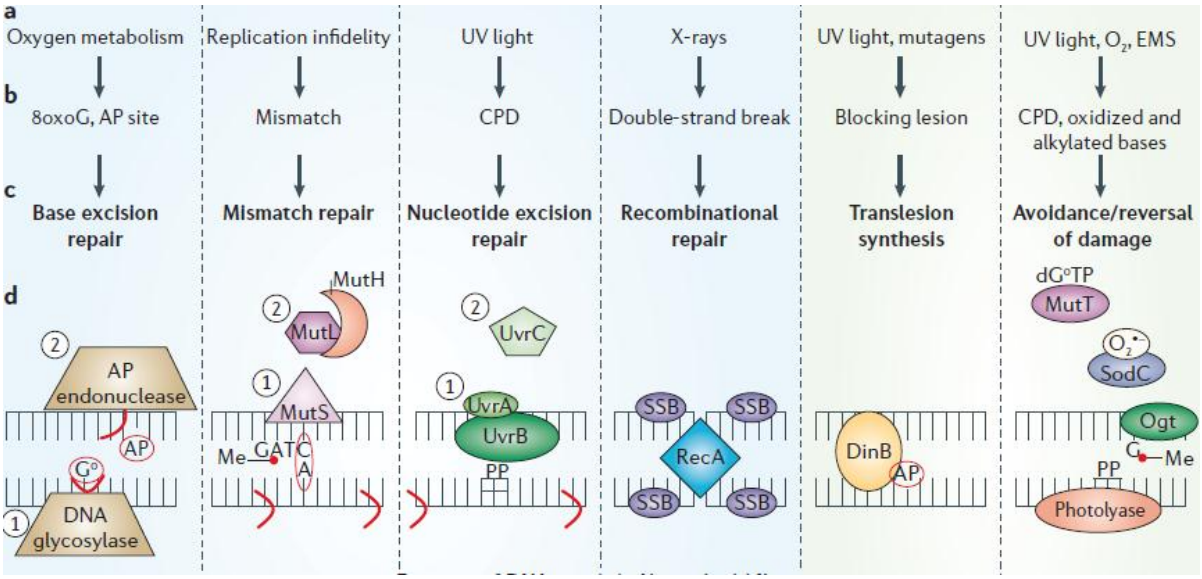
My thesis will focus on the study of the homologous recombination pathway with special emphasis on a critical step during homologous recombination, namely, the *strand exchange reaction* catalyzed by RecA. More broadly, the thesis consists of two parts: 1) Kinetics of basepair exchange and heteroduplex formation 2) Homology search mechanism.

In *Chapter 1,* I have given a broad overview of DNA repair processes in *E.coli* and more specifically on RecA. This chapter provides a biological context for the role of RecA during homologous recombination and other cellular pathways in *E.coli* with some insight into structural and biochemical studies on RecA. *Chapter 2* focusses on the methods I use to study the RecA mediated strand exchange reaction and also information regarding new methods which will prove useful for future single molecule fluorescence based studies. *Chapter 3* contains results regarding the interaction between RecA-ssDNA and a homologous dsDNA. *Chapter 4* contains preliminary observations of interactions between different DNA repair proteins and how single molecule FRET can address the mechanism underlying such interactions. *Chapter 5* introduces the second part of my thesis where I focus on the homology search process mediated by RecA in which I establish a role for sliding of dsDNA relative to the RecA filament leading into *Chapter 6* where I establish that basepairing can occur during sliding providing a physiologically relevant context for the sliding activity. *Chapter 7* introduces the prospect of future studies on the strand exchange

reaction using four color FRET to address some of the drawbacks of a conventional two color FRET approach to study the strand exchange reaction. The appendices at the end of the thesis provide supporting information on the experimental methods and analyses used.

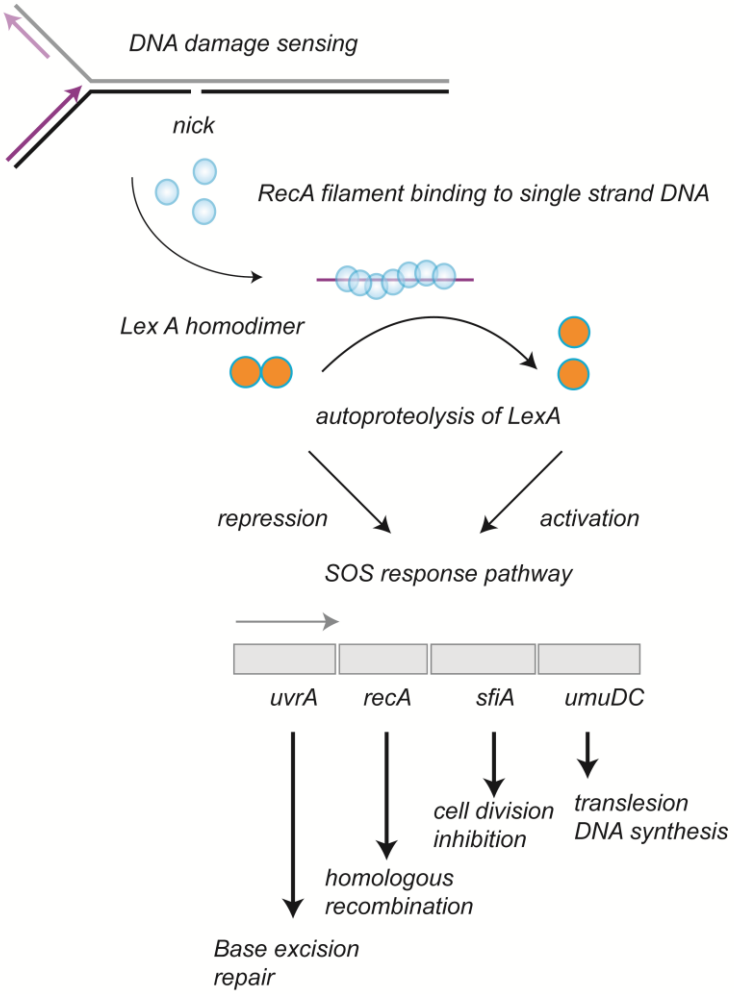
1.5 Figures

Figure 1.1 DNA repair pathways in *E.coli*



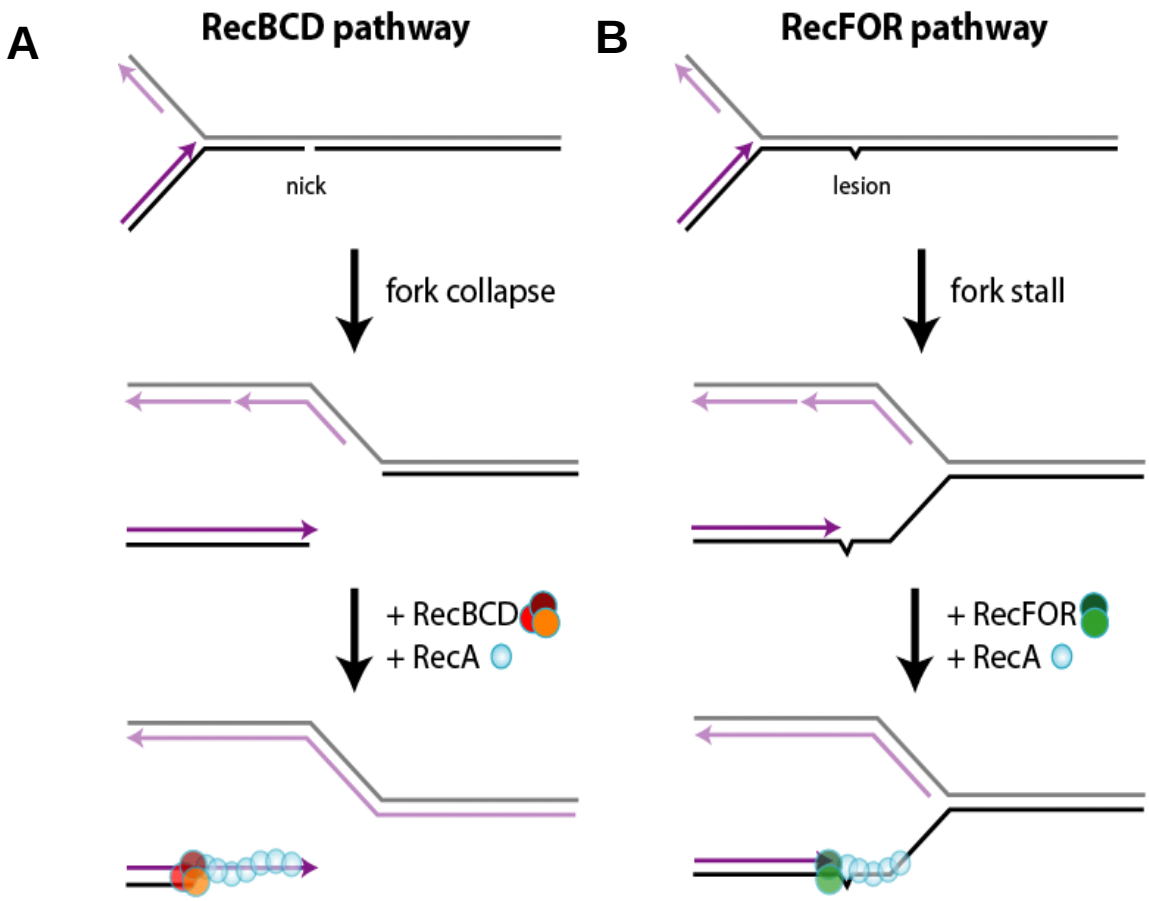
Major DNA-repair pathways in *Escherichia coli* **a.** DNA damaging agent **b.** Result of DNA damaging **c.** Repair pathway **d.** Proteins involved in a specific repair pathway

Figure 1.2 SOS pathway responds to DNA damage



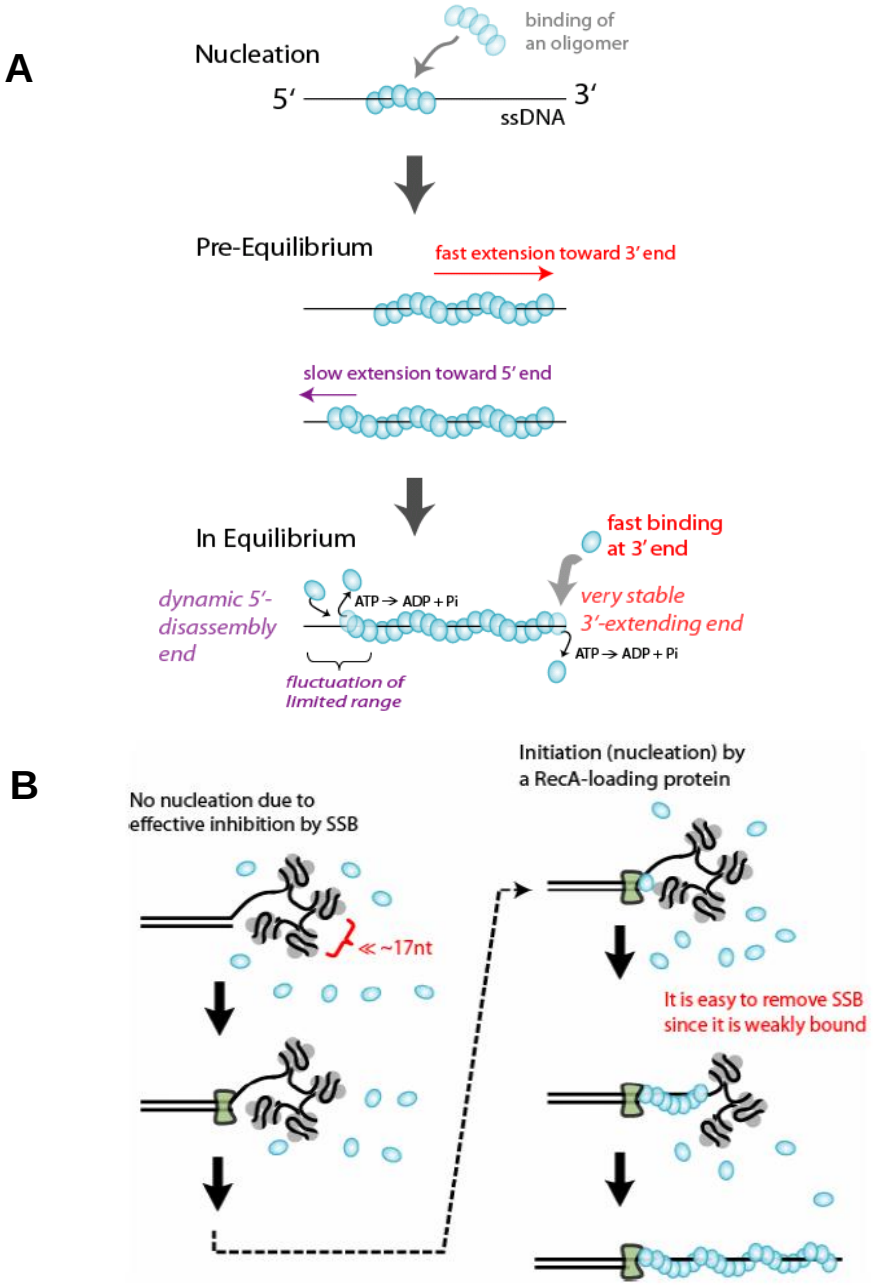
The initiation of the SOS response pathway involves the RecA protein binding form a filament on ssDNA upon encountering DNA damage. This is followed by the autocatalytic cleavage of LexA which then results in the expression of several genes that play critical roles in DNA repair, replication and mutagenesis.

Figure 1.3 DNA repair via homologous recombination pathways in *E.coli*



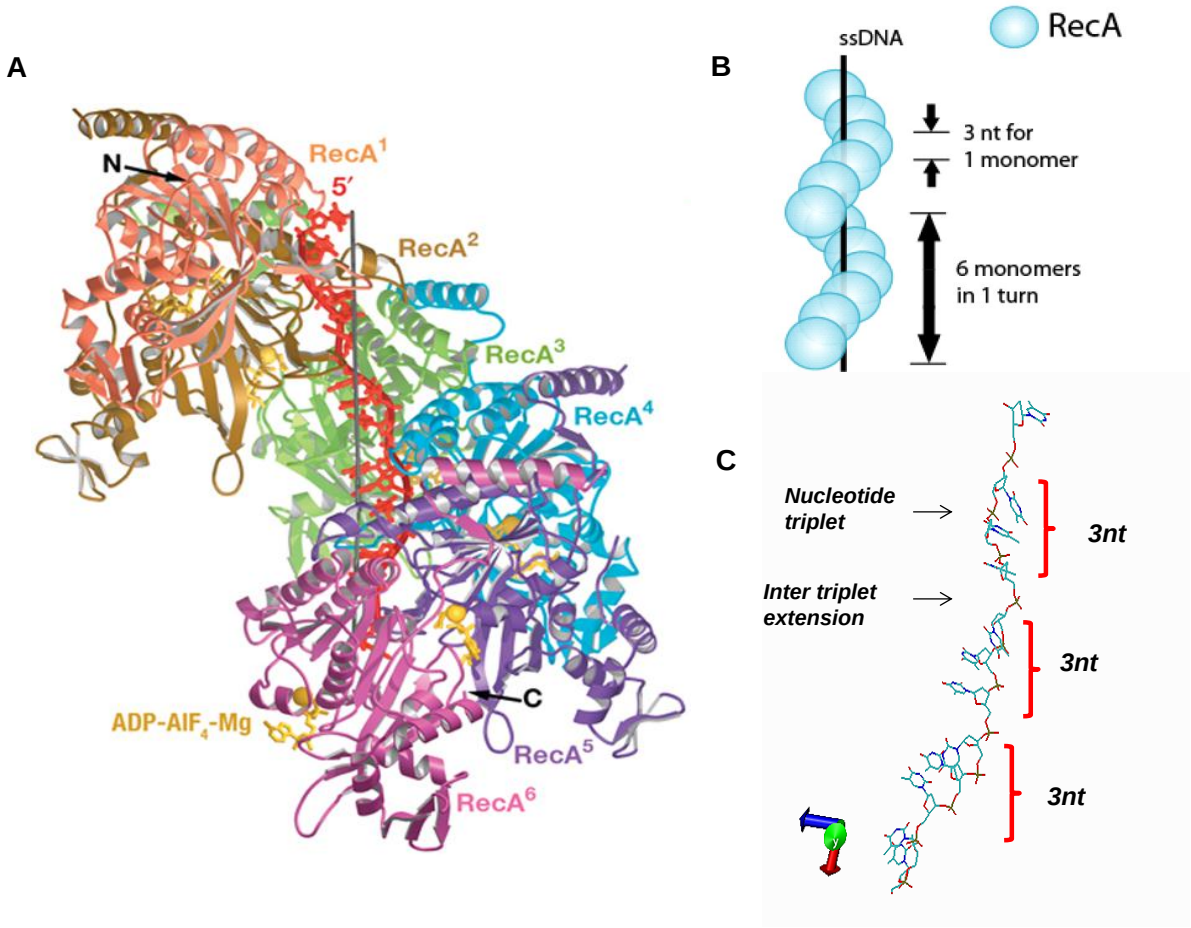
DNA repair via the homologous recombination pathway **A.** RecBCD pathway primarily involves the participation of the multi enzyme RecBCD complex which facilitates end resection of DNA containing double strand breaks **B.** RecFOR pathway primarily involves proteins RecFOR which facilitate the nucleation of RecA on ssDNA gaps encountered during the replication of lesion containing DNA

Figure 1.4 RecA filament dynamics and filament assembly



A. Dynamics of filament assembly where 5’ end is grows more slowly than the 3’ end due to an asymmetry in association rates of RecA monomers **B.** Filament assembly involves proteins that facilitate nucleation (such as RecFOR) and proteins that enable filament growth (such as SSB)

Figure 1.5 Description of RecA filament structure



A. High resolution structure of RecA-ssDNA complex (Chen et al., 2008) **B.** RecA binding stoichiometry and physical properties **C.** Structure of ssDNA embedded within a RecA filament displays triplets of nucleotides in B-form conformation and the RecA mediated extension between adjacent triplets

Chapter 2

Single molecule fluorescence methods[1]

Groups are capable of being (only) as moral and
intelligent as the individuals who form them
adapted from Aldous Huxley (1958)

2.1 Introduction to single molecule methods

Real time approaches that aid in the detection, visualization and manipulation of individual molecules have provided new perspectives on biological processes. In contrast to ensemble assays where the read out from the experiment is the average behavior of many molecules, single molecule methods allow for: 1) Direct detection of sub populations and dynamics 2) Detection of non-synchronizable dynamics 3) Detection of molecular heterogeneities 4) Correlation between multiple observables.

Single molecule fluorescence based methods are widely used in many different contexts. Measurements involving localization of a biomolecule tagged with a fluorophore can enable detection of protein movements with nanometer precision. In a different implementation, a pair of dyes can be used to measure distance changes on a short length scale (1-10nm) via FRET.

Using fluorescence to measure distance changes on a short distance changes over short resol the utilization of two fluorophores to meas to detect individual molecules and monitor their behavior.

2.2 Introduction to fluorescence

The excitation of fluorophore with an appropriate light source results in a sequence of events that leads to the emission of a photon (Figure 2.1). The absorption of a photon by a fluorophore present initially in the ground state (S_0) results in its transition to the excited state (S_1). Rapid internal conversion results in the molecule

[1] Parts of Chapter 2 have been published
• Lee, J., S. Lee, **K. Ragunathan**, C. Joo, T. Ha and S. Hohng, "Single-molecule four-color FRET", Angew. Chem. Int. Ed. 49(51), 9922-9925 (2010)
• Lee, J., S. Lee, **K. Ragunathan**, C. Joo, T. Ha and S. Hohng, "Single-molecule four-color FRET", Angew. Chem. Int. Ed. 49(51), 9922-9925 (2010)

occupying the lowest vibrational level of S_1. Eventually the molecule will return to the ground state. If this transition from the excited state to the ground state is accompanied by the emission of a photon, this process is referred to as fluorescence. The emitted photon is red shifted (Stoke's shift) relative to the excitation source. Other possible relaxation pathways exist such as the dissipation of energy as heat, collisional quenching or intersystem cross over (ISC). During ISC, the electron in the excited state must undergo a spin conversion resulting in its transition to a triplet state. The transition to the triplet state is of particular relevance to our single molecule studies since it results in intermittency in the fluorescence signal arising from a fluorophore. A molecule in the triplet state is also highly reactive and may undergo oxidation subsequently resulting in the complete loss of fluorescence capability.

2.2.1 Förster resonance energy transfer

While a molecule is in the excited state it has the capability to transfer energy in a non-radiative manner to a different molecule and subsequently return to its ground state. This phenomenon is called Förster resonance energy transfer (FRET) (Figure 2.2). The efficiency of energy transfer between two fluorophores displays a steep distance dependence which is described by the equation, $E= 1/[1+ (R/R_0)^6]$, where R_0 denotes the characteristic length for which the efficiency of energy transfer is 50% (Figure 2.3). The value of R_0 depends on the dye pair we use. FRET can thus serve as a spectroscopic ruler providing a way to measure distance changes on the scale of 30-80Å. An important requirement for two molecules to serve as a FRET pair is that the emission spectrum of one molecule (donor) overlaps with the excitation of a second molecule (acceptor). In this thesis, FRET with two or more than two colors will be the strategy used to measure dynamics of individual molecules.

2.3 Observing single molecules

The detection of single molecules requires solving three critical issues: 1) the use of a detector with high sensitivity to eliminate background arising from thermal noise. 2) an excitation method which attenuates background 3) Spatial separation of individual molecules. Confocal imaging is one of the methods used to visualize individual molecules and it involves the use of a focused laser to excite molecules within a diffraction limited spot. Detection of fluorescence emission is achieved by using an

avalanche photo diode (APD) and excluding background fluorescence using a pinhole. Confocal imaging allows for high time resolution (~100μs) but compromises on the ability to observe more than one molecule at a time. Alternatively we could use a wide field excitation approach such as total internal reflection where focusing light on the imaging surface at an angle beyond the critical angle results in an exponentially propagating excitation field allowing us to detect molecules that are close to the surface (~200nm). This method allows the simultaneous observation of many molecules over a 75 μm x 75 μm area but offers lower time resolution (typically 8-30ms) compared to confocal imaging.

2.3.1 Two color FRET instrumentation

Given our interest in measuring FRET between two fluorophores, it is imperative that we setup a detection system to measure the emission intensities of an individual donor and acceptor fluorophore (Figure 2.4). The relative ratio of the intensities of the two fluorophores provides a read out of the efficiency of energy transfer between them with a change in this ratio resulting from changes in distance between the two fluorophores barring contributions from other photophysical factors (also see Appendix C for details about FRET calculation). We typically immobilize the single molecules on a passivated quartz slide which serves as the imaging surface for measuring single molecules (see also Appendix A). The slide is mounted on an inverted microscope (Olympus, IX71) with a custom made sample holder stage. A pellin broca prism is placed on top of the quartz surface and the incoming excitation light entering the prism undergoes a change in angle resulting in total internal reflection at the interface between the quartz slide and the aqueous imaging solution. Three laser lines are available in the setup – 488nm, 532nm and 633nm which excite a range of commonly used fluorophores. The emission resulting from the molecules is passed through a long pass filter that attenuates the excitation light and only the fluorescence emission from the immobilized single molecules (typically Cy3 and Cy5) is refocused to the camera. Prior to refocusing the emission, the light is passed through a dichroic mirror which splits the total emission into donor and acceptor channels. We then map the molecules in the donor channel to the corresponding positions in the acceptor channel and record their intensities as a function of time. It is this ratio of donor and acceptor intensities which we report as FRET. The choice of dichroics depends on the colors being measured. For Cy3 and Cy5 separation of

emission is achieved using a dichroic with a cutoff at 630nm which causes a leakage of 15% donor signal to the acceptor channel. It is common practice to correct for such contributions from leakage during post processing of single molecule traces.

2.3.2 Expanding the single molecule color palette

The choice of fluorophores is subject to different experimental needs. I will specifically discuss the context in which we would like to obtain multiple distances simultaneously from a single experiment. The ability of Cy5 blinking has been exploited to enable measurement of multiple distances from a sample labeled with more than one Cy5 fluorophore (Uphoff et al., 2010). However given that blinking is stochastic, it becomes challenging to estimate the changes in conformation when the molecule when is switching between different conformations and we want to discover correlations between these changes in conformation.

The ability to measure multiple distances requires expanding the palette of fluorescent dyes beyond what is typically used for single molecule experiments. Figure 2.5 lists the emission spectra for some of the dyes typically used for single molecule imaging. Dyes which are red shifted to Cy5 (emission max.=667nm) such as Cy5.5 (emission max.= 695nm) have been previously employed to design a three color FRET measurement where distance changes between a single donor (Cy3) and two acceptors (Cy5 and Cy5.5) have been monitored (Hohng et al., 2004; Roy et al., 2009). However, such a configuration is limited by the ability to design experiments wherein the distance between Cy5 and Cy5.5 is large and the two dyes themselves exhibit no significant FRET interaction. This is because of the significant spectral overlap between Cy5 and Cy5.5 which does not permit for independent excitation of the different dye pairs. We need to use dyes which are spectrally separated from each other such that we can excite each dye pair using a unique excitation. The use of Cy7 (emission max.= 785nm) as an alternative dye has been explored in three color configuration (Lee et al., 2010b). The advantage of using Cy7 is its large spectral separation relative to Cy5 though the disadvantage is its poor detection efficiency using an EM-CCD and also transmission losses arising from optical components that are not ideally suited for dyes with far infra-red emission (Roy et al., 2008). On the other end of the spectrum, the possibility of using Cy2 and Alexa 488 (emission max= 519nm) which are blue shifted relative to Cy3 has also

been explored. Cy2 is commercially available only as a bis-NHS ester which poses a challenge in terms of single labeling of biomolecules and this is something which one should be careful in using this dye for single molecule experiments. On the other hand its photophysical characteristics are superior to Alexa 488 which suffers from low brightness due to its poor extinction coefficient and also exhibits fast photobleaching compared to Cy3 or even Cy5. It is important to note that the different dyes also have different overlap in emission and excitation spectra resulting in different R_0 values allowing for a different range of sensitivity depending on the dye pair used (Figure 2.6).

Given that tagging proteins with a fluorophore at a unique position remains a challenge, another option is the use fluorescent proteins such as GFP, YFP, mCherry etc. The advantage of fluorescent proteins is the relative ease with which the fusions with proteins of interest can be made and the fact that the purified protein is obtained in the labeled form. This principle has been utilized to develop approaches wherein the cell can be lysed and individual proteins can be readily visualized by recruiting them to the single molecule imaging surface by using immobilized antibodies (Jain et al., 2011; Yeom et al., 2011). However, the poor photophysical features of fluorescent proteins such as frequent blinking and fast photobleaching pose a challenge to the measurements of single fluorescent proteins. Despite their disadvantages, fluorescent protein fusions have given single molecule researchers an opportunity to study biomolecules in their native states of association and interaction with other proteins rather than relying only on the use of purified components.

2.3.3 Multicolor single molecule FRET and instrumentation

The following multicolor FRET scheme and instrumentation was developed in Sungchul Hohng's lab at Seoul National University and I implemented and designed an identical setup in the Ha laboratory. Given that we can now reliably estimate FRET between different dye pairs, we can simultaneously measure dynamics between and within different sections of a molecule and also measure how the dynamics may be correlated with each other (Lee et al., 2010a).

The instrumentation for a multicolor (three or four color) is virtually identical to the single molecule two color setup described previously. However, instead of a single

dichroic which separates the emission of the dyes into two channels for two color FRET implementation, we use two additional dichroics (bringing the total number to three dichroics) to create four different emission channels to simultaneously measure, for example, Alexa488, Cy3, Cy5 and Cy7 (Figure 2.7). In addition to a change in the optical configuration, there is also a change in the excitation configuration. There are a total of six inter-dye distances to be determined in four-color FRET. To recover all the information necessary to calculate the six inter-dye distances, we developed an approach that uses three independent excitation lasers: a red laser (633-nm or 640-nm) for Cy5, a green laser (532-nm) for Cy3, and a blue laser (473-nm) for Cy2 or Alexa488 excitations. The FRET efficiency of Cy5-Cy7 pair (E_{34}) is determined via red excitation. Next, FRET efficiencies of Cy3-Cy5 (E_{23}) and Cy3-Cy7 (E_{24}) are determined at green excitation on the basis of E_{34}. Finally, the remaining three FRET efficiencies (E_{12}, E_{13}, and E_{14}) are determined using blue excitation (Figure 2.8). The sequential determination of FRET efficiencies in real time was achieved by alternating three excitation lasers on a faster time scale than the conformational dynamics of molecules while synchronizing the detection of fluorescence signals with laser switching.

The current implementation in the Ha lab allows for alternating between two laser excitations allowing for the determination of three possible FRET efficienices in a three color format. However, with a careful experimental design one can still achieve four color detection in a scheme which is referred to as ‚dual FRET‘ approach (Lee et al., 2010a).

2.3.4 Implementation of alternating laser excitation (ALEX)

An alternating excitation scheme must acheive synchronization between the excitation source and the detector. The EM-CCD camera provides an output TTL pulse sequence which provides a readout of completion of data acquisition for each successive frame. We use this output as a readout to synchronize the shutters (Figure 2.9). We achieve the alternation of lasers by feeding the output TTL pulse sequence from the camera to a flip flop circuit. The logic of the flip flop circuit allows for alternating between two output 5V signals which can be fed to the shutter controller. Integration of the flip flop circuit with a multiplexer allows the user to specify whether the shutters are controlled using the computer or the flip flop circuit.

While this excitaiton and detection scheme achieves a high precision of synchronization between the detector and the excitation sequence, there is no flexibility to change the duration between successive frames and hence the utility of such a device is narrow. The maximum speed with which one can alternate shutters is limited to and fixed by the time scale of data acquisition. The highest time resolution we typically use is 30ms which is much slower compared to shutter opening time which can be achieved on the μs time scale.

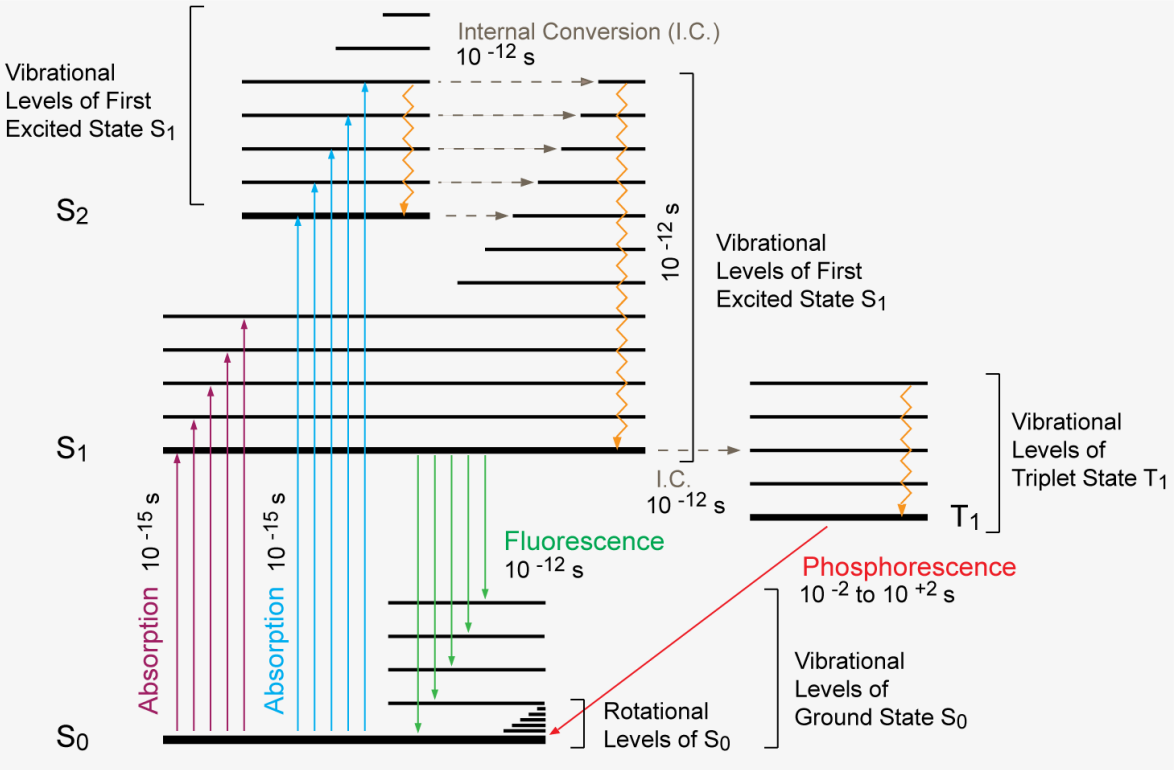
2.3.5 Perspectives and outlook on multicolor FRET

Single-molecule FRET (Förster Resonance Energy Transfer) has provided unprecedented details on fundamental processes in biology. However, information regarding single inter-fluorophore distances in conventional two-color FRET is insufficient to completely capture the intrinsic complexity of many biological systems. To cope with this challenge, single-molecule multi-color FRET techniques have been developed, and the unique capability of three-color detection has been utilized to solve a number of important biological problems. However, as single-molecule approaches are being expanded to include more complex biological systems, there is an ever increasing demand for more advanced FRET techniques.

Currently the use of multi color approaches has been indispensable for confirming sliding dynamics in several biological systems such as RecA (as described in Chapter 5), SSB sliding on ssDNA (Roy et al., 2009), Tar binding protein (TRBP) sliding on dsRNA substrate (unpublished, in collaboration with Hye Ran Koh) and also to study the correlations in dynamics of protein and RNA during ribosome assembly (unpublished, in collaboration with Hajin Kim and Sarah Woodson).

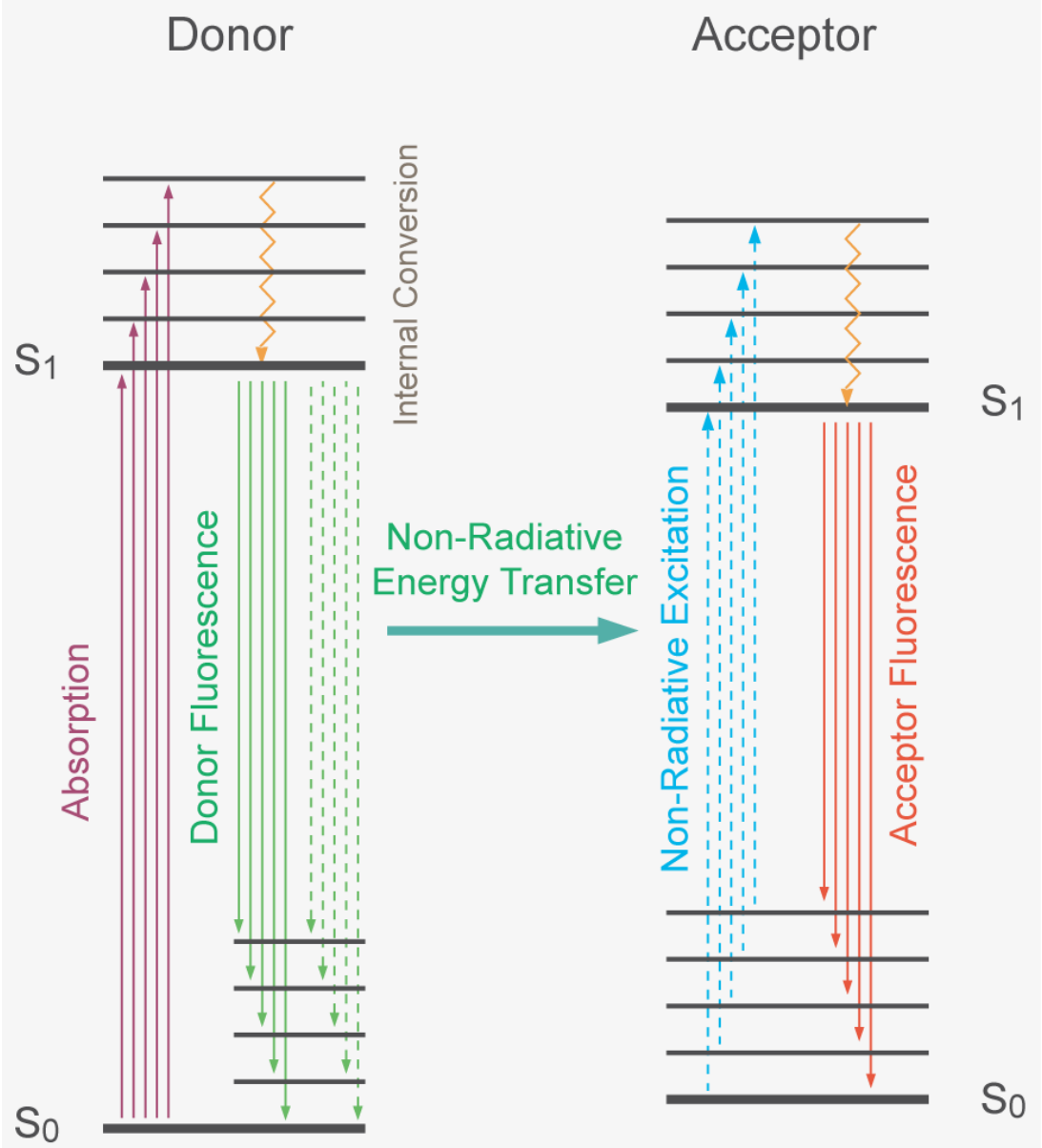
2.4 Figures

Figure 2.1 Jablonski diagram indicating the consequences of molecular exctiation



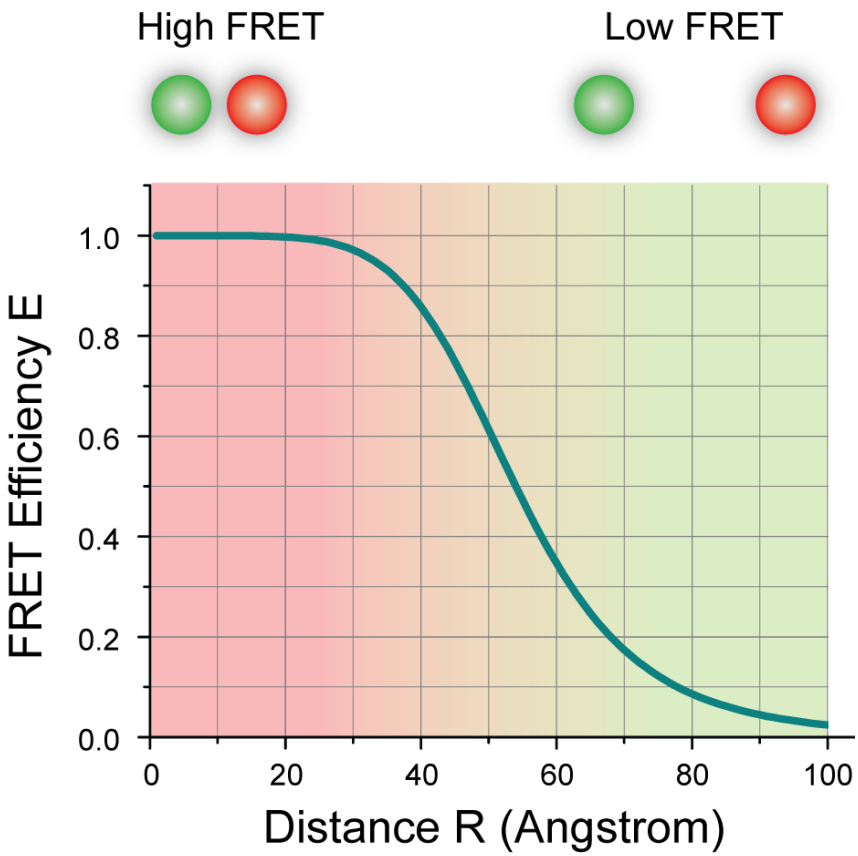
Excitaiton of a molecule from the ground state (S_0) leads to a transition the excited state (S_1) following which its decay to the ground state can occur via multiple pathways.

Figure 2.2 Forster resonance energy transfer involves a non-radiative interaction between a donor and acceptor fluorophore



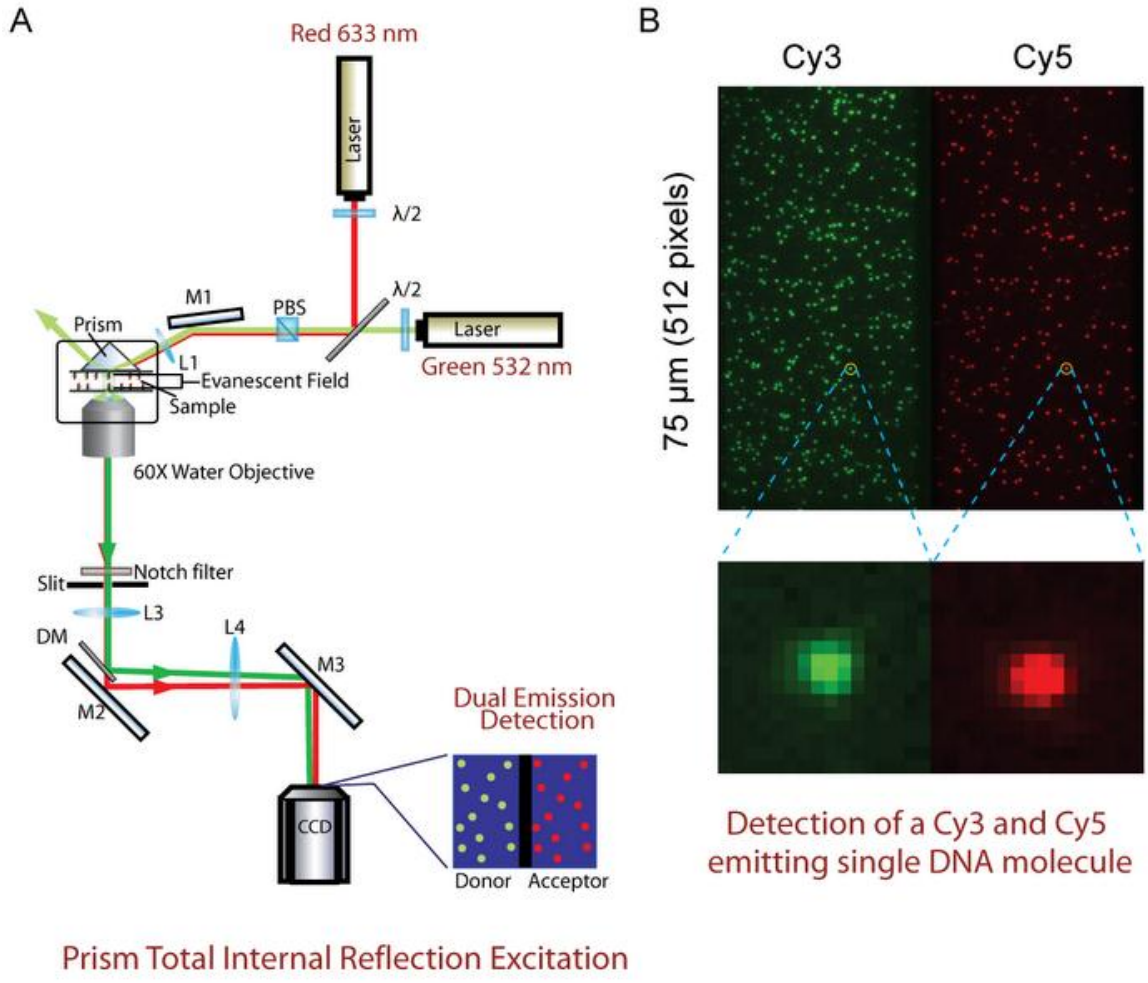
Excitation of a donor leads to a transition to the excited state (S_1) from which non-radiative energy transfer can occur leading to subsequent excitation of the acceptor.

Figure 2.3 FRET can be used as a spectroscopic molecular ruler



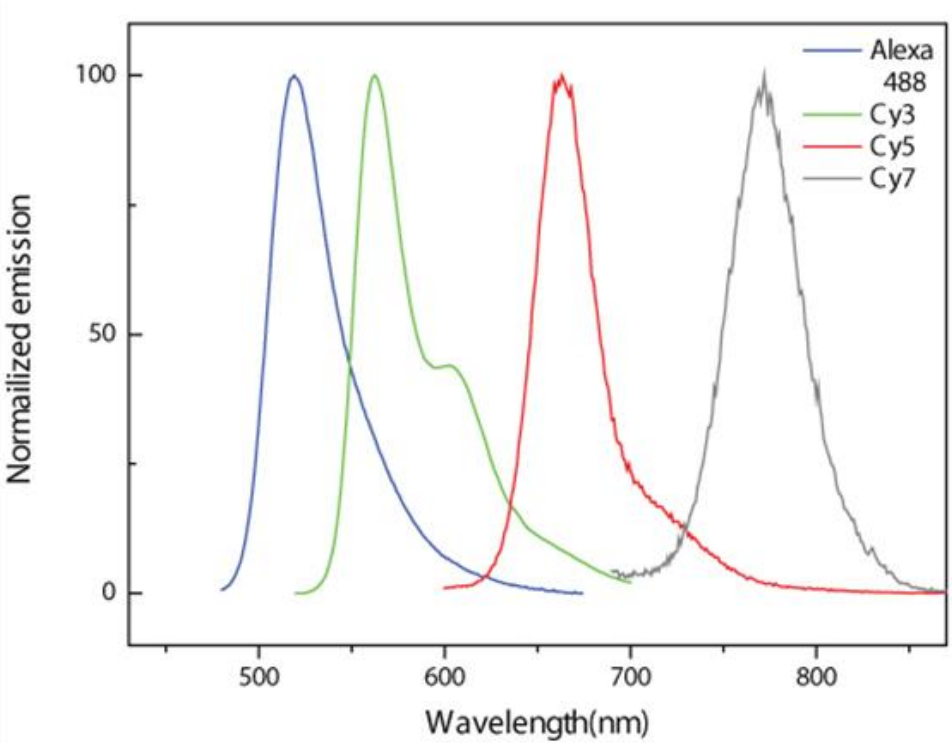
Efficiency of energy transfer between a typical donor acceptor pair (Cy3 and Cy5) displays a steep distance dependence with R_0 ~ 54Å.

Figure 2.4 Single molecule two color FRET detection



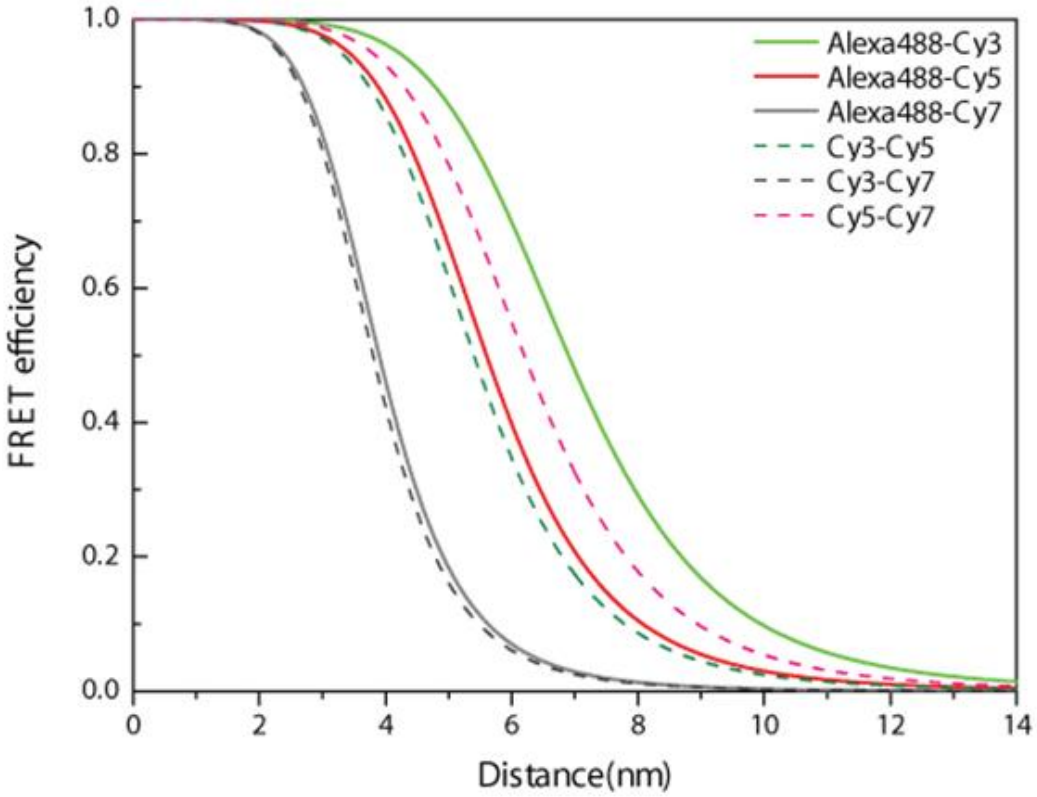
A. Excitation and emission scheme where molecules are excited via total internal reflection. The resulting emission is refocussed on to a CCD after splitting the emission between a donor and acceptor channel **B.** An example of the image acquired using th CCD showing donor and acceptor emission. A mapping function allows for determining the spatial locations corresponding to the donor and acceptor emissions arising from individual molecules or complexes

Figure 2.5 Excitation and emission spectra for typically used fluorophores



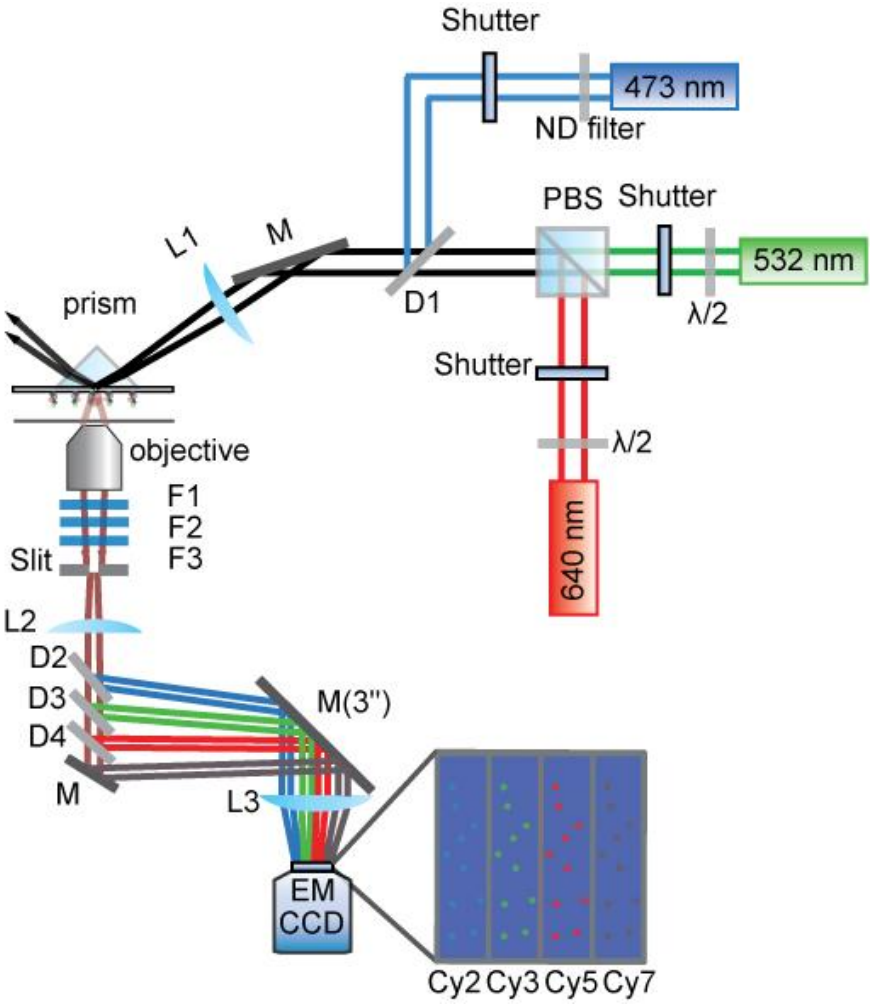
Emission spectra for a different fluorophores typically used for single molecule imaging. The key point is to note the spectral separation between the different fluorophores which permits independent excitatation and detection.

Figure 2.6 Predicted FRET efficiency curves for typically used fluorophores



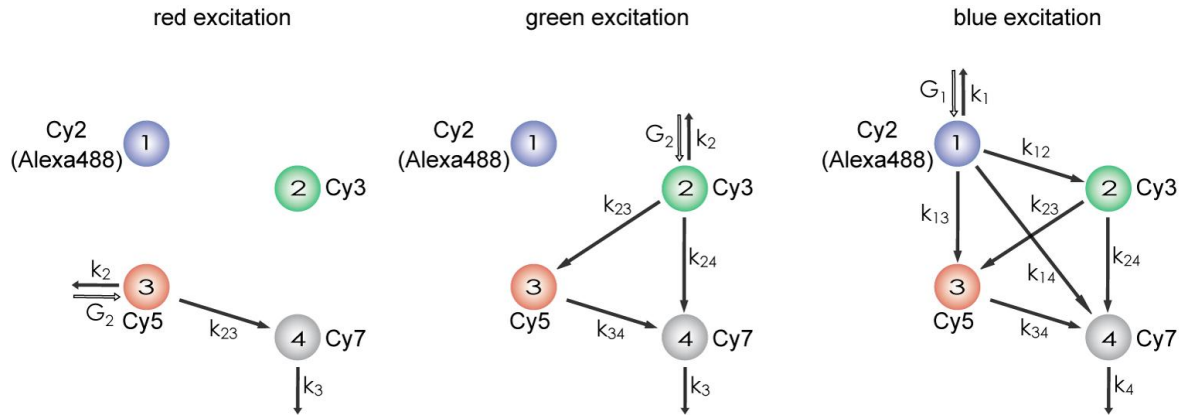
The energy transfer efficiency between two flurophores is a function of different factors that includes, distance between dyes, relative orientation of the dyes, quantum yield, overlap between the donor emission and acceptor excitation spectra.

Figure 2.7 Single molecule multi color FRET detection



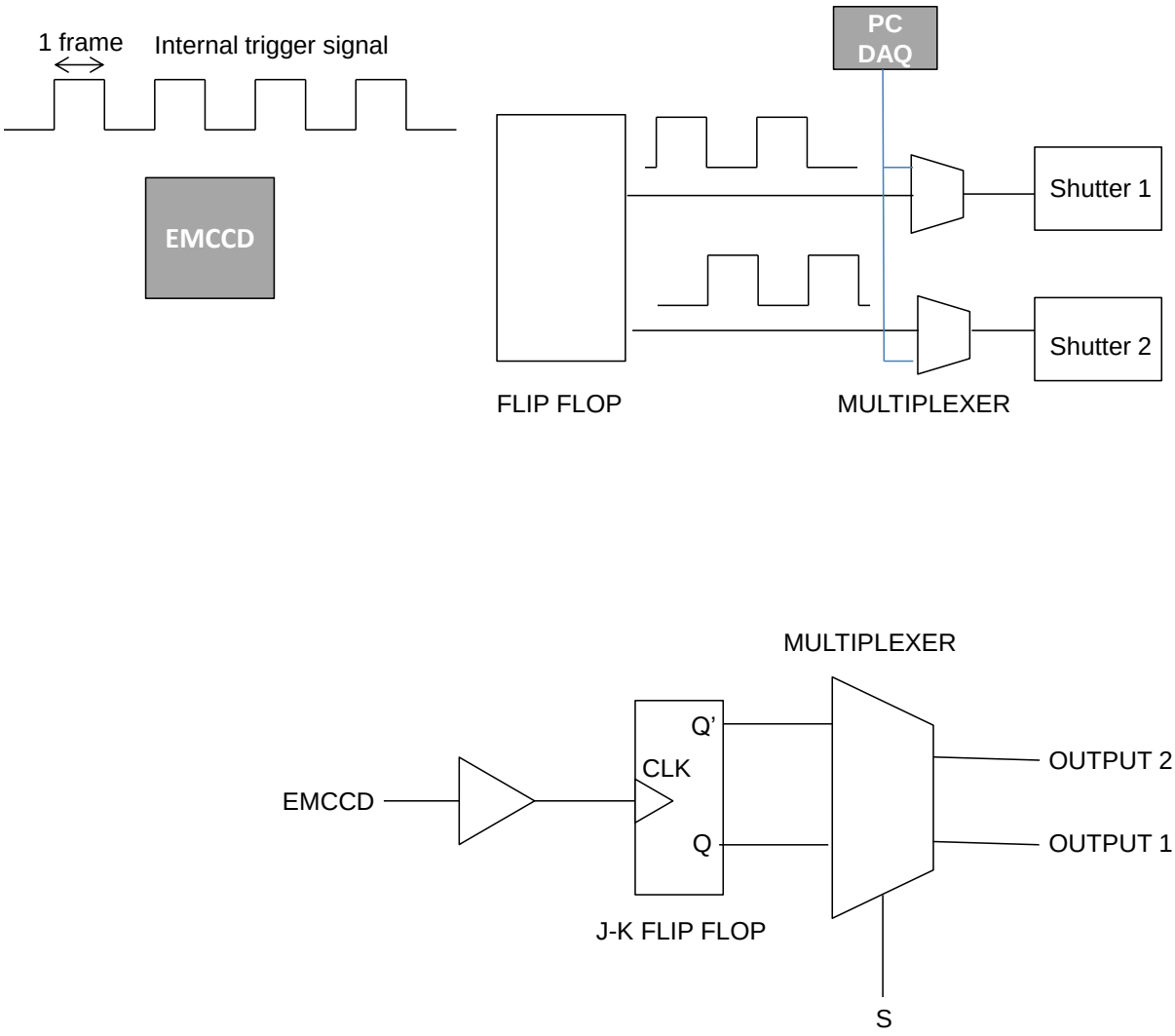
Excitation and emission scheme to detect more than two colors. In this scheme, we can detect four different colors and simultaneous detection of six possible FRET efficiencies.

Figure 2.8 Possible FRET interactions in a four color scheme



Excitation of different colors gives rise to different FRET interactions. Alternating between different excitation lines allows us to determine the distances between all fluorophores.

Figure 2.9 Design of a circuit for two color alternating laser excitation (ALEX)



Design of an electrical circuit to synchronize detection via an EM-CCD with alternating laser excitation

Chapter 3

Real time observation of strand exchange[2]

I had felt that recombination was far too
complex and abstruse a subject for me to master
Alvin Clark (1996)

3.1 Introduction to the strand exchange reaction

The isolation of recombination-deficient mutants of *E.coli* paved the way for the identification of RecA (Clark and Margulies, 1965) which catalyzes the strand exchange reaction- a key step in the homologous recombination pathway. Strand exchange involves pre-synapsis, synapsis and heteroduplex extension via branch migration. Presynapsis involves the assembly of RecA monomers on a ssDNA. Synapsis results in the interaction between a RecA bound ssDNA with a target dsDNA at a homologous site leading to the reciprocal exchange of basepairs (Figure 3.1).

We introduce the terms of initiation, propagation and completion to describe the three steps that occur during synapsis. *(1) Initiation:* Search for homology by the RecA filament and homologous alignment between the 'incoming' dsDNA and the RecA bound 'reference' ssDNA *(2) Propagation:* Base pair exchange between the reference ssDNA and incoming dsDNA molecule to form a 'joint molecule'. Joint molecules represent a protein bound, three-stranded intermediate state during strand exchange wherein basepair exchange may not have proceeded to completion (Menetski et al., 1990) *(3) Completion:* Release of the 'outgoing' displaced ssDNA from the post-synaptic complex resulting in a RecA bound heteroduplex and free ssDNA. Joint molecule formation can be carried out without hydrolyzing ATP (Kowalczykowski and Krupp, 1995; Menetski et al., 1990) but the release of RecA from the final heteroduplex product requires ATP hydrolysis (Rosselli and Stasiak, 1990).

[2] The work in Chapter 3 has been published
• **K. Ragunathan**, C. Joo and T.Ha, "Real time observation of strand exchange reaction with high spatiotemporal resolution", Structure 19(8) 1064-73 (2011)

From a structural perspective, the RecA filament presents triplets of nucleotides in the B-form configuration (Chen et al., 2008) raising the intriguing possibility that basepairing exchange between two DNA strands may proceed via Watson-Crick pairing involving destabilization of the 'incoming' dsDNA in 3bp increments. However, until now, there has not been an experimental test of this prediction.

Following the exchange of base pairs between homologous DNA strands, the outgoing ssDNA is thought to remain bound to the RecA filament via weak interactions with the RecA secondary binding site (Mazin and Kowalczykowski, 1996, 1998). Biochemical studies demonstrated that the RecA secondary binding site serves as a gateway for strand exchange mediating the exit and the entry of DNA strands from the RecA filament (Kurumizaka et al., 1996). There is presently little information on the characteristics of DNA bound to the secondary binding site and no clear consensus exists on the structural and dynamic properties of the complex formed by the three DNA strands and the RecA filament during and after strand exchange (Camerini-Otero and Hsieh, 1993; Chiu et al., 1993; Folta-Stogniew et al., 2004; Jain et al., 1995; Podyminogin et al., 1995; Zhou and Adzuma, 1997). While fluorescence and FRET (Fluorescence Resonance Energy Transfer) based ensemble measurements have been valuable in establishing the presence of multiple kinetic intermediates during strand exchange, the number and identity of each of these intermediates remains ambiguous (Bazemore et al., 1997; Folta-Stogniew et al., 2004; Lee et al., 2006; Xiao and Singleton, 2002).

Here, we studied the mechanism of RecA-mediated joint molecule formation using single molecule FRET (Ha et al., 1996) as described previously in Chapter 2. Our single molecule FRET assay can separate the initial docking from the subsequent propagation leading to joint molecule formation thereby enabling us to analyze the strand exchange kinetics with clarity and precision. We found that the initiation of joint molecule formation involves a synaptic complex of <14 bp in length. Our data suggests that the propagation of base pairing leading to joint molecule formation occurs in 3 bp increments with destabilization of the incoming dsDNA and concomitant pairing with the reference ssDNA. Unexpectedly, we discovered the formation of a highly dynamic complex between RecA and the displaced outgoing ssDNA which remained bound for a few seconds after basepair exchange was completed.

3.2 Results

3.2.1 Single molecule fluorescence assay for strand exchange

A biotinylated dsDNA (18bp) with a free 5' ssDNA overhang and an acceptor fluorophore (Cy5) at the ssDNA-dsDNA junction was immobilized on a polymer-passivated surface (Figure 3.2). The single stranded portion of a specified homology length, L_h (nt), is bound by RecA to form a stable pre-synaptic filament using ATPγS as a co-factor. Using ATPγS allowed us to monitor synaptic events without turnover of RecA monomers from the DNA during or after reaction completion. We then flowed in a solution containing donor (Cy3)-labeled homologous dsDNA (also of homology length, L_h (bp)) and ATPγS, while simultaneously washing away free RecA in solution. This procedure ensures that the incoming dsDNA interacts solely with the immobilized filament. The labeling sites on the incoming dsDNA were chosen so that the donor and acceptor fluorophores are in close proximity after joint molecule formation. This 'docking-and-pairing assay' monitors docking of the incoming dsDNA to the RecA filament via fluorescence signal appearance and pairing via FRET change (Figure 3.2). The completion of joint molecule formation was confirmed by the appearance of a high FRET population with an apparent FRET efficiency E of ~0.85 (Figure 3.3A).

A control with non-homologous DNA produced only a low FRET population at E~0.1 (Figure 3.3A). In addition, the homologous dsDNA case showed a rapid accumulation of reaction products (Figure 3.3B) in contrast to the non-homologous control thus recapitulating the specificity of the RecA strand exchange reaction. The CCD image obtained using homologous and non-homologous dsDNA provides a visual cue regarding the specificity of the strand exchange reaction (Figure 3.4).

In order to verify that the final product formed in the presence of homologous dsDNA was the expected heteroduplex, we carried out deproteinization of the joint molecules and incubation with a restriction enzyme whose restriction site was located between the donor and acceptor dyes in the final product. Over 95% of the reaction product could be cleaved off (Figure 3.5A-C) resulting in the loss of donor signal and confirming that the end product is the expected heteroduplex.

3.2.2 Direct observation of initial pairing and strand exchange

Real-time single molecule time traces showed the docking of Cy3 labeled homologous dsDNA to the RecA filament as an abrupt appearance of fluorescence signal (Figure 3.6A). One class of molecules showed a low FRET value (E~0.1) at the moment of docking and later transitioned to the high FRET state (E~0.85) (Figure 3.4.6A (top panel)). This low to high FRET transition signals successful joint molecule formation near the labeled end of DNA. The other class of molecules showed the high FRET state from the moment of docking (Figure 3.6A (bottom panel)), indicating that joint molecule formation initiated near the labeled end. Several controls showed that the low FRET state is not due to photophysical effects of the fluorophores (Appendix C).

We defined the term, $<\Delta t_{delay}>$, to represent the mean dwell time of the low FRET delay period (Δt_{delay}) prior to the final high FRET transition. We exclude the molecules that directly display high FRET upon binding while calculating the $<\Delta t_{delay}>$. The low FRET period cannot be attributed to the homology search process because the mean dwell time, $<\Delta t_{delay}>$, among those events showing non-zero Δt_{delay}, did not change appreciably when the ssDNA length was varied while keeping the homology length, L_h at 39 nt (Figure 3.6B (inset)). Therefore, homology search after docking must be instantaneous within our time resolution (30 ms) so that the initial synaptic complex formation essentially coincides with docking.

It is notable that the Δt_{delay} distributions showed a clear shift to longer times with an increase in L_h (Figure 3.6B). The average of Δt_{delay}, $<\Delta t_{delay}>$, also showed a strong dependence on L_h (Figure 3.6C). Therefore, we attribute the initial low FRET period (Δt_{delay}) to the time it takes to propagate basepair exchange from the initial synaptic complex to the labeled end of DNA (see diagrams in Figure 3.6A). Consistent with this interpretation, the fraction of molecules which exhibit zero delay (Figure 3.6A (bottom panel)) decreased with increasing L_h (Figure 3.6C inset). This observation is likely due to more initiation sites being available thus decreasing the probability to initiate the reaction from the labeled end.

Furthermore, the histogram of the time delay (Δt_{delay}) exhibited by molecules showing the initial low FRET state displayed a non-exponential distribution which we will analyze in more detail in the following section.

3.2.3 Evidence for 3 nt step size of basepair exchange

Having established that Δt_{delay} represents the propagation of joint molecule formation, we further analyzed the distribution of Δt_{delay} to extract information regarding the number and identity of kinetic intermediates prior to reaction completion. If there was only one rate limiting step which needed to be overcome to complete the exchange of all the basepairs in the incoming dsDNA, we would obtain a distribution of delay times described by a single exponential curve. However, the distribution of delay times in our measurements is non-exponential and the data displays a rise phase followed by a decay (Figure 3.7A). If the time for joint molecule formation involved *N* hidden rate limiting steps prior to reaction completion, we can recapitulate the key features of our dwell time distribution plot. This approach to modeling dwell time distributions has been useful in estimating the kinetic step size of motor proteins in several biophysical studies (Myong et al., 2007; Park et al., 2010; Yildiz et al., 2003; Yildiz et al., 2004). Hence, assuming that joint molecule formation involves base pairing exchange in well-defined increments, the distribution of Δt_{delay} can be used to estimate how many basepairs are exchanged per rate-limiting step.

$$\text{begin} \xrightarrow{\overbrace{\quad k \quad k \quad \ldots\ldots \quad k \quad k \quad}^{\text{N steps}}} \text{end}$$

In this model, the histogram of Δt_{delay} should follow a gamma distribution, ${\Delta t_{delay}}^{N-1} e^{-k\Delta t_{delay}}$. *N* (the number of steps) and *k* (the reaction rate per step) are free parameters obtained after fitting the dwell time histogram with a gamma distribution for each DNA length. From fitting the data for L_h of 31, 39 and 45 nt (Figure 3.7A) we found that *N* increases with increasing L_h while *k* does not change significantly (Figure 3.7B and inset).

The slope of the linear fit of *N* vs. L_h, gave ~1/3 (step/bp) (Figure 3.7B) suggesting that basepairing occurs in a stepwise manner in 3 bp increments. The non-zero x-intercept (Figure 3.7B), ~14 bp, specifies the number of base pairs already exchanged for *N*=0, providing an estimate for the upper limit of the initial synaptic complex size.

It is important to note that while Figure 3.7A shows the distribution of the delay time for L_h=39 nt to be peaked at about 100-120 ms, there are more data points at longer times outside of the major peak which are responsible for the inflated averaged delay time that we plot in Figure 3.6C. Our step size analysis was restricted to the major peak assuming that molecules outside of this distribution may arise from a kinetically distinct species. Data from the DNA molecules with larger L_h were not analyzed in the same manner since they exhibited broad distributions with long time tails possibly due to multiple initiation sites along the DNA (Figure 3.6B).

3.2.4 Filament dissociation via ATP hydrolysis and heteroduplex formation

In order to ensure that the kinetics of joint molecule formation is not influenced by RecA turnover from the DNA, we carried out identical measurements using ATP as a cofactor. The same $<\Delta t_{delay}>$ delay was observed when ATP was used as the co-factor suggesting that the rate of joint molecule formation measured is independent of the co-factor used.

To test if ATP hydrolysis mediated dissociation of RecA from the heteroduplex product might affect the rate of joint molecule formation, we modified the donor and acceptor positions in the 'docking and pairing assay. For this measurement, the reference ssDNA was labeled at an internal position (position 8 of L_h= 39nt). The homologous incoming dsDNA (L_h= 39bp) was also labeled internally such that after heteroduplex formation, the dyes are finally separated by 9bp (Figure 3.8A).

Upon docking, molecules exhibited the initial low FRET delay period (Δt_{delay}) prior to the appearance of a mid-FRET state (E~ 0.4) corresponding to the stretched conformation of the heteroduplex product indicating that RecA still remains bound. Eventually, when the joint molecule is converted to a protein free heteroduplex due to RecA dissociation via ATP hydrolysis ($\tau_{dissociation}$), molecules exhibited a high FRET state (E ~ 0.7) (Figure 3.8B).

We first analyzed the delay period between initial docking and FRET change (Δt_{delay}) to obtain a dwell time histogram which after gamma distribution fitting gave us a value of N= 4 steps (Figure 3.8C). In this reaction since the reference ssDNA is internally labeled, we are monitoring basepair exchange only until position 8 of L_h= 39nt. After accounting for the initial synaptic complex size of ~ 14 bp, this DNA

construct would be effectively measuring propagation over a 17 nt region (i.e. 39 minus 8 minus 14). This corresponds to ~ 4 nt step size in propagation of joint molecule formation which is within 30% of the value supported by the main data of the paper obtained from constructs labeled at the terminal ends of the DNA (Figure 3.7B).

To measure the time for RecA dissociation via ATP hydrolysis after completion of joint molecule formation, we post-synchronized the FRET trajectories followed by fitting with a single exponential curve to obtain a dissociation time ($\tau_{dissociation}$) of ~ 30s (Figure 3.8D). Hence, RecA filament dissociation from the heteroduplex via ATP hydrolysis occurs on a much slower time scale than the time for joint molecule formation.

3.2.5 Strand separation and joint molecule formation occur concomitantly

To test our model further and measure the correlation between the kinetics of strand separation and joint molecule formation which was measured using the previous assay, we designed an alternative labeling strategy by attaching both fluorophores on the incoming dsDNA so that FRET reports on its local strand separation process (Figure 3.9A). For this measurement, we immobilized a DNA molecule of homology length, L_h= 39nt, with no fluorescent label and formed a RecA filament with ATPγS as a co-factor. We observed DNA docking to the RecA filament as an abrupt appearance of fluorescence signal in the high FRET state, E~0.85 (Figure 3.9B). Surprisingly, the disappearance of high FRET, an indication of strand separation (Ha et al., 2002; Myong et al., 2007), was followed by a period of rapid FRET fluctuations (marked by Δt_2 in Figure 3.9B) which lasted 3.3 s on average (Figure 3.11). Analogous to the docking and pairing assay, the high FRET period where the incoming dsDNA binds to the RecA filament in an intact conformation can be attributed to events where the initiation of joint molecule formation occurs at a position distal to the labeled end. The distance between the dyes is insensitive to the propagation of joint molecule formation until the reaction proceeds to the labeled end.

We measured the dwell time of the high FRET period (Δt_1) which would represent the time taken for strand separation at the labeled end. The dwell time of the initial high FRET state upon incoming DNA binding, Δt_1, displayed a narrowly peaked

distribution (Figure 3.10A), similar to the Δt_{delay} distribution observed in the previous docking-and-pairing assay. In addition, $N_{separation}$, the number of rate limiting steps present during Δt_1 obtained by a gamma distribution fit and $<\Delta t_1>$ were in agreement with those obtained for Δt_{delay} (Figure 3.10A-B). The strong correlation between the kinetic rates measured in these two assays suggests that strand separation in the incoming dsDNA and joint molecule formation with the reference ssDNA proceed concomitantly.

3.3 Conclusions

3.3.1 Formation of the initial synaptic complex

Our observations suggest that the formation of the initial synaptic complex is coincident with homology recognition. The upper limit to the size of the initial synaptic complex, which we estimated as <14 bp, is in good agreement with the earlier estimate of ~15 nt as the minimum length required for homology search and strand exchange (Hsieh et al., 1992). The size is also consistent with the upper limit of 16 bp placed on the synaptic length during homology search (van der Heijden et al., 2008).

Perturbations that destabilize the incoming dsDNA accelerate the formation of the initial synaptic complex (Lee et al., 2006). As for oligonucleotide substrates, the thermal breathing of duplex ends may provide preferred initiation sites for strand exchange. Based on the presence of single molecule time traces which can be categorized into two kinetically distinct populations (Figure 3.6A), we propose that the reaction primarily initiates from either the proximal or distal end relative to the fluorescent labels. Our data also shows that the preference for the ends decreases for longer homology lengths (Figure 3.6C inset) consistent with the idea that more initiation sites become available with increasing DNA homology length.

3.3.2 Propagation of basepair exchange

The kinetic analysis of our data for $L_h \leq 45$ nt allowed us to propose a model wherein joint molecule formation occurs in 3bp steps. This finding provides support for a model in which the homologous alignment of the incoming dsDNA with the RecA-bound ssDNA occurs in increments of 3bp and involves the local exchange of basepairs (Adzuma, 1992). Because of the gap between adjacent triplets observed

in the RecA crystal structure (Chen et al., 2008), the incoming dsDNA needs to be stretched locally to mediate alignment of successive triplets. The gap between adjacent triplets in the RecA-DNA complex might provide an explanation for the rate limiting step which we propose exists with the periodicity of 3bp. The distortion upon stretching the 3bp segment would unstack the bases at the gap and melt three basepairs. The newly freed triplet of nucleotides would then basepair with the triplet in the reference ssDNA.

As the DNA becomes longer, the delay time increased nonlinearly (Figure 3.6C) indicating that joint molecule formation becomes substantially slower for longer homology lengths. The non-linear dependence of the delay period is possibly due to the effect of multiple synaptic complexes and/or DNA topology. The current two-color FRET assay cannot distinguish between initiation from the end versus initiation in the middle of the filament. Previous magnetic tweezers studies on topologically constrained DNA could not observe strand exchange without negative supercoiling and the apparent rate of strand exchange was much slower, ~2 bp/sec, indicating that DNA topology may pose a significant barrier to the propagation of strand exchange (van der Heijden et al., 2008). Another possibility for the slower global strand exchange rate measured is the effect of torsional stress due to the concomitant rotation of dsDNA and the ssDNA-RecA filament complex (Honigberg and Radding, 1988; Rosselli and Stasiak, 1990) which could result in the slower joint molecule formation rates in the case of longer homology lengths (Figure 3.6B).

3.3.3 Dynamic interactions between DNA and RecA secondary binding site

The secondary binding site of RecA binds to the outgoing ssDNA during strand exchange (Mazin and Kowalczykowski, 1998). The binding of the outgoing ssDNA to the secondary binding site signals the completion of strand exchange (Mazin and Kowalczykowski, 1996). Given that our assay can detect the completion of joint molecule formation (Figure 3.6A), we propose that the fluctuations in the outgoing ssDNA interaction which we detect after strand separation (Δt_2 ,Figure 3.9B) might represent the bound state of DNA to the RecA secondary binding site.

The structure of the three-stranded complex upon joint molecule formation remains ambiguous with evidence pointing to the existence of a metastable three-stranded structure (Folta-Stogniew et al., 2004; Voloshin and Camerini-Otero, 2004) and an

alternative model proposing that the outgoing ssDNA strand is stabilized by the RecA secondary binding site where it remains bound until dissociation (Mazin and Kowalczykowski, 1998). We can rule out the existence of a stable triplex structure since our data demonstrates that the three-stranded structure formed in the presence of RecA is highly dynamic with the outgoing ssDNA displaying large excursions in FRET. Hence, our data is best explained by a model in which, following the separation of the two strands of the incoming dsDNA, the outgoing ssDNA is relayed to the secondary binding site of the RecA filament where it remains bound until dissociation or until its removal is facilitated by SSB (see Chapter 4 for results related to the role of SSB in strand exchange).

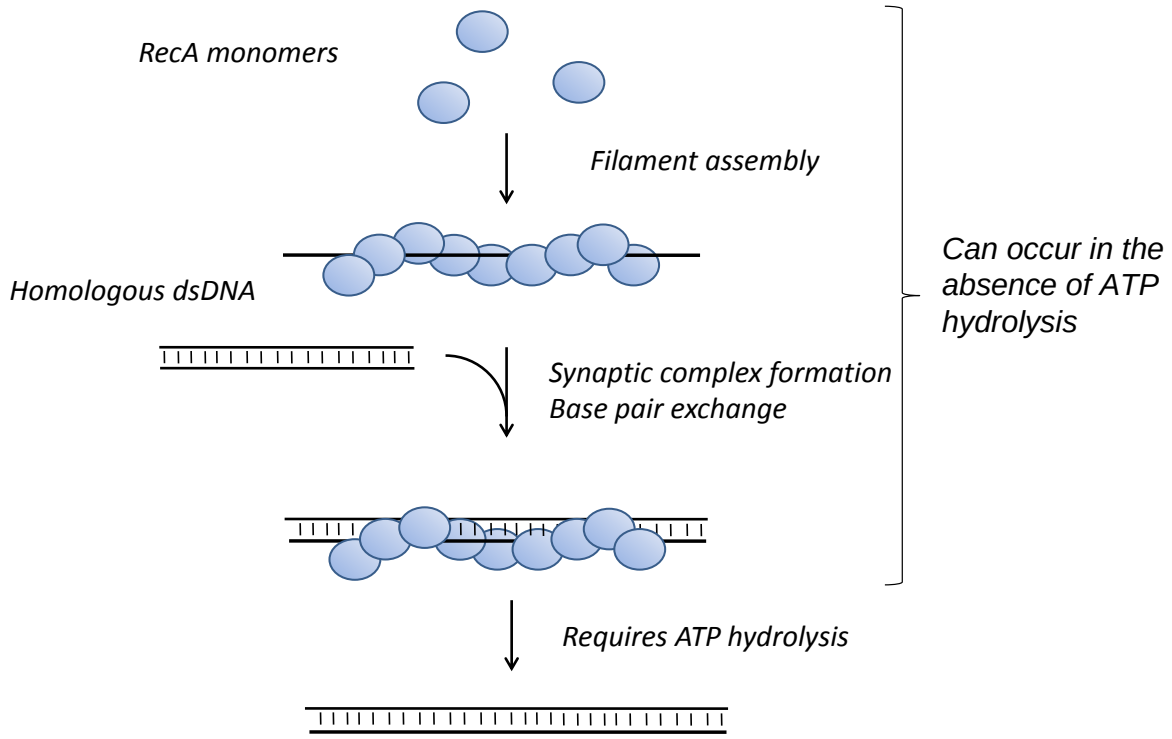
Because strand exchange is nearly isoenthalpic in terms of basepairing with the reference ssDNA and the basepairing which is lost in the incoming dsDNA, it remains unknown how strand exchange can be irreversible even in the absence of ATP hydrolysis. The dynamic nature of the outgoing ssDNA occupying the secondary site indicates its high conformational entropy even while it still remains bound to the RecA filament. We suggest that the entropy gain from the dynamic mode of interaction between the outgoing ssDNA and the RecA secondary binding site may provide a driving force for making the propagation of basepair exchange irreversible (i.e. unidirectional) in the absence of ATP hydrolysis.

3.3.4 Broader implications

Recent structural modeling of Rad51 with DNA proposed that the DNA bases were non-uniformly stretched with triplets being maintained in B-form configuration (Reymer et al., 2009) suggesting that the mechanism of strand exchange could be evolutionarily conserved across different recombinases. The tools and assays developed to study RecA mediated strand exchange can be applied to the study of Rad51 filaments and its accessory protein partners. Our single molecule approach can also be easily extended to other systems where homology or target search processes occur. For example, in eukaryotes, RNA interference is executed when small RNA-loaded RISC binds to a target mRNA guided by sequence recognition of 7-8 nt (Ameres et al., 2007; Bartel, 2009; Wang et al., 2009). It should be possible to observe the initial complex formation followed by the propagation of basepairing, the cleavage of the target mRNA, and the release of the decay products in real time.

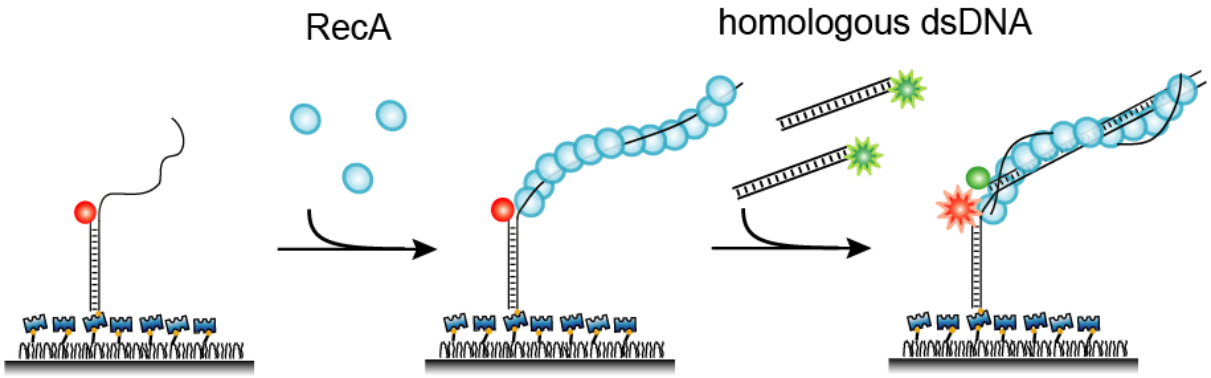
3.4 Figures

Figure 3.1 An overview of the RecA mediated strand exchange reaction



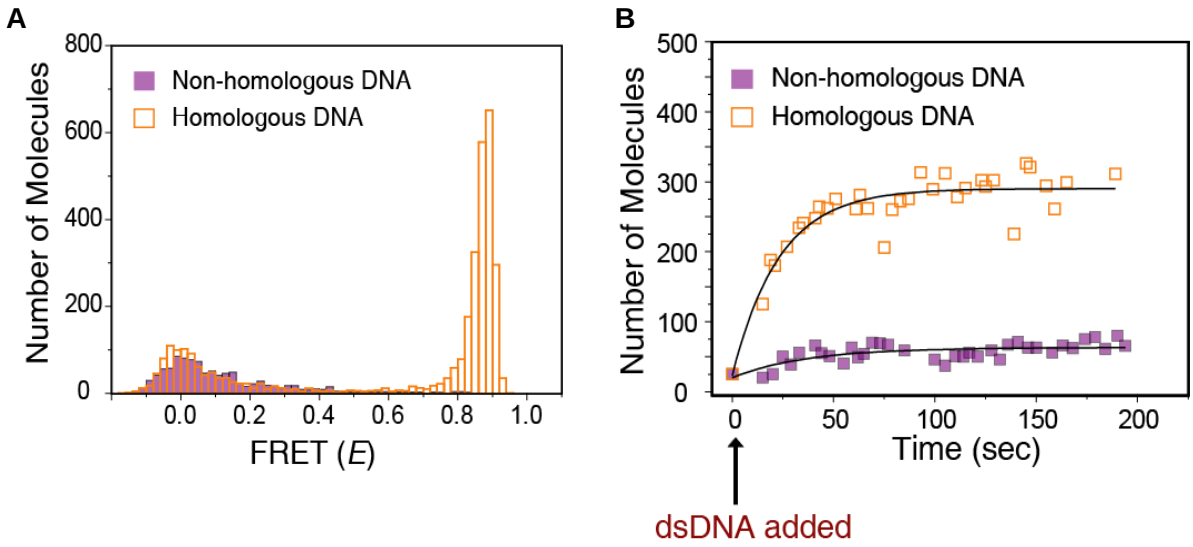
The strand exchange reaction involves multiple steps leading to the exchange of DNA strands between a RecA bound ssDNA and a target dsDNA containing a homologous site.

Figure 3.2 Single molecule assay to study RecA mediated strand exchange



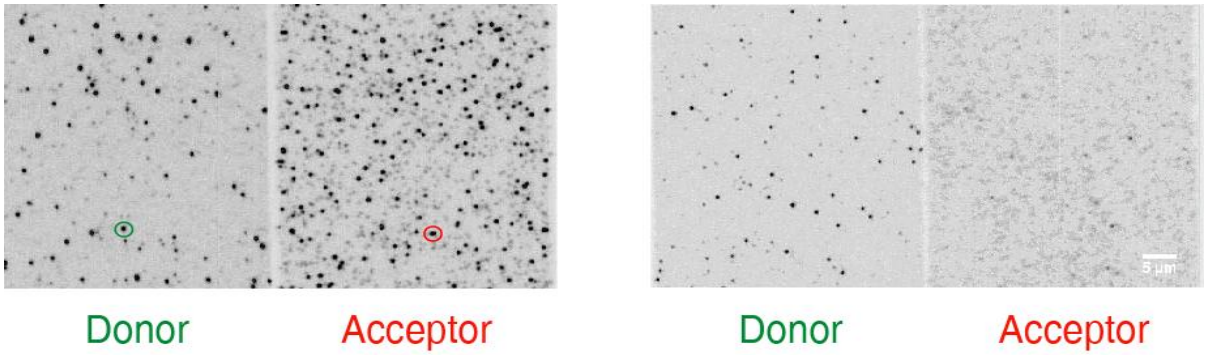
A schematic of the FRET-based strand exchange assay. The final heteroduplex product would have the donor (green) and the acceptor (red) in close proximity.

Figure 3.3 Single molecule strand exchange assay preserves reaction specificity



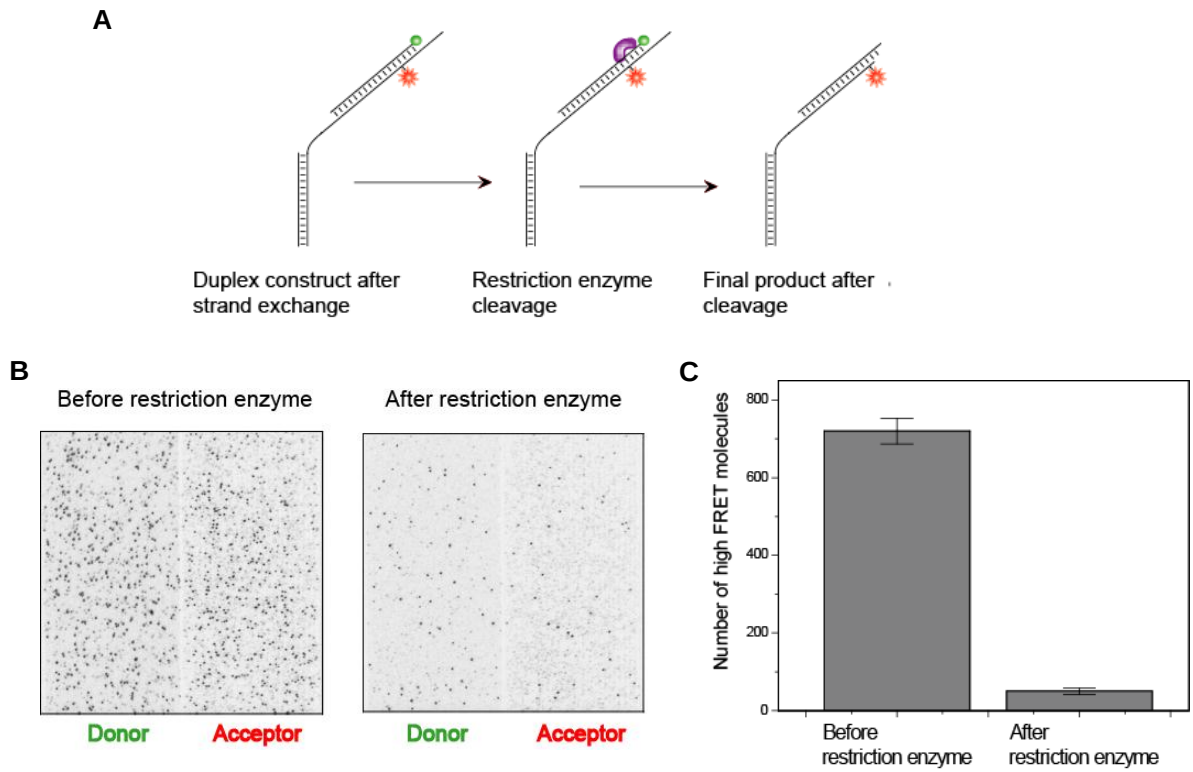
A. FRET efficiency histograms with immobilized homologous and non-homologous DNA (L_h= 39nt) obtained after 10 min reaction. dsDNA of length 39bp was used for each measurement. Data was obtained from 15 imaging areas each **B.** Number of Cy3-labeled dsDNA molecules (L_h= 39bp) bound to the RecA filament (L_h= 39nt) per imaging area vs. time. dsDNA was added at t=0.

Figure 3.4 Images of donor and acceptor emission channels



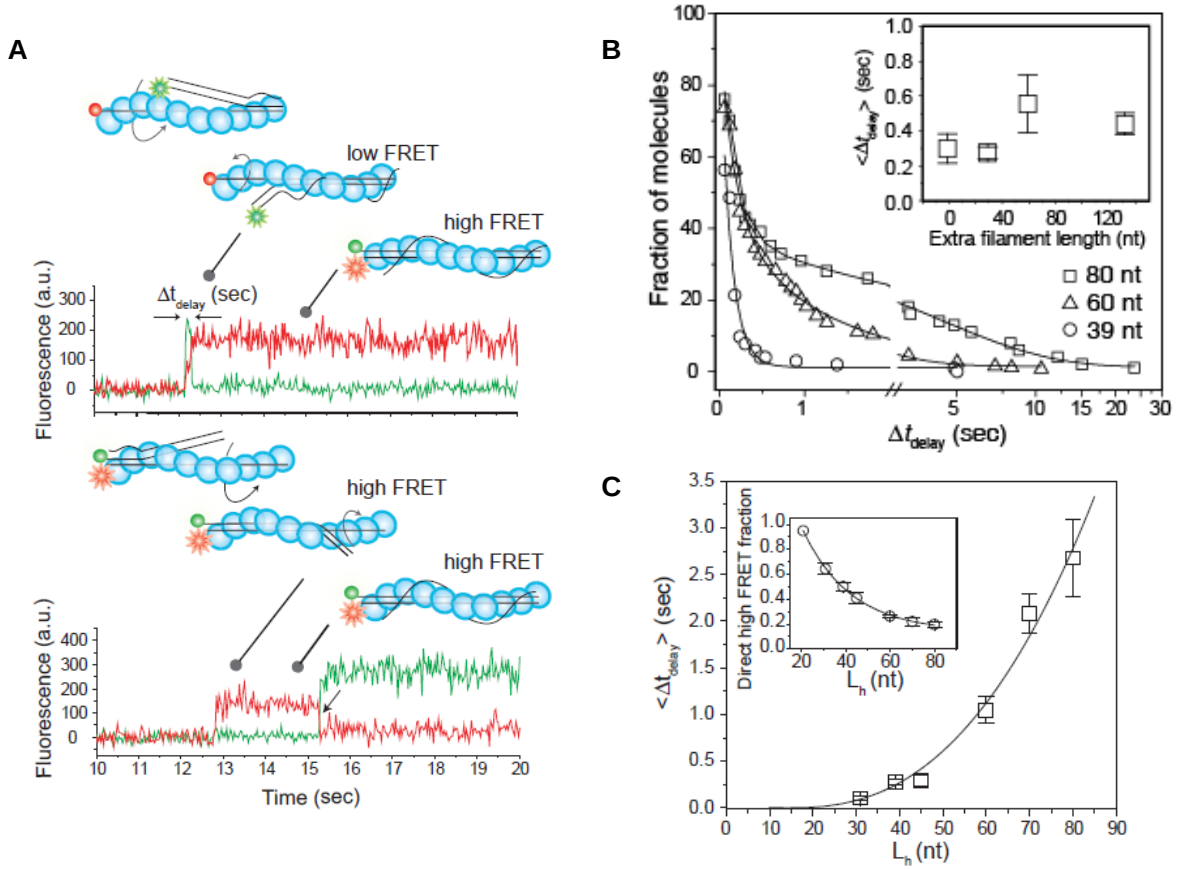
The green and red circles show an example of donor and acceptor spots. Scale bar = 5 μm.

Figure 3.5 Confirming heteroduplex formation after strand exchange



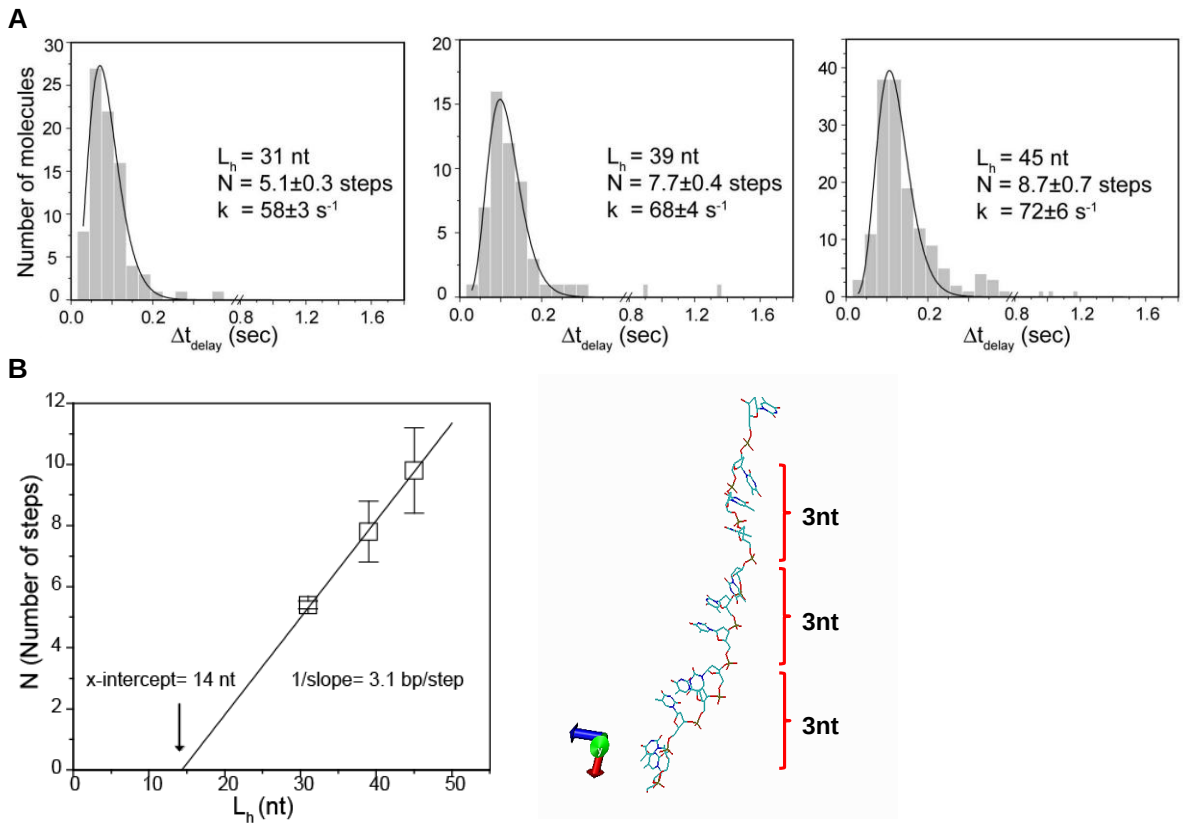
A. A reference ssDNA (L_h= 39nt) with an acceptor fluorophore (Cy5) internally labeled was immobilized. After RecA filament formation, strand exchange reaction was carried out using a homologous dsDNA labeled internally with a donor fluorophore (Cy3). A restriction site for the enzyme DdeI is present between the donor and the acceptor dyes such that the donor dye is released after cleavage of the DNA by the restriction enzyme **B.** CCD image of single molecules after strand exchange and RecA protein removal, prior to the addition of restriction enzyme (first panel) and after incubation of the strand exchange product with DdeI for 30 mins (second panel) C. Comparison of the number of FRET molecules (molecules exhibiting Cy5 fluorescence via FRET from Cy3) before and after the restriction enzyme reaction. Error bars denote the standard deviation of the number of molecules from multiple CCD images.

Figure 3.6 Real time observation of strand exchange propagation



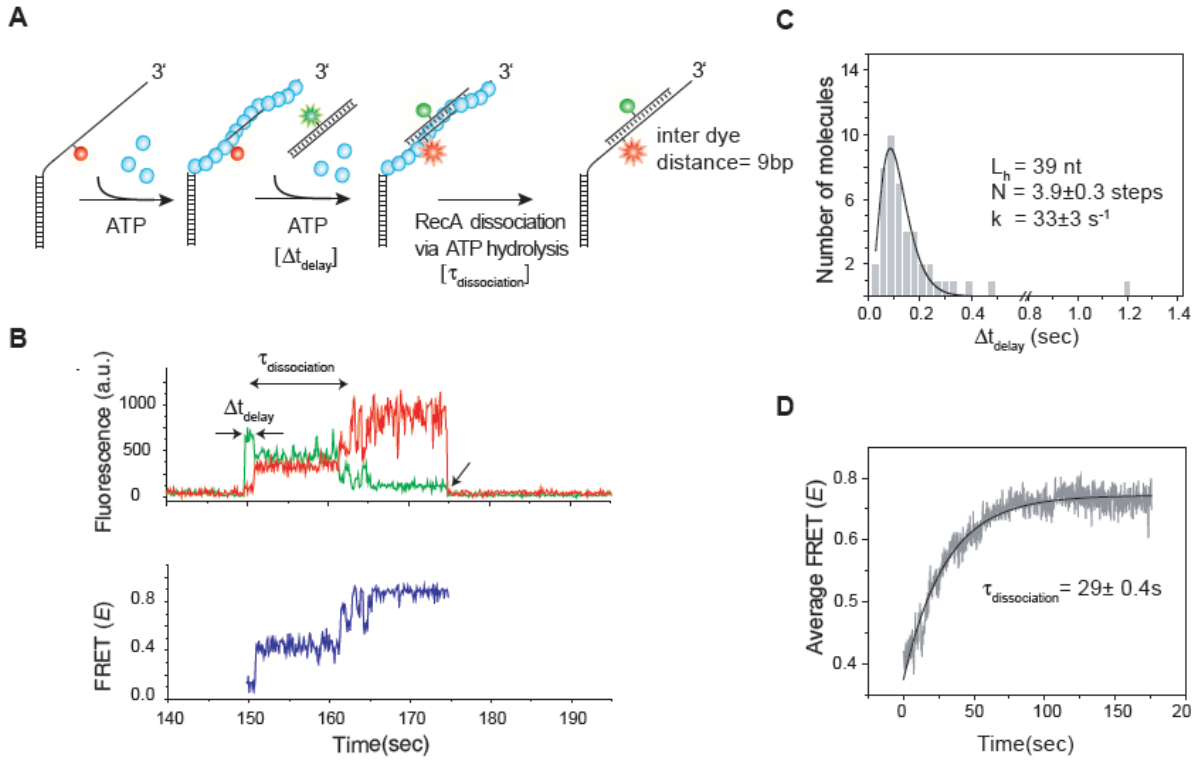
A. Single molecule time traces of donor (green) and acceptor (red) intensities showing docking and pairing for L_h= 39nt. Homologous dsDNA, L_h= 39bp was used in this measurement. In one class of events, a low FRET period of Δt_{delay} precedes the appearance of high FRET (top panel). In the other class of events, high FRET appears from the moment of docking (bottom panel). The cartoons above the time traces show the proposed reaction stages. The arrow in the bottom panel marks acceptor photobleaching. **B.** Survival probability of the initial low FRET state vs. time for L_h = 39 nt, 60 nt, 80 nt. dsDNA of corresponding lengths, L_h (bp), was used in each measurement**. C.** $\langle\Delta t_{delay}\rangle$ vs. L_h. A Power function fit was used as a guide. The population of Δt_{delay}=0 was not included when calculating the average. Inset shows fraction of molecules with Δt_{delay}=0 vs. L_h. Exponential fitting was used as a guide. Error bars are standard errors of the mean determined from three independent experiments.

Figure 3.7 Model for strand exchange propagation involving 3nt steps



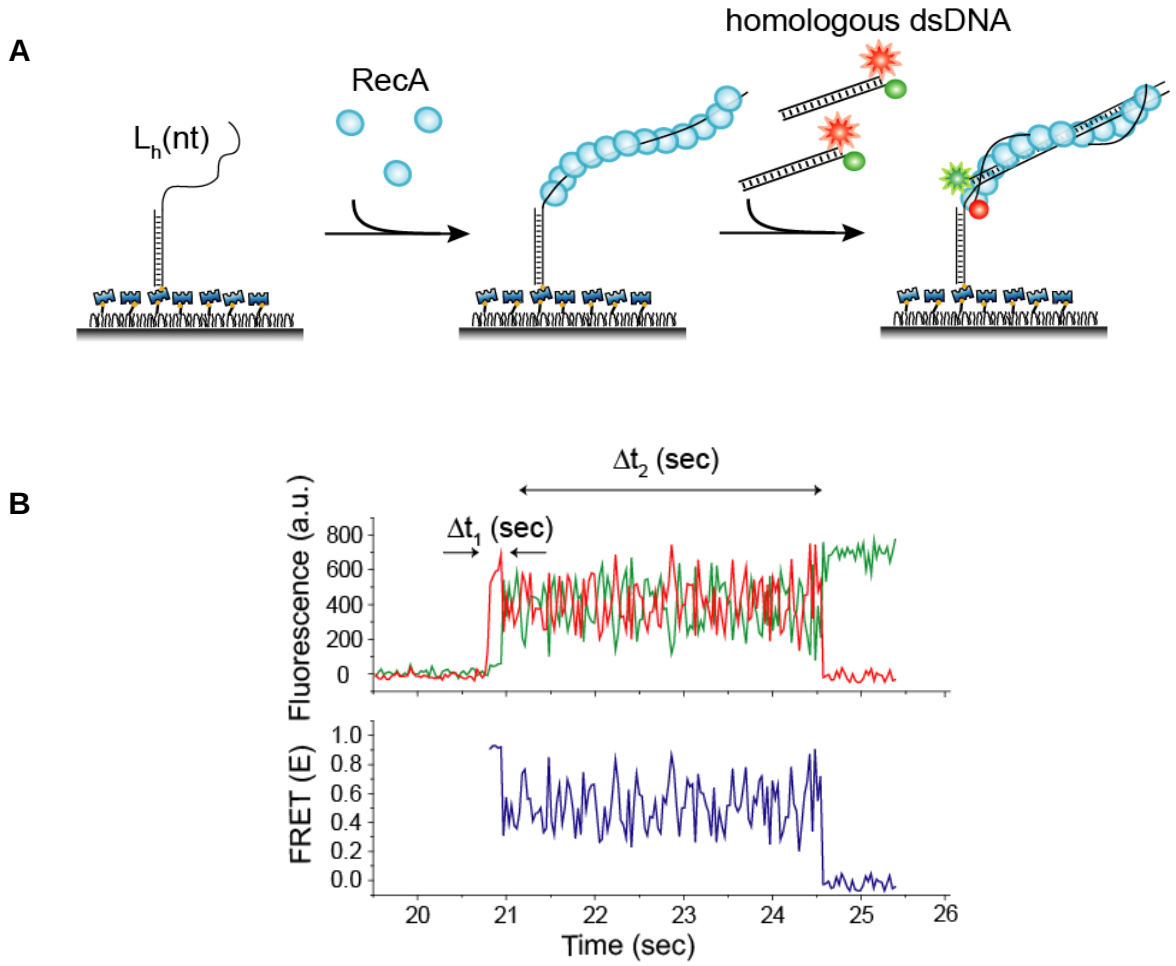
A. Gamma distribution fit of Δt_{delay} histograms for L_h= 39 nt. dsDNA of corresponding length, L_h=39bp, was used in this measurement. **B.** Number of steps, N vs. L_h. A linear fit of the data is shown. Inset shows the stepping rate k vs. L_h. Error bars are standard errors of the mean determined from three independent experiments.

Figure 3.8 ATP hydrolysis mediated dissociation of RecA from the nascent heteroduplex



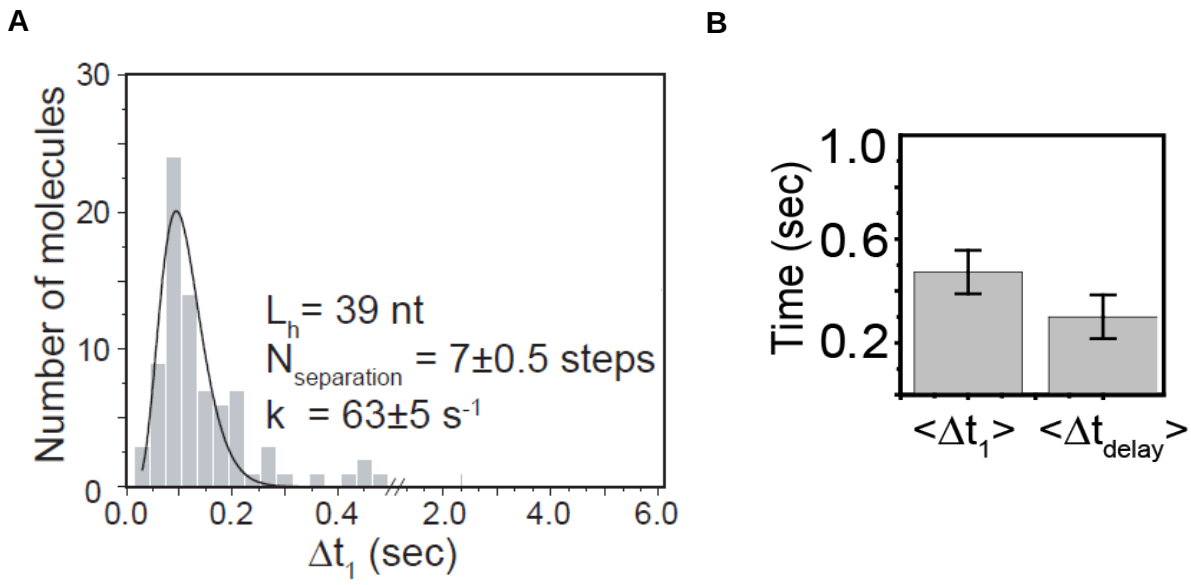
A. A schematic of the modified FRET-based docking and pairing assay. Cy3 labeled dsDNA (L_h= 39bp) binds to a RecA filament formed on an acceptor labeled reference ssDNA (L_h= 39nt) such that the distance between the dyes in the final product is 9bp. **B.** Single molecule time traces of donor (green) and acceptor (red) intensities display a low FRET(~0.1) upon initial binding (Δt_{delay}) corresponding to dsDNA docking followed by FRET change to a mid-FRET state (~0.4) indicating the RecA bound state of the heteroduplex finally followed by RecA dissociation via ATP hydrolysis ($\tau_{dissociation}$) leading to a high FRET state (~0.7). The arrow in the top panel marks donor photobleaching and loss of FRET. **C.** Gamma distribution fit for the initial low FRET period prior to heteroduplex formation. This construct monitors propagation over a 17nt segment of the DNA. **D.** Averaged FRET trajectory for the time period, $\tau_{dissociation}$ and a single exponential fit for the RecA filament dissociation time.

Figure 3.9 Strand separation assay



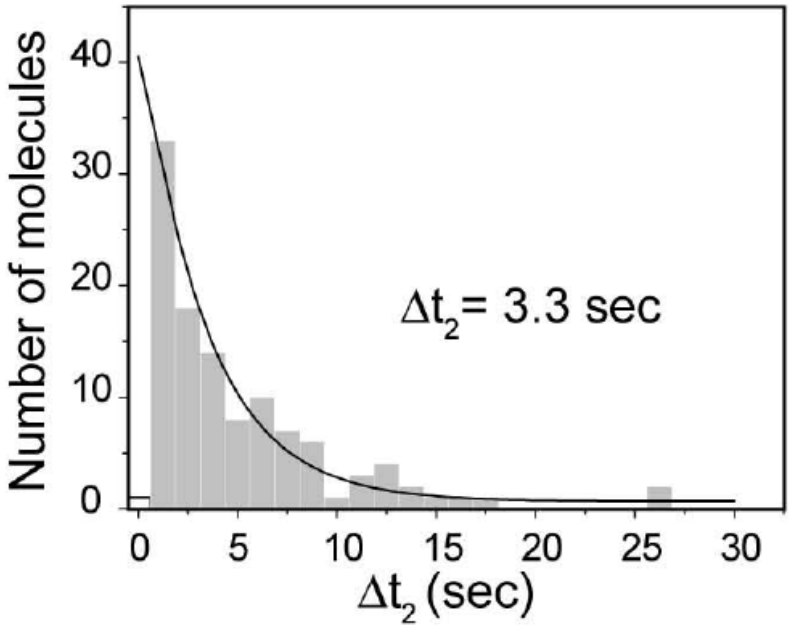
A. A schematic of the FRET-based strand separation assay. A doubly labeled dsDNA (L_h= 39bp) binds to a RecA filament formed on an unlabeled reference ssDNA (L_h= 39nt). Only the donor-labeled strand remains on the surface after the reaction and the acceptor-labeled outgoing ssDNA is eventually released. **B.** Single molecule time traces of donor (green) and acceptor (red) intensities display a high FRET period upon binding (Δt_1) followed by rapid fluctuations in FRET for a period Δt_2 as shown until disappearance of acceptor signal.

Figure 3.10 Strand separation kinetics



A. Histogram of Δt_1 and a fit to Gamma distribution. The population of $\Delta t_{delay}=0$ was not included when calculating the average. **B.** $\Delta t_1>$ and $<\Delta t_{delay}>$ are similar. Error bars are standard errors of the mean determined by bootstrapping.

Figure 3.11 Histogram of Δt_2 and a single exponential decay fit.



Chapter 4

Protein-protein interactions facilitating DNA repair[3]

No protein is an island entire of itself,
As a piece of the continent;
A part of the main
adapted from John Donne(1624)

4.1 Protein interactions assist DNA repair

The orchestration of DNA repair processes engages a variety of proteins that coordinate with each other to successfully complete repair at the site of DNA damage. In the introductory chapter (Chapter 1), I provided examples of how RecA filament interactions with proteins such as LexA influence the response of the cell to DNA damage by initiating the SOS response pathway. Here I will outline other examples of interactions between proteins involved in DNA repair and show how we can use single molecule FRET approaches to study their mechanisms.

RecA interacts with proteins that assist in filament formation, filament disassembly and play key roles in assisting the strand exchange reactions. This chapter will provide some insights into a few examples of protein-protein interactions involving RecA and other accessory proteins. Most of the observations serve as a platform for future work. 1) RecA interactions with SSB after strand exchange completion. which can 2) RecA interactions with RecX – a protein involved in RecA disassembly 3) Interaction between a helicase (Rep) and SSB

4.2 SSB interaction with outgoing ssDNA during post-synapsis

4.2.1 Introduction

Single Stranded DNA Binding protein (SSB) has both a pre-synaptic and post-

[3] Portions of the work in Chapter 4 have been published
• **K. Ragunathan**, C. Joo and T.Ha, "Real time observation of strand exchange reaction with high spatiotemporal resolution", Structure 19(8) 1064-73 (2011)

synaptic function during strand exchange. During pre-synapsis, SSB binds to single stranded (ss) and by virtue of being able diffuse along ssDNA via reptation (Roy et al., 2009; Zhou et al., 2011), SSB disrupts secondary structures and allows for uninterrupted assembly of RecA filaments. During post-synapsis, SSB is proposed to interact with the strand which is poised to depart from the three strand complex after basepair exchange and heteroduplex formation (Mazin and Kowalczykowski, 1996, 1998).

4.2.2 Single molecule measurements of SSB interaction with outgoing ssDNA

Rapid FRET fluctuations observed in the strand separation assay persisted long after joint molecule formation must have finished, that is $\Delta t_2 >> \Delta t_1$ (see Chapter 3, Figure 3.9). Therefore, the outgoing strand remains associated with the heteroduplex product after base pair exchange has been completed. In order to confirm that the rapid fluctuations in FRET arise from conformational changes between the outgoing ssDNA and the nascent heteroduplex product, we used an alternative labeling scheme where the donor is attached to the outgoing strand and the acceptor to the reference strand immobilized on the surface (Figure 4.1). We reproduced the observation that the outgoing DNA strand remains bound to the RecA filament after base pairing exchange. This assay also showed the extended period of FRET fluctuations which we observed in the previous strand separation assay implying large scale relative motion between the outgoing ssDNA and the heteroduplex product (Figure 4.2). In addition, the lifetime of the three-stranded complex was identical within error for these two configurations (Figure 4.3). In comparison, photobleaching timescale was about an order of magnitude longer (See Appendix C). Furthermore, the lifetime of the three-stranded complex decreased substantially when SSB was added (Figure 4.4) (Mazin and Kowalczykowski, 1998). To monitor the removal of the outgoing ssDNA by SSB, we formed a pre-synaptic filament on a reference ssDNA (L_h= 80 nt) and flowed a solution containing homologous dsDNA along with SSB protein. L_h of 80 nt was used for the SSB analysis because an SSB tetramer requires a minimum of ~65 nt for binding ssDNA under our buffer conditions using 10 mM Mg^{+2} and 100 mM Na^{+} (Lohman and Overman, 1985).

4.3 RecX interacts with RecA to stimulate filament disassembly

4.3.1 Introduction

RecX is an *E.coli* protein that mediates RecA filament disassembly. The protein consists three tandem three helix repeats and acts as a negative regulator of the homologous recombination pathway. Crystal structures of RecX (Ragone et al., 2008) and subsequent docking simulations with a RecA filament showed that the basic residues within RecX are important for promoting its interaction with the acidic portion of the RecA filament. The mechanism of RecX mediated dissociation of the RecA filament is not clear. One model proposes that RecX caps the filament ends hence promoting filament disassembly (Drees et al., 2004). Alternatively, another model posits that RecX can trigger filament disassembly by increasing the number of ends available for RecA disassembly (Ragone et al., 2008). In effect, such a model proposes that RecX interrupts the coupling between adjacent RecA monomers. Alternatively, the faster rate of RecA disassembly could be a consequence of an increase in the ATP hydrolysis rate mediated by RecX association with RecA. We sought to test these models by using our single molecule FRET platform for which several features have been optimized to observe RecA binding and dissociation.

4.3.2 Preliminary single molecule data

We first immobilized a DNA with an acceptor fluorophore at the ssDNA/dsDNA junction with 3' polarity and an overhang of length, 17nt. The DNA only on gives a FRET value of ~ 0.8 (Figure 4.5A**)**. In the presence of RecA and ATPɣS the FRET value shifts to the low FRET state resulting from RecA binding and extension of the ssDNA (Figure 4.5B). In control experiments, RecX does not dismantle filaments bound with ATPɣS. In order to observe filament formation on a short DNA strand (17nt in this experiment) using ATP as a co-factor, we needed to utilize a strategy of creating a hybrid filament (Joo et al., 2006). Briefly, we first needed to form a filament with RecA and ATPɣS. This results in complete filament formation along the ssDNA portion and the dsDNA portion. Then we washed away the excess RecA and ATPɣS which results in complete dissociation of the ssDNA portion of the filament but preserves the RecA/dsDNA filament (Figure 4.5C). This is intrinsic to the property of RecA filament assembly and it is unclear why filaments on duplex DNA are relatively more stable compared to those formed on ssDNA. We then add RecA

in the presence of ATP which uses the pre-formed dsDNA portion of the RecA filament as a nucleation cluster to initiate filament formation (Figure 4.5D).

After forming filaments on the short DNA substrate, we added RecX together with RecA and ATP (since removing free RecA in solution would shift the equilibrium to the dissociated state in a RecX independent manner). RecX induces rapid disassembly of RecA filaments as evidenced by the presence of a high FRET peak (Figure 4.6) similar to the peak observed in the histogram with DNA only (Figure 4.5A). Single molecule traces provide a very interesting view of the disassembly process. The contrast between single molecule time traces in the presence of RecA (Figure 4.7A) and those in the presence of RecA and RecX (Figure 4.7B) is striking. In the presence of RecA the DNA is maintained in a stretched conformation shifting the histogram to a low FRET state. In the presence of RecX, the DNA is mostly in the high FRET state with frequent attempts of filament formation (low FRET) and subsequent disassembly (high FRET). Further analysis is required to understand the mechanistic basis for RecX induced RecA disassembly. The results presented here establish a single molecule assay for testing RecX induced RecA disassembly.

4.4 Interaction between a helicase and SSB

4.4.1 Introduction

Helicases play important roles in a variety of physiological contexts that include repair and replication. SSB is recruited to bind ssDNA during repair and replication and the tight binding of SSB poses barrier to the function of other proteins which utilize ssDNA. Rep belongs to the SF1 family of helicases and exhibits 3'->5' translocation activity along ssDNA. We tested the effect of adding Rep to ssDNA in the presence of a bound SSB. There is no known functional interaction between Rep and SSB. Hence, the mechanisms of a helicase overcoming such a roadblock might serve as a canonical example of how a translocase might move along a ssDNA track in the presence of bound proteins.

4.4.2 Preliminary single molecule FRET data

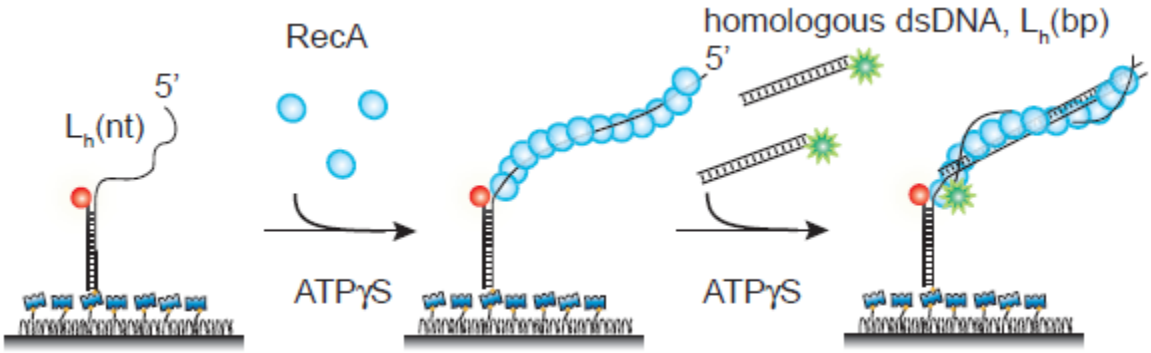
We utilized a ssDNA with an overhang of length, 80nt and an acceptor fluorophore at the ssDNA/dsDNA junction. This length provides an ideal substrate SSB binding given that under high salt conditions SSB wraps around 65nt of ssDNA. We first

tested the effect of adding Rep labeled with a donor (Cy3) fluorophore. Protein binding results in the appearance of FRET changes (Figure 4.8A) which reflect the characteristic patterns observed during translocation and snapback as previously reported (Myong et al., 2005). We then challenged Rep by providing it with the same DNA substrate but this time we pre incubated the ssDNA with SSB. In this case, we observed a dramatic decrease in the number of events exhibiting FRET indicating that SSB binding to DNA excludes Rep from gaining access to the ssDNA substrate. However, a fraction of molecules (~5%) exhibited FRET trajectories indicative of Rep translocation (Figure 4.8B). Interestingly the translocation did not involve multiple cycles and Rep binding to ssDNA pre bound with SSB seemed exhibit one or few translocation events with a larger translocation periods compared to free ssDNA.

It is likely that the number of events may be increased if we were to use ssDNA which is longer than 80nt to provide a docking site for Rep while SSB is still bound. Experiments using labeled SSB will also be required to confirm that the translocation events observed are bonafide events of translocation of Rep along SSB bound ssDNA. It is interesting that even in the presence of a bound SSB, Rep might still translocate along ssDNA suggesting that proteins in the path of translocation may not have an all or none sort of an effect but may modulate access to ssDNA while still permitting protein function (Honda et al., 2009).

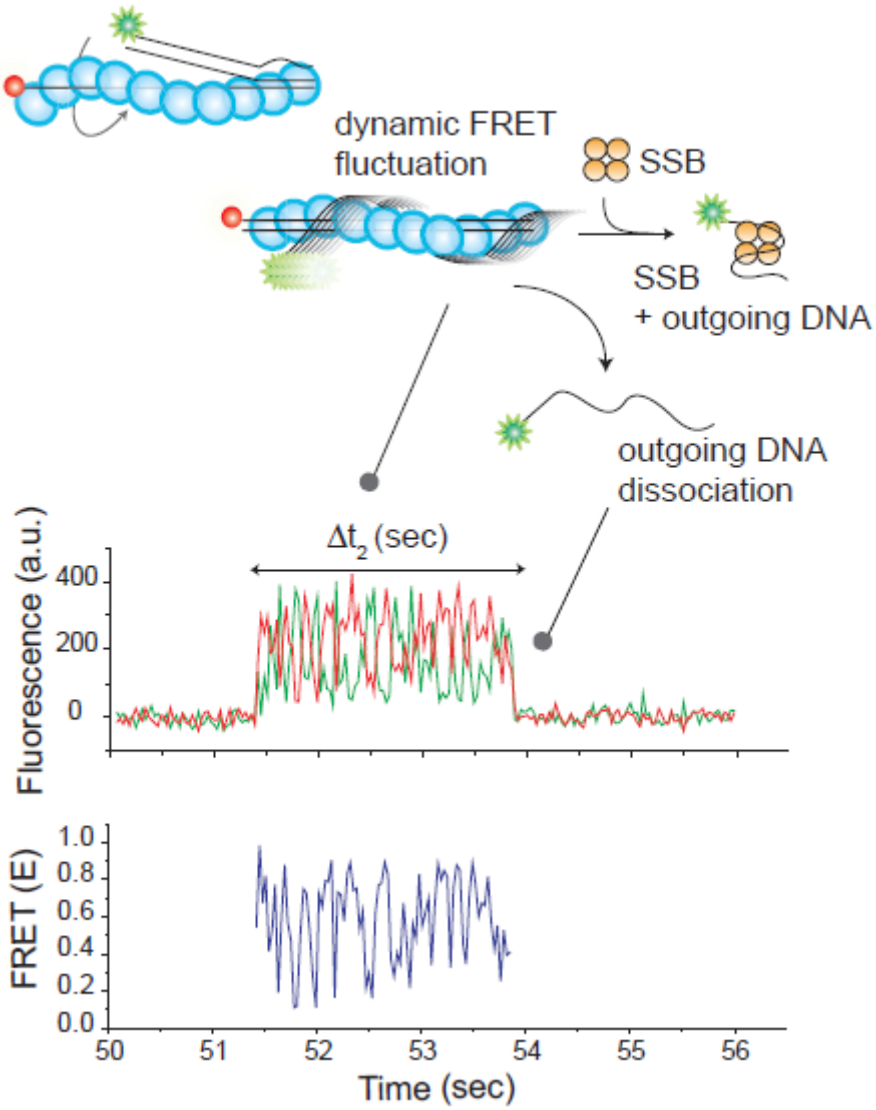
4.5 Figures

Figure 4.1 Outgoing strand assay



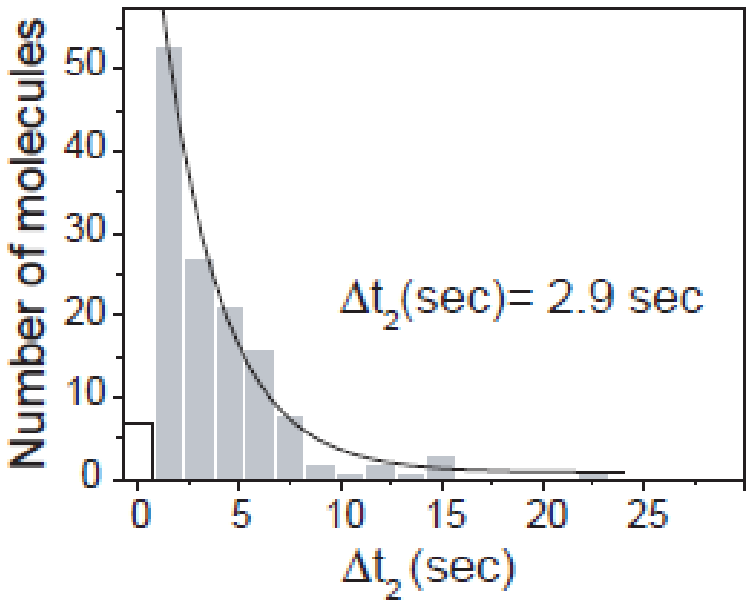
A schematic shows the binding of a donor-labeled dsDNA (L_h= 39bp) to a RecA filament formed around an acceptor-labeled reference ssDNA (L_h= 39nt) and the release of the donor-labeled outgoing ssDNA.

Figure 4.2 Outgoing ssDNA bound to RecA exhibits rapid FRET fluctuations



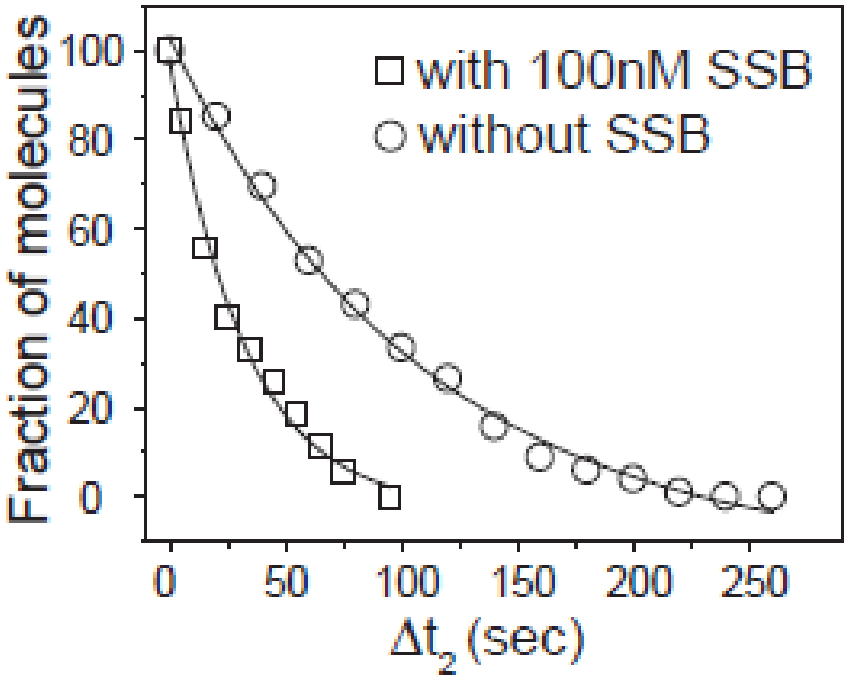
Single molecule time traces of donor (green) and acceptor (red) intensities exhibit rapid fluctuations in FRET over a period marked by Δt_2 until signal disappearance likely due to outgoing ssDNA release.

Figure 4.3 Dwell time of outgoing ssDNA bound to RecA filament



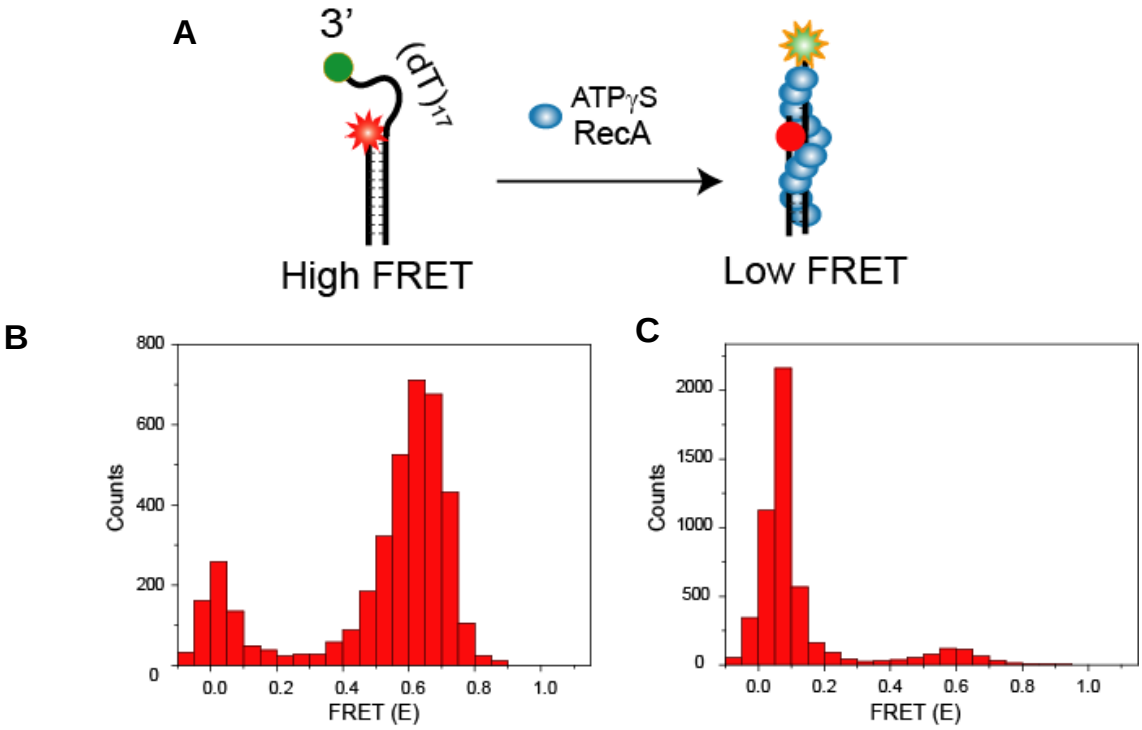
Histogram of Δt_2 and a single exponential decay fit

Figure 4.4 SSB assists in removal of outgoing ssDNA bound to RecA



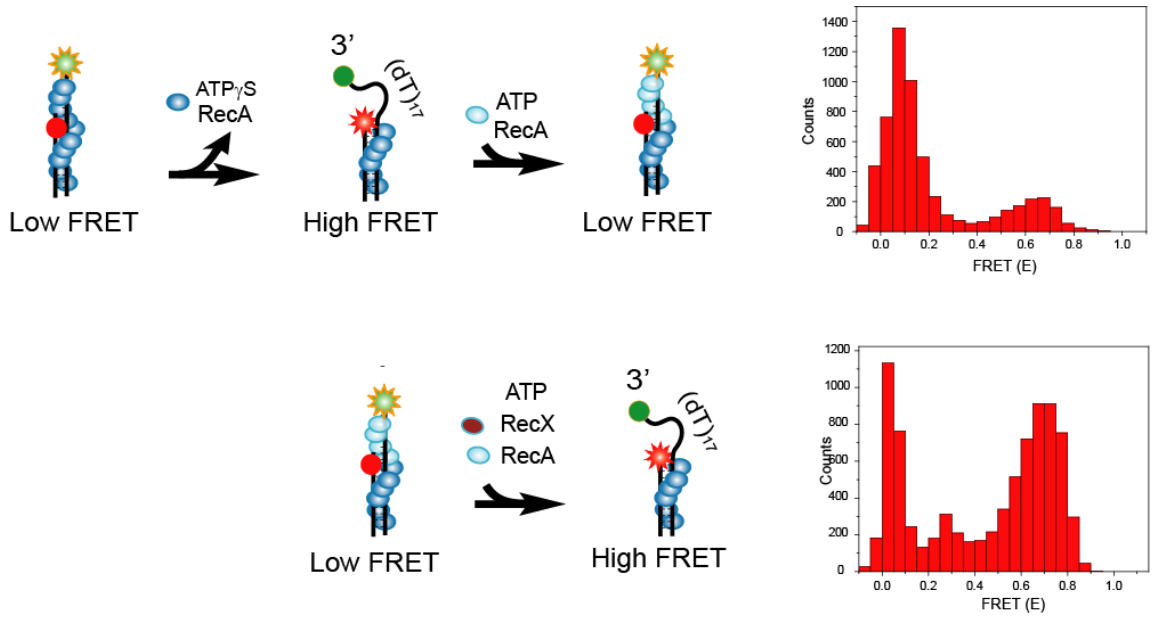
SSB (100nM) flowed along with homologous dsDNA facilitates the removal of the outgoing DNA strand (L_h= 80 nt).

Figure 4.5 Construction of mixed RecA filaments with pre-nucleation cluster



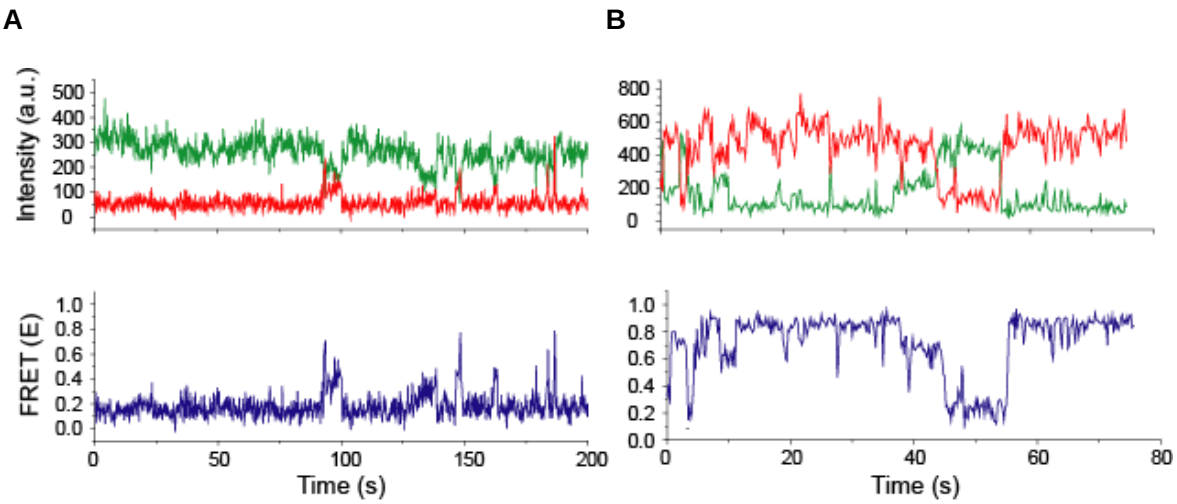
A. Schematic demonstrating the formation of RecA filaments in the presence of ATPɣS. **B.** DNA with a poly T sequence (17nt) with donor and acceptor dyes was immobilized on the surface. FRET histogram indicates a high FRET peak ~ 0.6 in the absence of protein **C.** Addition of RecA and ATPɣS stretches the DNA resulting in a shift in the FRET towards the low FRET state ~ 0.1.

Figure 4.6 RecX addition triggers RecA disassembly



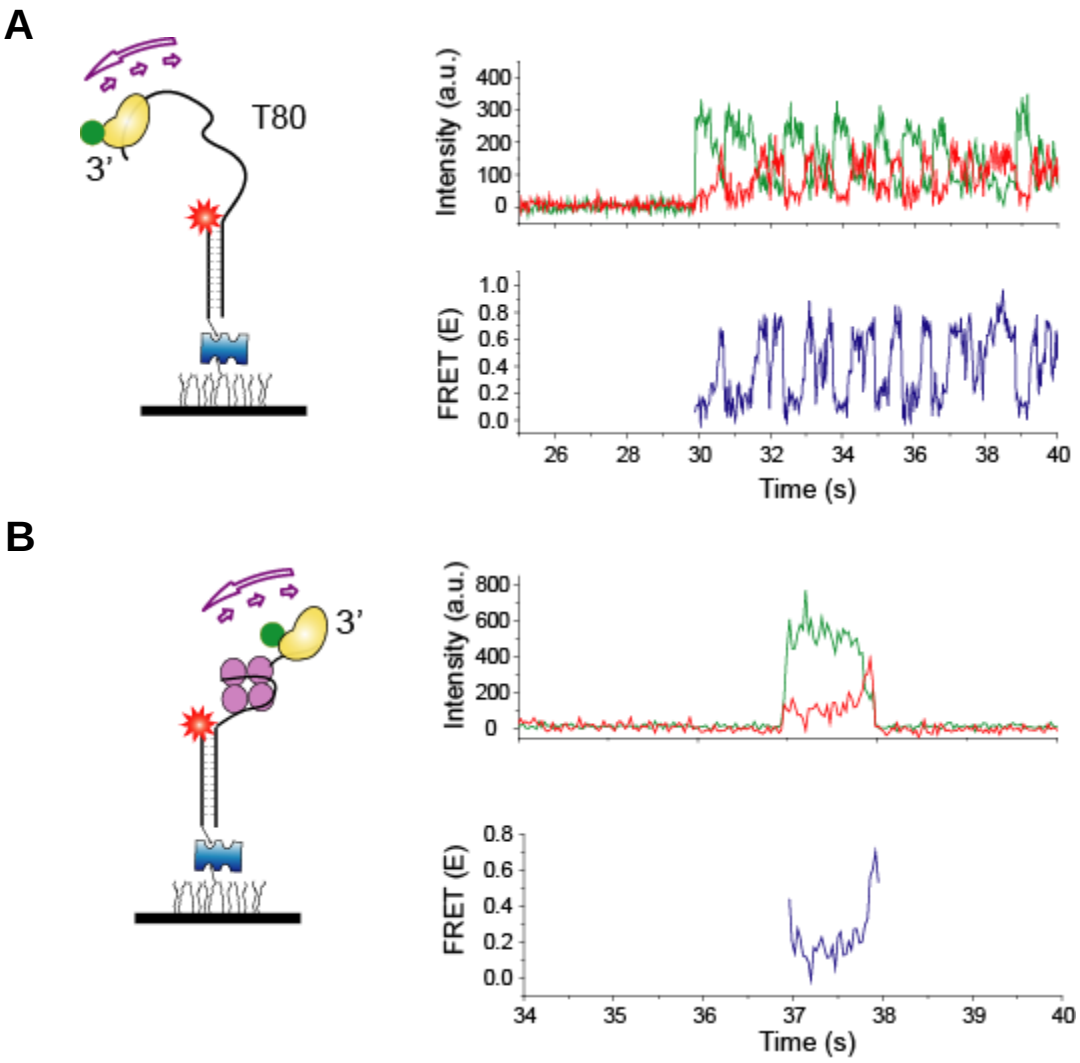
A. Schematic demonstrating the formation of hybrid RecA filaments (Joo et al., 2006) using a RecA pre-nucleation cluster formed on dsDNA. The presence of a low FRET state demonstrates that the DNA is bound by RecA **B.** Addition of RecX (1μM) with RecA (1μM) to the hybrid filament triggers filament disassembly resulting in a shift in the FRET histogram to the free DNA only state (Figure 4.5B)

Figure 4.7 Single molecule time trace of RecX mediated disassembly of RecA filaments



A. Single molecule time trace with donor (green) and acceptor (red) intensities with the corresponding FRET values (blue) with RecA (1μM) and ATP in solution. **B.** Same as A, except with RecX (1μM) in solution in addition to RecA and ATP

Figure 4.8 Rep translocation on ssDNA in the presence of bound SSB



A. Single molecule time trace with donor (green) and acceptor (red) intensities with the corresponding FRET values (blue) of Rep translocation along ssDNA **B.** Same as A, except with bound SSB

Chapter 5

Sliding facilitates RecA mediated homology search

During a search of homology, sliding
of the PSC along a dsDNA molecule
does not occur to any significant extent.
Kenji Adzuma (1998)

5.1 Introduction to the homology search problem

While the binding affinity of free RecA monomers to ssDNA does not have a strong sequence dependence during presynapsis, the formation of a complex between ssDNA and RecA results in a sequence specific binding entity which is capable of searching and forming complementary basepairs at a homologous site within a target dsDNA (Kuzminov, 1999) containing as few as ~ 8 complementary basepairs (Hsieh et al., 1992).

However, whether homology recognition within a target dsDNA occurs by 3-D diffusion of the RecA-ssDNA complex and iterative collisions with a target dsDNA or whether 1-D sliding of the filament along the target ssDNA can also facilitate homology search, remains unclear. Initial studies addressing this question ruled out a role for long range 1-D sliding along the dsDNA (~ several kb) (Adzuma, 1998). A recent study suggested that homology search occurs via intersegmental transfer wherein structural properties of the RecA filament allow for weak and iterative interactions with non-contiguous sequences of a dsDNA with emphasis on a random coil-like conformation for the duplex (Forget and Kowalczykowski, 2012). However, this study was restricted to the visualization of the end products after completion of strand exchange and no real time information of the homology search process was provided.

Target site search involving sliding has been studied using single molecule approaches for a wide variety of DNA binding proteins which includes restriction enzymes (Bonnet et al., 2008), transcription factors (Wang et al., 2006), reverse transcriptases (Liu et al., 2008) DNA repair proteins such as DNA glycosylase (Blainey et al., 2006), mismatch repair complexes (Gorman et al., 2007) and Single

Strand DNA Binding (SSB) protein (Roy et al., 2009; Zhou et al., 2011). In addition, single molecule approaches have also been utilized to decipher details of the structural and functional properties of RecA (Fulconis et al., 2006; Galletto et al., 2006; Jain et al., 2011; Joo et al., 2006; Mani et al., 2009; Shivashankar et al., 1999; van der Heijden et al., 2008; van Loenhout et al., 2009) and Rad51 filaments (Modesti et al., 2007; Yeykal and Greene, 2006). In this study we use single molecule FRET (Ha et al., 1996) to study the mechanism of RecA-mediated homology search. Our ability to observe the interaction between RecA and dsDNA with high spatiotemporal resolution revealed an unexpected dynamic mode of interaction between the dsDNA and the RecA filament. Using two and three color FRET measurements, we show that the dynamics observed are a consequence of RecA filament sliding along dsDNA. We show that sliding is primarily mediated by the electrostatic interactions between dsDNA and the RecA filament suggesting that the same structural features of the RecA filament that are involved in homology search via intersegmental transfer may also facilitate 1-D sliding over short length scales of dsDNA. Furthermore, by using sequences with short stretches of homology, we were able to track repeated events of homology recognition and basepairing in real time without full dissociation of dsDNA from the RecA filament. Our study demonstrates that RecA-mediated homology search serves as an example of how a DNA bound multi protein complex can present a vehicle for enabling sliding and target search process. This system could serve as a canonical example for other proteins (eg. telomerase and Argonaute) which are bound to nucleic acid sequences that act as ‘guide’ strands conferring target site specificity analogous to the manner in which a ssDNA within the RecA filament specifies the site of homology within a dsDNA.

5.2 Results

5.2.1 Dynamic interactions between RecA filament and non-homologous dsDNA

For short (< 80 bp) homologous dsDNA substrates, homology search is rapidly completed (within 30 ms) once it encounters a RecA filament . To understand the mechanism by which the RecA filament recruits an incoming dsDNA during the homology search phase, we diminished the effect of complementary base pair

interactions to measure repeated events of homology search in the absence of stable product formation.

We immobilized a partial dsDNA molecule on a passivated quartz surface (Figure 5.1). The immobilized DNA construct consists of a ssDNA overhang region ($L_{filament}$= 39 nt) labeled with an acceptor dye (Cy5) at the junction. The DNA is immobilized on a polymer-passivated surface via a biotin-neutravidin interaction. The addition of RecA and ATPɣS results in the formation of a stable filament allowing us to observe the interactions between dsDNA and RecA since RecA dissociation is inhibited by the slowly hydrolyzing ATP analogue. After filament formation, we added a solution containing non-homologous dsDNA (L_{dsDNA}= 39 bp) and ATPɣS while simultaneously removing free RecA from solution which allowed us to observe the interaction solely between the incoming non- homologous dsDNA (free of RecA) and single immobilized RecA filaments. We previously demonstrated that RecA filaments formed under such solution conditions are stable and can carry out the strand exchange reaction with a homologous dsDNA.

Docking of non-homologous dsDNA to the RecA filament is detected as an abrupt appearance of fluorescence signal from the background level. After docking, we observed large and rapid fluctuations in FRET detected as anti-correlated changes of donor and acceptor intensities. The large fluctuations in FRET are indicative of extensive distance changes between the donor on the dsDNA and the acceptor on the RecA/ssDNA filament. Single molecule time traces show multiple binding and dissociation events suggesting that the observed interactions are transient so that the binding events do not result in the formation of a stable product (Figure 5.2A). The lifetime of the binding events is exponentially distributed with an average of 3.5 seconds (Figure 5.2B). Filaments formed in the presence of ATP displayed the same large fluctuations in FRET upon docking of non-homologous dsDNA and exhibit comparable dissociation times to filaments formed in the presence of ATPɣS (Figure 5.3A-B). Hence, the properties of the RecA filament are independent of the co-factor used and the fluctuations observed do not require ATP hydrolysis. Using a different length of the immobilized ssDNA ($L_{filament}$= 50 nt or 99 nt) did not significantly change the lifetime of the dsDNA interaction (Figure 5.4A) while increasing dsDNA length resulted in an increase in the lifetime presumably due to a larger number of contacts between the dsDNA and the RecA filament (Figure 5.4B).

We measured whether the lifetime of the complex between non-homologous dsDNA and the RecA filament depends upon the magnesium or sodium concentration. Lowering Mg^{+2} concentration from 10 mM to 1 mM caused an ~8 fold increase in the lifetime of the complex (Figure 5.5A-B). Rapid FRET fluctuations remained under all the solution conditions tested. Eliminating negative charge from the RecA C-terminus (Lusetti et al., 2003) by deletion of acidic residues resulted in a longer lifetime of the complex between RecA filament and non-homologous dsDNA (Figure 5.6A-B). Overall, our results suggest that the interaction between dsDNA and the RecA filament is likely to be electrostatic.

Previously, we have shown that the outgoing ssDNA, upon completion of strand exchange, exhibits rapid fluctuations in FRET which we attributed to the high entropy state of the outgoing ssDNA still remaining bound to the RecA filament. Since the dominant mode of interaction between the dsDNA and the RecA filament is electrostatic, we hypothesized that non-homologous ssDNA might also interact in the same manner as dsDNA bound to RecA filaments. Indeed, when we added non-homologous ssDNA to a pre-formed RecA filament, we observed multiple binding and dissociation events similar to those observed with non-homologous dsDNA (Figure 5.7A-B). Each binding event was characterized by large fluctuations in FRET with lifetime of the complex between non-homologous ssDNA and the RecA filament ($\Delta t_{\text{non-homologous}}^{\text{ssDNA}}$ = 2.8 sec) was comparable to that using non-homologous dsDNA (Figure 5.2B). Thus, the FRET fluctuations observed as a could represent a general property of interaction between a DNA molecule and the pre-formed RecA filament.

5.2.2 Sliding of RecA filament along dsDNA results in rapid FRET fluctuations

We wanted to address the basis for the large fluctuations in FRET upon docking of dsDNA to the RecA filament. We considered two possible models which could explain our observations related to non-homologous dsDNA interactions with the RecA filament.

The first model involves limited unwinding of the incoming dsDNA (Bianchi et al., 1985) by the RecA filament resulting in separation of the duplex ends. Repeated melting and annealing transitions at the labeled ends of the dsDNA could in principle result in rapid FRET fluctuations. We tested this possibility using a dsDNA (L_{dsDNA}=

39bp) labeled at one duplex end (donor and acceptor at the two opposing strands) so that local melting would cause FRET decrease and vice versa. Upon interaction of the dual labeled non-homologous dsDNA with an unlabeled RecA filament immobilized on the surface, we observed stable high FRET (Figure 5.8A-B) suggesting that separation of the duplex ends is unlikely to contribute to the large fluctuations in FRET that we observed in the previous assay (Figure 5.2A). Thermal fraying of the duplex ends, if it occurs, is likely to be faster than the time resolution of our measurement (30 ms).

A second model involves 1-D sliding of the RecA filament along the dsDNA in which case the observed FRET changes would result from motion of the non-homologous dsDNA relative to the RecA filament. In order to test the sliding model, we measured whether changes in the length of the RecA filament would affect the time scale of FRET fluctuations since such a change would modulate the encounter frequency between the donor and the acceptor fluorophores. In contrast, conformational changes either within the dsDNA or the protein would presumably exhibit the same time scale of FRET fluctuations independent of the RecA filament length.

Upon binding of non-homologous dsDNA We found that shorter lengths of the RecA filament (L_{filament} = 21 nt) exhibited more rapid changes in FRET and of smaller amplitudes (Figure 5.9A) compared to longer lengths of immobilized ssDNA (L_{filament} = 99 nt) which exhibited larger and comparatively slower changes in FRET (Figure 5.9B). Using cross correlation of the donor and acceptor fluorescence intensities to measure the time scale of FRET fluctuations, we tested four different lengths of immobilized ssDNA (L_{filament} = 21 nt, 39 nt, 69 nt and 99 nt) while using the same length of non-homologous dsDNA (L_{dsDNA}= 39bp) for each measurement. We found that the average cross correlation time increases with increasing filament lengths (Figure 5.10A-B) consistent with a model that involves 1-D sliding motion of dsDNA relative to the RecA filament. Furthermore, the histogram of FRET shifted towards the low FRET state with increasing filament lengths (Figure 5.10C).

To estimate the diffusion coefficient of the 1-D sliding process, we performed Monte Carlo simulations of dsDNA diffusing along a RecA filament (Figure 5.10D). For the purposes of our simulation, we treated the RecA filament as a rigid rod given its

large persistence length (persistence length= 784 nm) (Hegner et al., 1999). We then simulated time traces of donor and acceptor intensities of dsDNA bound to the various RecA filament lengths using different pre-assigned diffusion coefficients (D_{slide}) and then calculated the average cross correlation times for each case. By comparing the simulation results with the experimental data we estimated the diffusion constant for 1-D sliding of dsDNA relative to the RecA filament to be ~ 0.9 x 10^{-3} μm^2/sec or 7,700 bp^2/sec.

5.2.3 Three color experiments support sliding during homology search

To further test the hypothesis of 1-D sliding of the RecA filament, we designed a three-color FRET assay. This experimental design involved an immobilized RecA filament labeled with two acceptor fluorophores-Cy5 and Cy7 along the ssDNA embedded within the filament (Figure 5.11A). The two acceptors are separated by 33 nt ssDNA, which upon RecA binding results in a large separation and negligible FRET between them (Joo et al., 2006). The sliding model predicts that motion of the dsDNA relative to the RecA filament would result in anticorrelated changes between the two FRET efficiencies, one between Cy3 and Cy5 and the other between Cy3 and Cy7 because when Cy3 on the dsDNA approaches Cy5, it should move away from Cy7 and vice versa.

Following RecA filament formation on an immobilized ssDNA labeled with the two alternative acceptor fluorophores (Cy5 and Cy7), docking of the donor labeled non-homologous dsDNA to the RecA filament resulted in large and rapid fluctuations in FRET from Cy3 to Cy5 ($E_{Cy3\text{-}Cy5}$) and FRET from Cy3 to Cy7 ($E_{Cy3\text{-}Cy7}$) (Figure 5.11B (bottom panel)). Consistent with the sliding model, the time trace of $E_{Cy3\text{-}Cy5}$ *and* $E_{Cy3\text{-}Cy7}$ exhibits anticorrelation between the two FRET efficiencies supporting the idea that the Cy3 labeled non-homologous dsDNA moves between the two spatially separated acceptor positions. We found the timescale of fluctuations between the two FRET efficiencies ($T_{xcorr\text{-}3color}$= 0.04 s) (Figure 5.12A) in the three color measurement to be similar to that measured using the previous two color assay and an immobilized ssDNA of length 39 nt (Figure 5.10B). The scatter plot of $E_{Cy3\text{-}Cy5}$ versus $E_{Cy3\text{-}Cy7}$ and calculation of Pearson's correlation coefficient for the two variables ($r_{pearson}$= -0.6) provides further evidence for a negative association between $E_{Cy3\text{-}Cy5}$ and $E_{Cy3\text{-}Cy7}$ (Figure 5.12B). Hence, cumulatively the two and three color

FRET results support the hypothesis that the rapid changes in FRET observed in our assays using non-homologous dsDNA reflects the ability of the dsDNA to slide relative to the immobilized RecA filament.

5.3 Discussion

Our study establishes that sliding is a possible mechanism which can assist homology search mediated by RecA. In contrast to models where diffusion in solution alone limits the efficiency of homology search (Adzuma, 1998; Gonda and Radding, 1986), a 1-D sliding mechanism serves to enlarge the target site available to a protein upon each binding trial. Until now, there has been no example of a multi protein complex bound to DNA such as the RecA filament, which can present a vehicle for sliding and target search processes. We estimated the diffusion constant, D_{slide} to be ~ 0.9 x 10^{-3} μm^2/sec or 7700 bp^2/sec. Our measured value implies that under conditions typically used for *in vitro* strand exchange reactions with high Mg^{+2} (10mM Mg^{+2}) and interaction times between the RecA filament and dsDNA (this study) ranging from 0.5-10 sec (Forget and Kowalczykowski, 2012; Mani et al., 2009), sliding would allow homology search to occur over a distance of 60-300bp prior to dissociation.

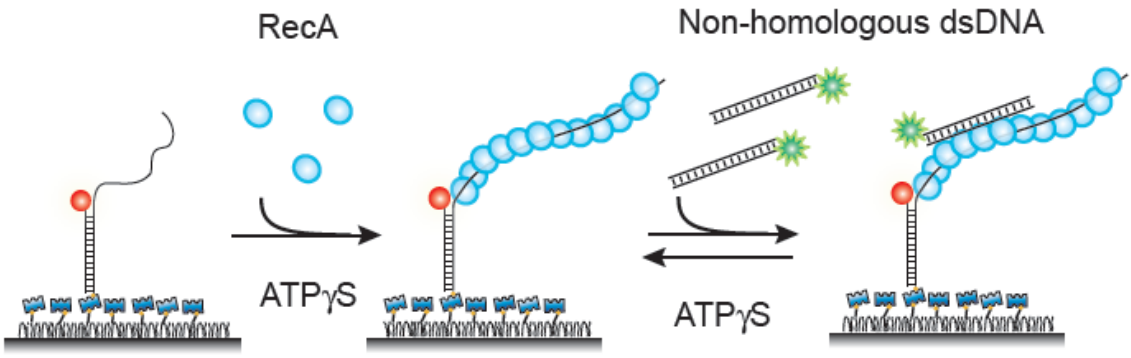
We showed that the interaction between dsDNA and the RecA filament in the absence of stable basepairing and product formation is primarily mediated by electrostatics (Cazaux et al., 1998; Gourves et al., 2001). Given that the RecA secondary binding site plays a critical role during the initial phase of synapsis where it might be involved in recruiting the incoming dsDNA to the RecA filament (Kurumizaka et al., 1996; Kurumizaka et al., 1999), we hypothesize that the RecA secondary binding site may be involved in mediating 1-D sliding. The fact that the secondary binding site is enriched in lysine and arginine residues (Chen et al., 2008) is also consistent with our observation that the dsDNA-RecA filament interaction in the absence of homology is primarily mediated by electrostatics.

It is likely that the same structural features of the RecA filament which enable intersegmental transfer (Forget and Kowalczykowski, 2012) with dsDNA, may also be involved in facilitating homology search via 1-D sliding. Previous reports of the 1-D sliding motion of Rad51 filaments along dsDNA (Graneli et al., 2006) provides a precedent for RecA and its homologs to possess the capability of sliding along DNA.

Here, we demonstrate that the sliding activity we observe can have a functional role in the context of facilitating homology search. The next chapter (Chapter 6) describes how basepairing interactions can occur during sliding of dsDNA relative to the RecA filament.

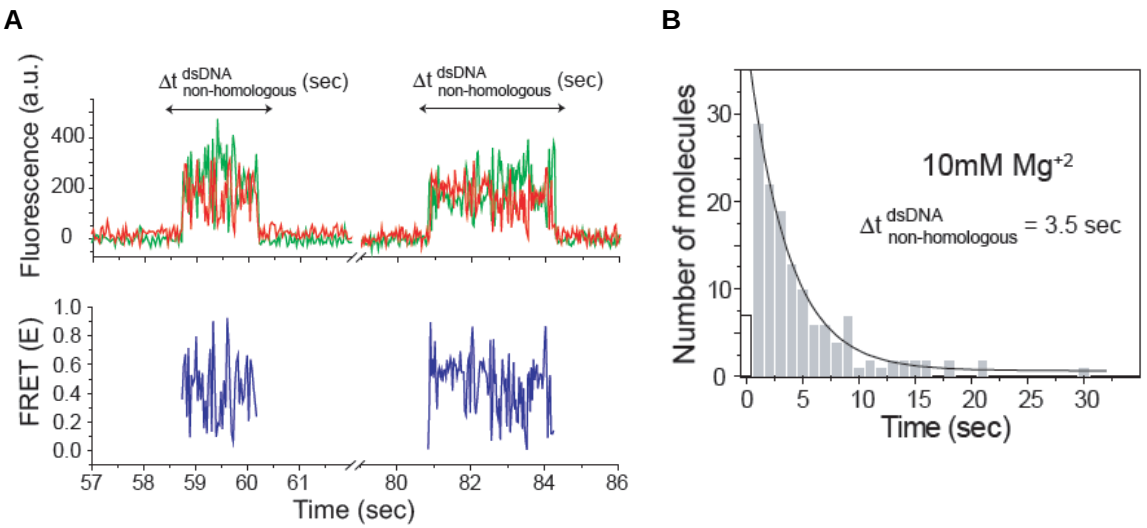
5.4 Figures

Figure 5.1 Single molecule assay to detect interactions between non-homologous dsDNA and RecA filament



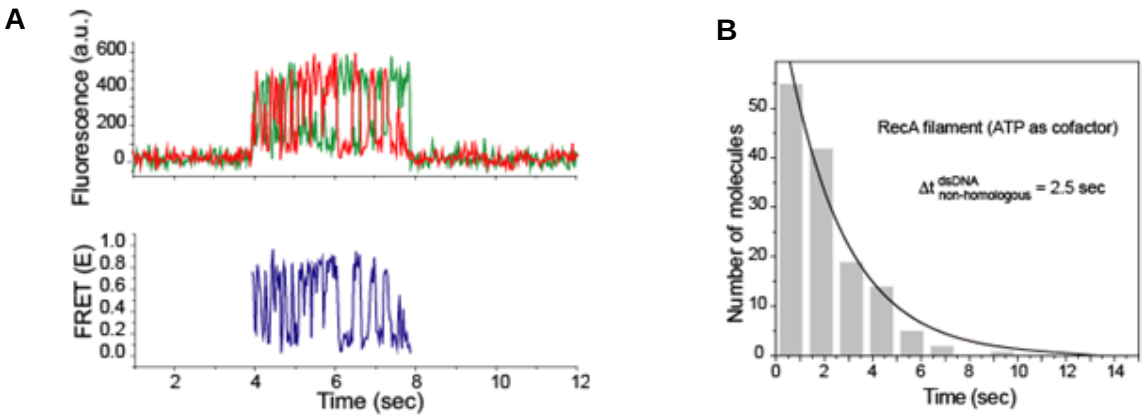
A schematic of the single molecule FRET based assay to detect interactions between RecA filament and non-homologous dsDNA. After RecA filament formation on ssDNA ($L_{filament}$= 39 nt) labeled with an acceptor (red), a non-homologous dsDNA (L_{dsDNA}= 39 bp) labeled with a donor was added. DNA docking results in appearance of donor(green) signal with FRET reporting on the changes in distance

Figure 5.2 Dynamic interactions between RecA filament and non-homologous dsDNA



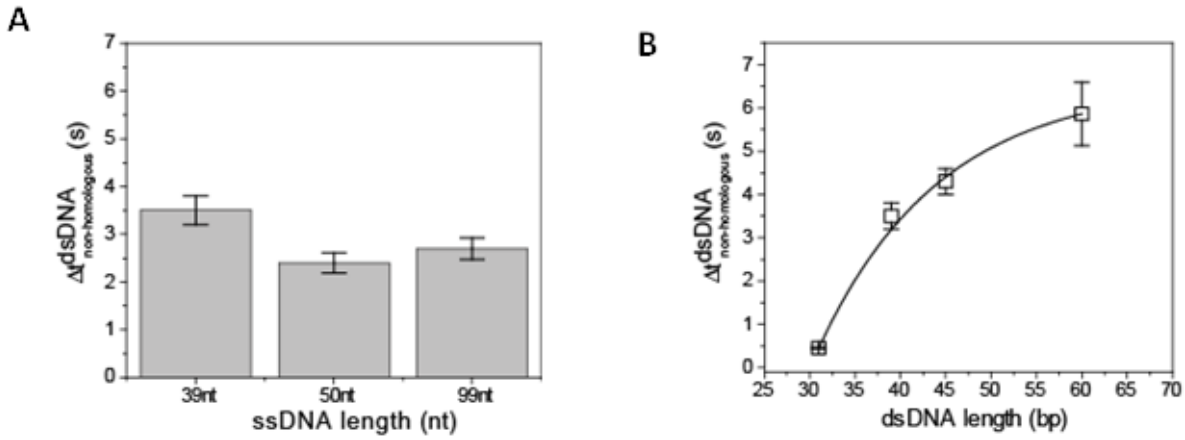
A. Single molecule time traces showing donor (green) and acceptor (red) intensities exhibits rapid FRET fluctuations with multiple binding and dissociation events within a single time trace (top panel). Corresponding FRET time traces (blue) are shown in the bottom panel **B.** Histogram of the duration of the bound state for non-homologous dsDNA ($\Delta t^{\mathrm{dsDNA}}_{\mathrm{non-homologous}}$) and a single exponential decay fit.

Figure 5.3 Non-homologous DNA interactions with the RecA filament are independent of ATP hydrolysis



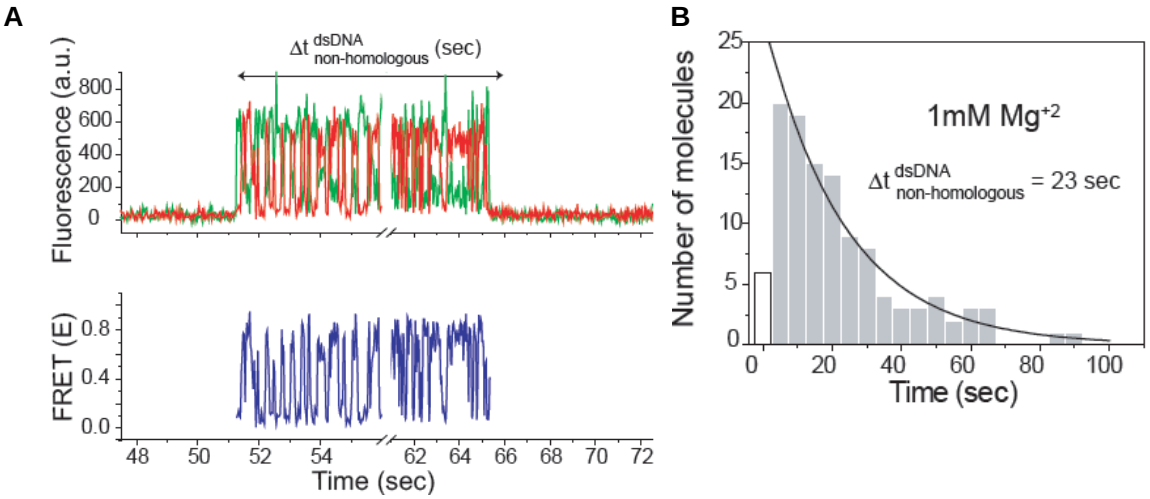
A. Single molecule time trace showing donor (green) and acceptor(red) intensities (top panel) and the corresponding FRET values (bottom panel) **B.** Histogram of the dwell time of DNA docking to the RecA filament and a single exponential decay fit

Figure 5.4 Non-homologous dsDNA interaction with RecA filament: Dependence on filament length, $L_{filament}$ (nt) and dsDNA length, L_{dsDNA} (bp)



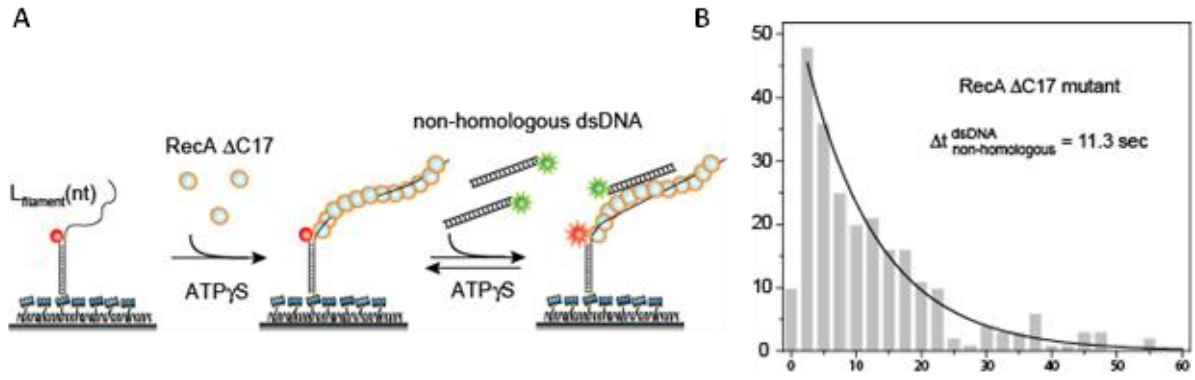
A. Off-time of non-homologous dsDNA binding versus ssDNA length. RecA filament formation was initiated on ssDNA of different lengths followed by addition of non-homologous dsDNA (L_{dsDNA}= 39bp). Error bars are standard errors obtained from single exponential decay fitting. **B.** Off time of non-homologous dsDNA versus dsDNA length. RecA filament formation was initiated on ssDNA ($L_{filament}$= 39nt) followed by addition of non-homologous dsDNA (L_{dsDNA}= 31, 39, 45 and 60bp). Error bars are standard errors obtained from single exponential decay fitting. Dwell time for L_{dsDNA}= 60bp represents the mean lifetime of all binding events. Exponential fit was used as a guide.

Figure 5.5 Binding time dependence on magnesium concentration



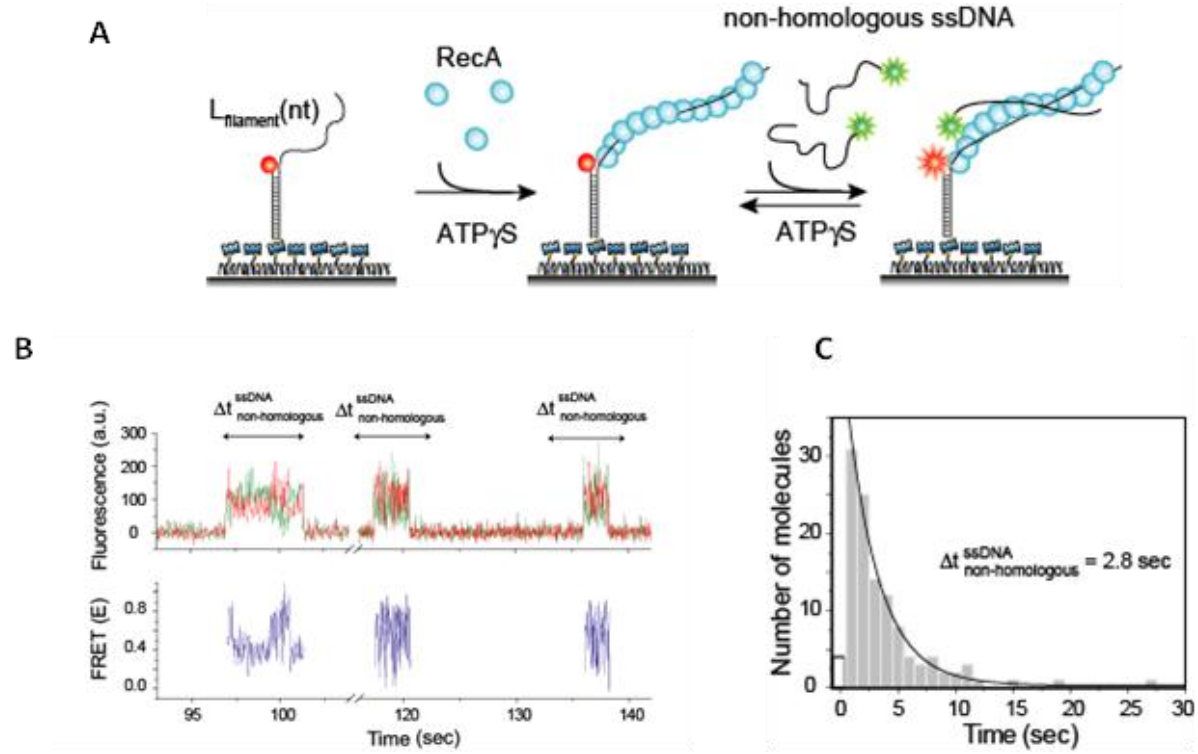
A. Single molecule time traces showing donor (green) and acceptor (red) intensities exhibits rapid FRET fluctuations with multiple binding and dissociation events within a single time trace (top panel). Corresponding FRET time traces (blue) are shown in the bottom panel **B.** Histogram of the duration of the bound state for non-homologous dsDNA ($\Delta t_{\mathrm{non-homologous}}^{\mathrm{dsDNA}}$) and a single exponential decay fit with 1mM Mg^{+2} in solution.

Figure 5.6 Deletion of acidic residues enhances RecA affinity for non-homologous dsDNA



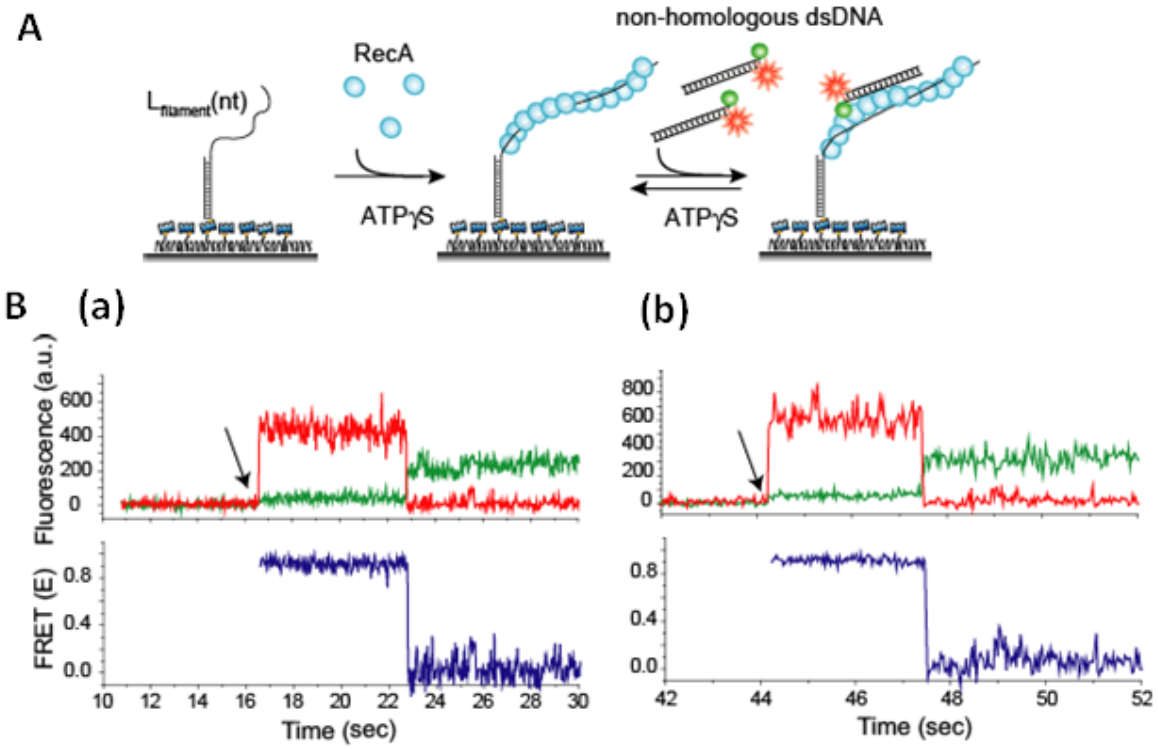
A. A ssDNA ($L_{filament}$= 39nt) with an acceptor fluorophore (Cy5) at the junction was immobilized. After filament formation using a C terminal deletion mutant of RecA in the presence of ATPɣS as a co-factor, non-homologous dsDNA with a donor fluorophore (Cy3) was added **B.** Histogram of the dwell time of DNA docking to the RecA filament and a single exponential decay fit

Figure 5.7 Non-homologous ssDNA also exhibits sliding dynamics upon interaction with RecA filament



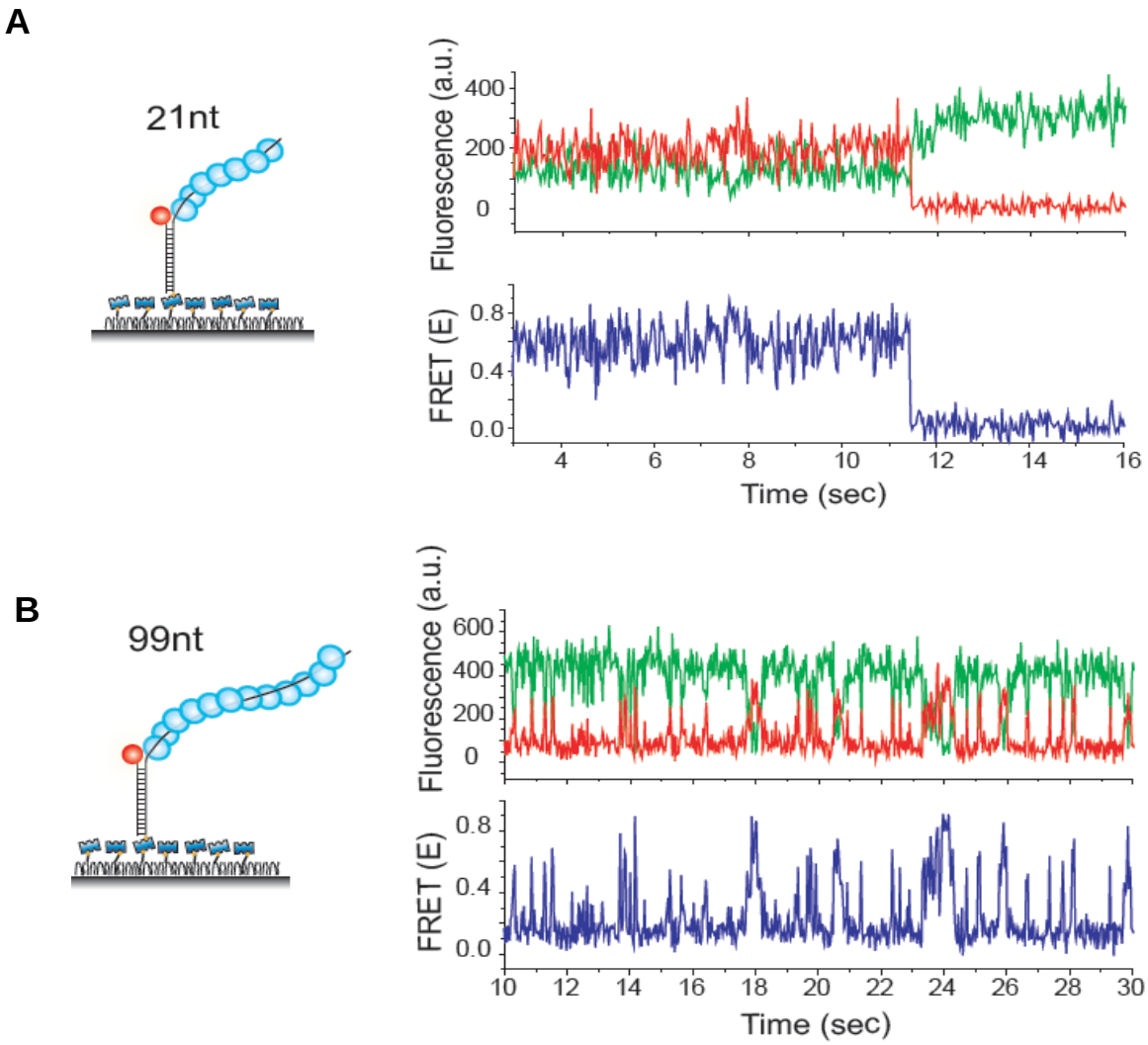
A. A ssDNA ($L_{filament}$= 39nt) with an acceptor fluorophore at the junction was immobilized. After RecA filament formation in the presence of ATPɣS as a co-factor, non-homologous ssDNA labeled with both a donor (Cy3) was added. **B.** Single molecule time trace showing donor (green) and acceptor(red) intensities (top panel) and the corresponding FRET values (bottom panel). Docking events of ssDNA to the RecA filament are reversible and multiple events can be observed in a single time trace. C. Histogram of binding events of non-homologous ssDNA and a single exponential decay fit.

Figure 5.8 Thermal breathing of DNA ends does not contribute to the observed fluctuations in FRET



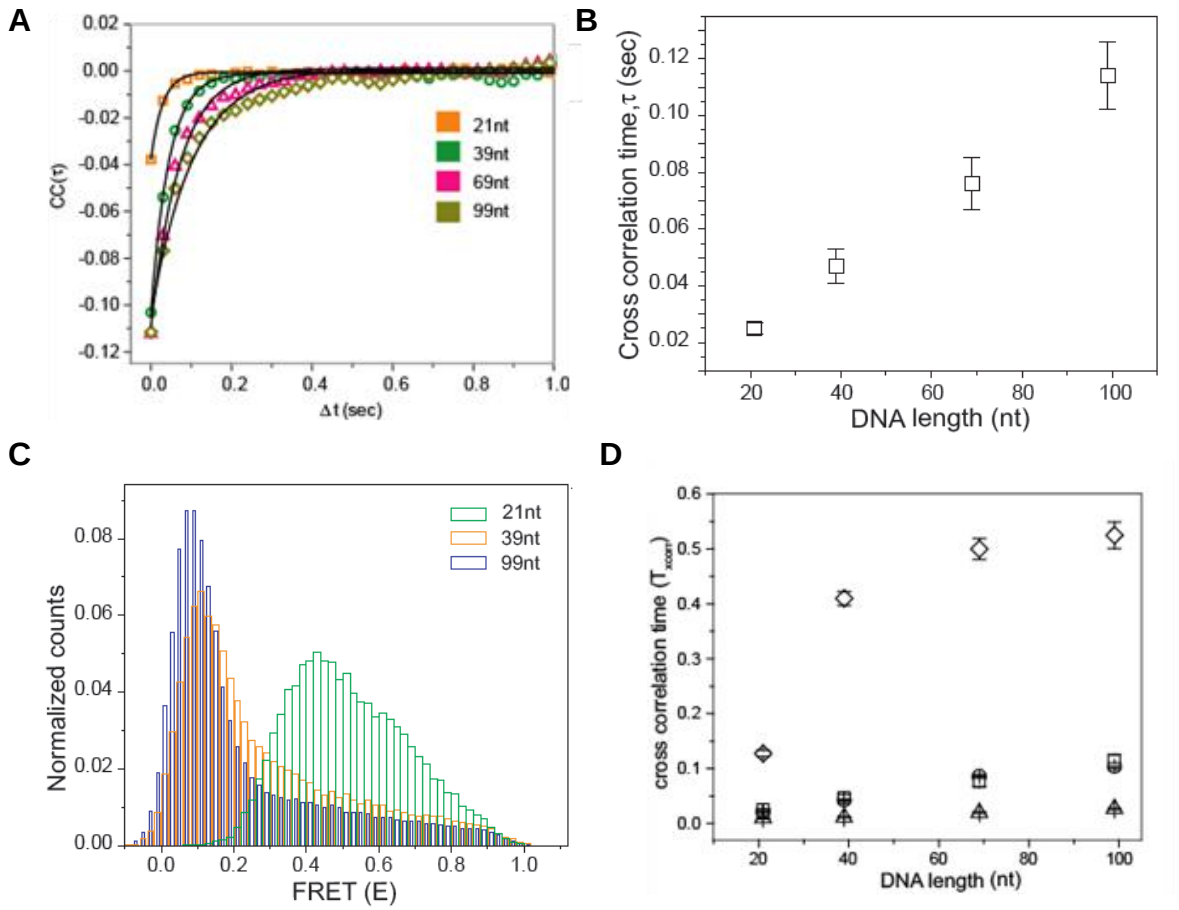
A. A ssDNA ($L_{filament}$= 39nt) with an no fluorophore was immobilized. After RecA filament formation in the presence of ATPɣS as a co-factor, non-homologous dsDNA labeled with both a donor (Cy3) and acceptor (Cy5) fluorophores was added. The labeling scheme was such that the appearance of high FRET indicates the intact nature of the dsDNA. **B.** (a) and (b) Single molecule time trace showing donor (green) and acceptor(red) intensities (top panel) and the corresponding FRET values (bottom panel). Arrows represent the time at which non-homologous dsDNA docking occurs. FRET remains high (~0.9) until photobleaching of the acceptor.

Figure 5.9 Dependence of filament length on time scale of FRET fluctuation



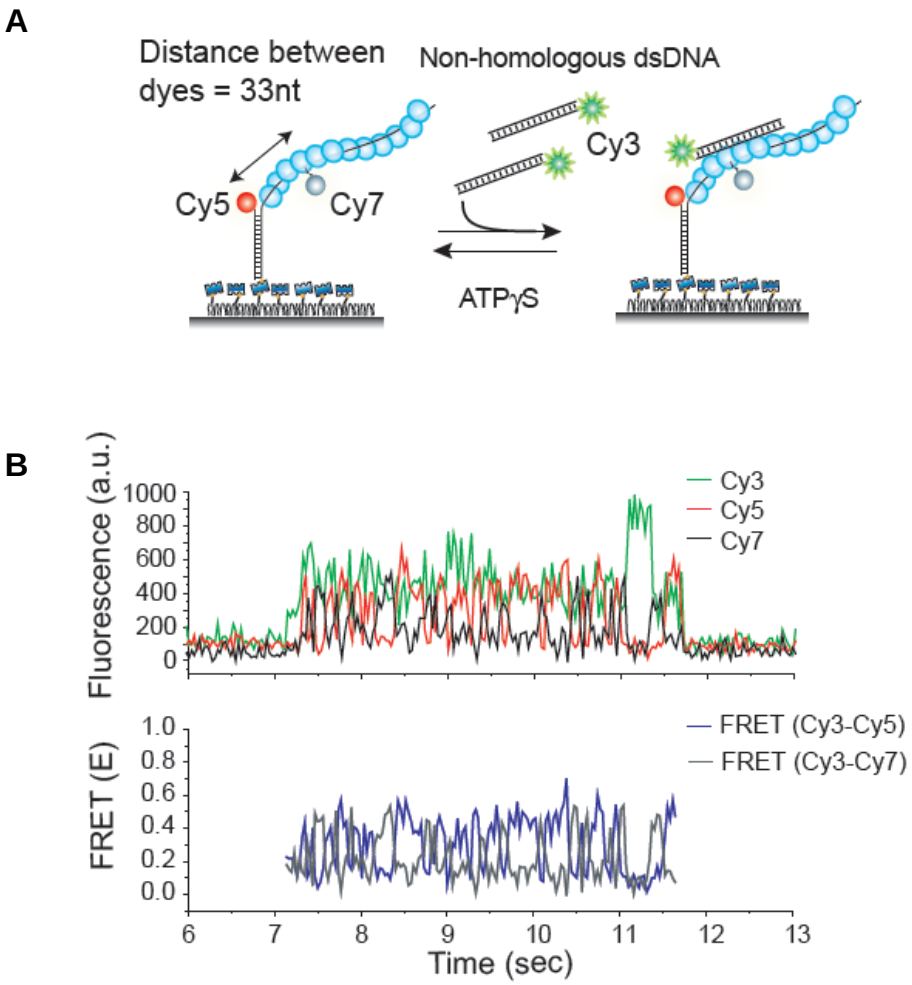
A. Single molecule traces showing donor(green) and acceptor(red) intensities (top panel) upon docking of non-homologous dsDNA (L_{dsDNA}= 39bp) to a RecA filament assembled on a ssDNA overhang, $L_{filament}$= 21nt. Corresponding FRET time traces (blue) are shown in the bottom panel. **B.** Same as A, except that the RecA filament is assembled on a ssDNA, $L_{filament}$= 99nt.

Figure 5.10 Cross correlation analysis of FRET fluctuation and Monte Carlo simulation



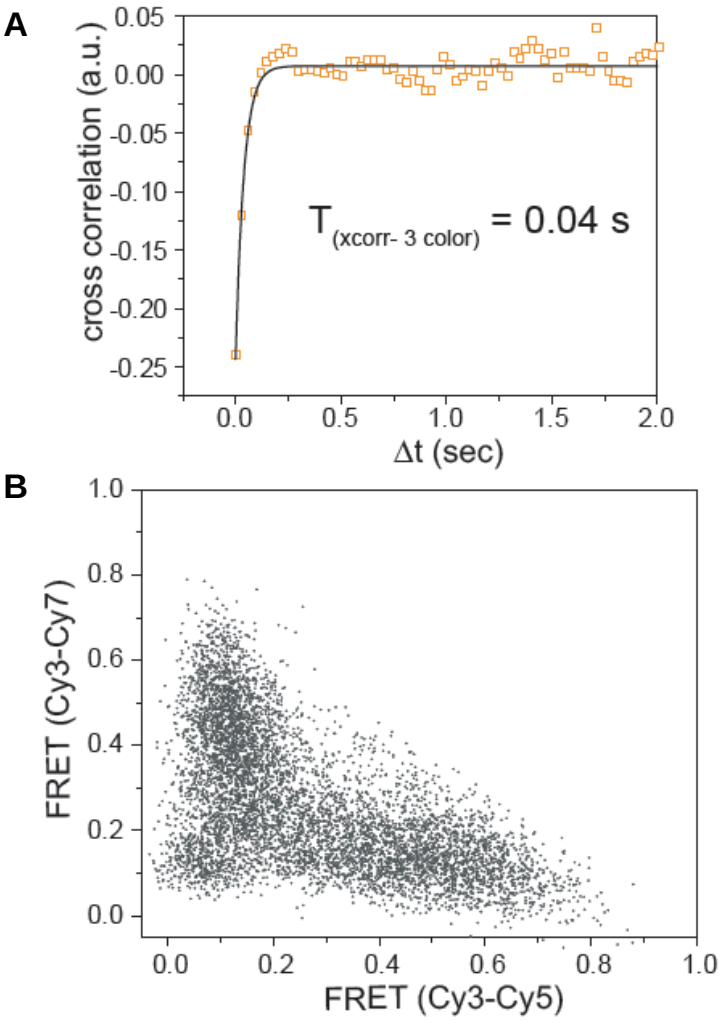
A. Single exponential fitting of average cross correlation curves for various filament lengths **B.** Average cross correlation time, versus $L_{filament}$. Error bars are standard errors of the mean determined from three independent datasets. **C.** FRET efficiency (*E*) histograms of single molecule traces for individual docking events of non-homologous dsDNA to RecA filaments assembled on ssDNA, $L_{filament}$= 21nt, 39nt and 99nt. **D.** Monte Carlo simulations of RecA filament sliding were carried out as described in the *Methods* section(□ experimental data (same as Figure 5.10B), simulation with diffusion coefficient 0.9×10^{-3} μm^2/sec (◯), simulation with diffusion coefficient 0.09×10^{-3} μm^2/sec (△) and simulation with diffusion coefficient 9×10^{-3} μm^2/sec (◇).

Figure 5.11 Three color FRET observations support RecA filament sliding



A. A schematic of the single molecule three color FRET assay to measure RecA filament sliding. ssDNA ($L_{filament}$= 99nt) labeled with two acceptor fluorophores (Cy5-red and Cy7-black) with a separation of 33nt between the fluorophores, was immobilized on the surface. Upon docking of non-homologous donor (Cy3) labeled dsDNA to the pre-formed RecA filament formation, sliding predicts anticorrelated emissions between the two acceptors. **B.** Single molecule time traces of Cy3(green), Cy5(red) and Cy7(black) intensities (top panel). Corresponding FRET time traces of FRET between Cy3 and Cy5 ($E_{Cy3\text{-}Cy5}$-blue) and FRET between Cy3 and Cy7 ($E_{Cy3\text{-}Cy7}$-grey).

Figure 5.12 Analysis of three color FRET measurements



A. Normalized cross correlation plot of $E_{Cy3\text{-}Cy5}$ and $E_{Cy3\text{-}Cy7}$ averaged over 30 molecules and a single exponential fit of the data is overlaid (black). **B.** Scatter plot of $E_{Cy3\text{-}Cy5}$ and $E_{Cy3\text{-}Cy7}$ for 30 molecules showing unique high FRET regions along both axes.

Chapter 6

Homology recognition during sliding

If you can't go home, there is nowhere to go, and nowhere is the biggest place in the world-indeed, nowhere is the world.

Aleksandar Hemon (2008)

6.1 Introduction to basepairing within RecA synaptic complex

The mechanism of basepair recognition during RecA mediated homology search has remained enigmatic despite several efforts to understand how RecA senses homology. Successful recognition of homology leads to the formation of a synaptic complex between a homologous dsDNA and a RecA filament. There are two models proposed for basepair recognition 1) a triplex intermediate structure involving RecA-ssDNA invasion at the major groove of the dsDNA requiring non Watson-Crick base contacts to recognize homology (Chiu et al., 1993; Hsieh et al., 1990) 2) Melting-annealing transitions which lead to homology recognition via Watson-Crick basepairing interactions (Adzuma, 1992). The key difference between the two models is whether initial melting of the duplex is a pre-requisite for recognition and basepairing.

Chemical crosslinking methods to detect the position of the outgoing ssDNA were the commonly used approach to propose models for homology recognition and basepairing within the synaptic complex. The position of the outgoing ssDNA following heteroduplex product formation in turn places constraints on the topology of the initial interaction between the RecA filament and dsDNA. Interaction with the major groove would result in the released ssDNA positioned at the minor groove of the heteroduplex (Malkov et al., 2000) while initial interaction with the minor groove (or rather the absence of a major groove interaction) places the outgoing ssDNA at the major groove of the final heteroduplex product (Baliga et al., 1995; Podyminogin et al., 1995; Zhou and Adzuma, 1997). The conclusion regarding the mechanism of homology recognition based on the read out of outgoing ssDNA position within the heteroduplex product remains ambiguous.

Studies based on NMR measurements of the RecA filament in complex with ssDNA suggested that RecA induces sugar-base stacking interactions within the bound DNA wherein deoxyribose sugars contribute a methylene moiety that stabilizes adjacent nucleotides (Nishinaka et al., 1997). A change in the conformation of the sugar (via sugar puckering) causes basepairs within the dsDNA to flip outward which primes the DNA for basepair interactions with the complementary ssDNA within the RecA filament (Nishinaka et al., 1998). A model where basepairs flip outward to test for pairing interactions with the ssDNA offers partial support for the second model based on melting annealing transitions within the dsDNA as the mechanism that drives homology recognition and basepairing.

The experimental scheme described in this chapter allows us to observe homology recognition and basepairing in real time and this assay may serve as a platform for asking future questions regarding the nature of the basepairing intermediate formed between RecA-ssDNA and a target dsDNA.

6.2 Results

6.2.1 Observing homology recognition and basepairing

Can the sliding of RecA filament on dsDNA be a physiologically relevant activity, that is, can the RecA filament recognize a homologous sequence during sliding? We tested whether sliding might facilitate recognition and basepairing at a short homology site located within a target dsDNA. We embedded two repeats of a homologous sequence at positions HS1 and HS2 within an otherwise non-homologous ssDNA. If the dsDNA containing a sequence complementary to the repeat slides back and forth between the two homologous sites (HS1 and HS2) present within the RecA bound ssDNA, it would support the proposal that homology search may be facilitated by 1-D sliding (Figure 6.1). Both HS1 and HS2 were complementary to a sequence in close proximity to the donor labeled end of the incoming dsDNA and the location of and spacing between the sites were chosen to be within a FRET sensitive regime (~ 20-80Å) relative to the acceptor fluorophore at the junction.

We first confirmed that dsDNA docking to a RecA filament bound to a ssDNA homopolymer sequence (poly T, $L_{filament}$= 50 nt) preserved the large and rapid FRET

fluctuations as previously observed (Figure 6.2). We then inserted two identical 6 nt sequences (L_h= 6 nt) at HS1 and HS2 resulted in the appearance of distinct FRET states in the single molecule time traces (Figure 6.3A) and a histogram of molecules exhibiting FRET displayed distinct peaks in contrast to the broad distribution observed using poly T DNA sequence (Figure 6.3B). We used a statistical approach based on Hidden Markov Model (HMM) analysis to make unbiased assignments of the various FRET states in the single molecule time traces (Joo et al., 2006). The resulting FRET transitions as obtained from HMM analysis of 236 molecules exhibiting 13020 transitions were plotted in the form of a transition density plot (TDP) which is a two dimensional histogram reflecting the frequency of transitions between different FRET states (Figure 6.4A-B). The TDP displays the presence of transitions between, three distinct FRET states, *E* ~ 0.1, 0.5 and 0.9. In order to confirm that the observed FRET states arise from basepairing and recognition at HS1 and HS2, we performed measurements where the presence of only one homology site- either HS1 or HS2- leads to the appearance of a distinct FRET peak at either ~0.9 or ~0.5 respectively (Figure 6.5). Hence, we assigned the highest FRET state (~0.9) observed in the TDP (Figure 6.4B) to basepairing and recognition at HS1 and the next FRET state (~0.5) to basepairing and recognition at HS2. The lowest FRET state, *E*~ 0.2 (NH) is likely to correspond to a location of the dsDNA along the RecA filament outside the boundary of HS1 and HS2 in a FRET insensitive region indicating lack of recognition and basepairing at either of the two specific homology sites.

From the TDP, we can extract the transition rates between the various FRET states (Figure 6.4C). The use of identical sequences at HS1 and HS2 results in comparable dissociation rates ranging between 0.7 s^{-1}–1.4 s^{-1}. In contrast, the lack of homology when the dsDNA dwells in the low FRET state (NH) leads to consistently faster transition rates from NH to either HS1 or HS2 (2.8 s^{-1}–4 s^{-1}). Surprisingly, transitions between NH and the two homology sites, HS1 and HS2 are independent of each other indicating that sliding may involve events where the dsDNA can slide across a regions of homology (HS2) without engaging in recognition and basepairing. Transitions could also be observed between HS1 and HS2 without NH as intermediate indicating successful basepair recognition during transitions between the two homology sites. Homology recognition and basepairing at either HS1 or HS2

appears to be a stochastic process and we can now estimate the recognition probability associated with a single 6 bp homology sequence embedded within the target dsDNA. The conditional probability that a dsDNA present initially at NH slides to the HS1 position without basepairing at HS2 is given by,

$$P_{\text{recognition}}^{\text{NH->HS1}} = \frac{N^{NH->\text{HS1}}}{\left(N^{NH->\text{HS2}} + N^{\text{NH}->\text{HS1}}\right)}$$

$N^{\text{NH->HS1}}$ and $N^{\text{NH->HS2}}$ denote the number of transitions from NH to HS1 and HS2 respectively. The conditional probability of recognizing HS1 given the molecule is initially present at NH, $P_{\text{recognition}}^{\text{NH->HS1}}$ is 45%. Similarly, we can calculate the conditional probabilities associated with the recognition of a homology site given the DNA is initially bound at HS1 (48%) or HS2 (47%). We can perform the same analysis on DNA constructs with only one homology site (either HS1 or HS2). The TDP for single homology site constructs indicates the presence of peaks consistent with the FRET values corresponding to HS1 and HS2 (Figure 6.6). The same asymmetry in transition rates was observed where the rate of transition from NH to a homology site (either HS1 or HS2) was faster than the transition rate from HS1 or HS2 to NH due to basepairing at that location (Figure 6.7).

6.2.2 Dependence of sliding on homology length

Having shown that 6bp homology is sufficient to act as a molecular brake for the observed sliding behavior, we tested the effect of changing the length of homology at HS1 and HS2. We tested the effect of L_h= 5nt and 8nt at HS1 and HS2. The presence of 5bp homology at the two sites did not appear to perturb the sliding behavior significantly relative to the case with a homopolymer (poly T) resulting in fast FRET fluctuations and a broad FRET histogram which does not exhibit the presence of distinct FRET peaks (Figure 6.8A). In contrast 8bp, results in greater stability of basepairing at HS1 and HS2 resulting in distinct peaks as seen in the case of the 6bp homology DNA construct (Figure 6.8B). However, in contrast to the case where the homology length was 6bp, the presence of 8bp of homology results in longer dwells and rare transitions between different states precluding accurate analysis of the transition rates. Furthermore, the fraction of molecules exhibiting sliding increased with a decrease in homology length at each site consistent with the

fact that increase in number of complementary basepairs contributes to increased stability at each homology site (Figure 6.9).

6.2.3 Orientation of dsDNA binding to the RecA filament

Single molecule methods are useful to detect binding orientations of protein-DNA complexes (Rasnik et al., 2004). We tested whether labeling the 5' end of the two complementary strands of the duplex (plus strand and minus strand) would result in similar scale of FRET fluctuations upon docking of DNA to the RecA filament (Figure 6.10). In the event that there is no preferred binding orientation, both labeling positions must exhibit the same type of FRET dynamics.

The binding of dsDNA to the RecA filament bound to a poly T (50nt) ssDNA was detected as large and rapid FRET fluctuations and labeling of either strand with Cy3 at the 5' position resulted in virtually identical FRET histograms indicating that there is no preference for DNA binding orientation. We then tested whether the same symmetry in binding is also exhibited upon 3' end labeling of the dsDNA. The binding of 3' end labeled DNA resulted in rapid FRET fluctuations and labeling either the 3' end of the plus or minus strand resulted in similar FRET distributions.

Surprisingly, the distribution of FRET values for 3'end labeled DNA was different from that of 5' end labeled dsDNA. Interestingly, labeling of the 3' end resulted in a larger fraction of high FRET in the resulting histogram (compared to 5'end labeling) implying that 3'end labeling position was in closer proximity to the acceptor dye. While overall DNA binding displays symmetry, there is a preferential association of RecA filament with the strand which is in parallel conformation relative to the DNA embedded within the RecA filament. This preliminary data further suggests that some intrinsic feature of the RecA filament bound to ssDNA could dictate the orientation preference of dsDNA to bind to the RecA filament.

6.3 Discussion

The aforementioned data strongly supports the idea that basepairing interactions can occur during sliding. The extended structure of dsDNA upon interaction with the RecA filament introduces sugar-base stacking interactions that allows basepairs within the dsDNA to flip outward (Shibata et al., 2001) permitting frequent testing for homology during the 1-D sliding process. We attribute a physiologically relevant role

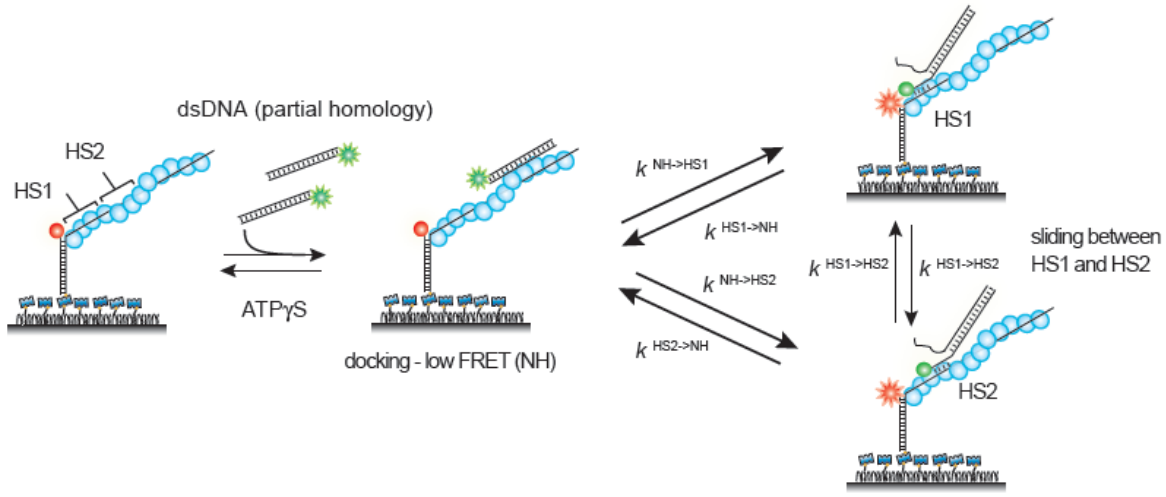
to the 1-D sliding process by demonstrating the capability of RecA filaments to engage in homology recognition and basepairing. Furthermore, we now show that 6 bp of homology is sufficient to represent the minimum recognition unit for RecA mediated homology search thus revising previous estimates of 8bp being the minimum unit of homology recognition (Hsieh et al., 1992). In the absence of additional homology beyond 6bp, the synaptic complex will quickly disintegrate (after ~1 sec) and the search for homology continues until full dissociation of dsDNA from RecA filament.

The mechanisms for target search mediated by proteins are not mutually exclusive and can involve combinations of sliding and 3-D diffusion based mechanisms (Gorman and Greene, 2008). We propose that while intersegmental transfer serves to bring non-contiguous segments of dsDNA close to each other, 1-D sliding during homology search could help RecA reorganize the initial synaptic complex in *cis* without the need to fully dissociate and rebind to a nearby homology site located within a distance of a few hundred basepairs. Given the ability of RecA to slide between nearby homology sites (Figure 6.3), sliding could serve as a mechanism to rapidly scan neighboring sequences for the existence of an optimal seed sequence from which basepair propagation and heteroduplex extension reactions can proceed.

The asymmetry in FRET between 5’ and 3’ labeling positions suggests that the polarity of the ssDNA bound by RecA may dictate the binding orientation of a dsDNA. RecA filament might have a preferential interaction with the strand having polarity which is parallel to the immobilized ssDNA. The strand with parallel polarity to the immobilized ssDNA is the strand that will eventually be the outgoing ssDNA after synapse formation at a homologous site. An asymmetric interaction with one of the two strands in a duplex has previously reported in the case of a viral packaging motor (Φ29) (Aathavan et al., 2009) where charge neutralization along 5’-3’ strand stalls dsDNA packaging while the same perturbation along the 3’-5’ strand has no significant effect on DNA packaging capabilities. One possible explanation is that the asymmetric interaction between the two strands of the dsDNA in the case of RecA may serve to stabilize the ‘soon-to-be-released’ outgoing ssDNA. This may result in the complementary DNA (incoming ssDNA) being available for basepairing interactions with the reference ssDNA embedded within the RecA filament.

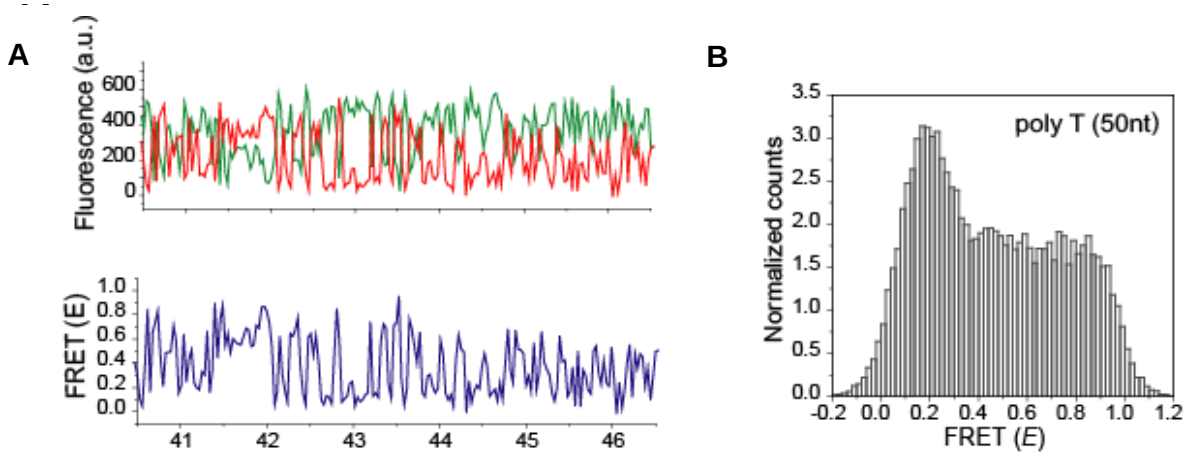
6.4 Figures

Figure 6.1 Testing basepairing during RecA filament sliding



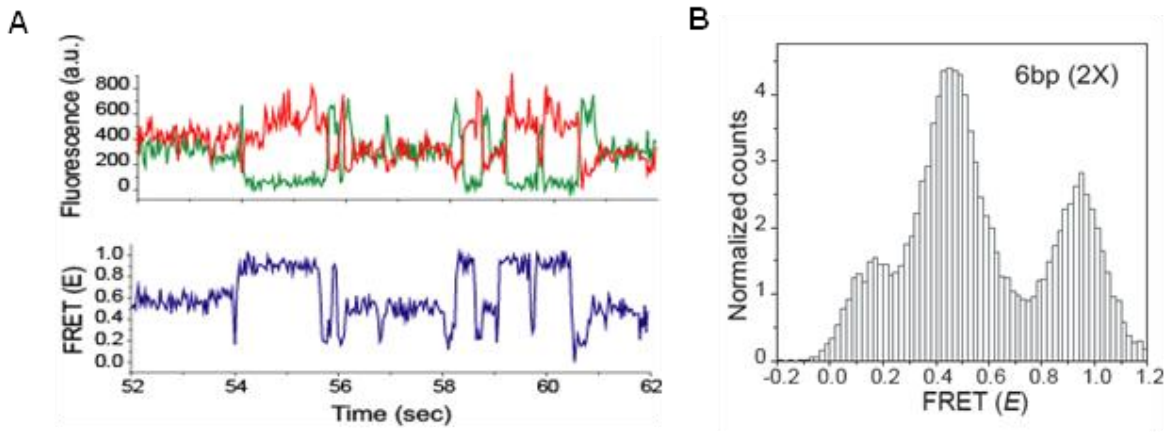
A schematic of the single molecule FRET based assay to detect homology recognition between RecA filament and dsDNA. After RecA filament formation on ssDNA ($L_{filament}$= 50 nt) labeled with an acceptor (red), a dsDNA (L_{dsDNA}= 39 bp) labeled with a donor was added. Docking at an arbitrary location along the RecA filament (NH) results in low FRET. Recognition of homology site 1 (HS1) or homology site 2 (HS2) results in the appearance distinct FRET states depending on their relative distances from the acceptor. Transitions between the three states are shown with their corresponding rates.

Figure 6.2 Non-homologous dsDNA interaction with a poly T ($L_{filament}$= 50nt) coated by RecA



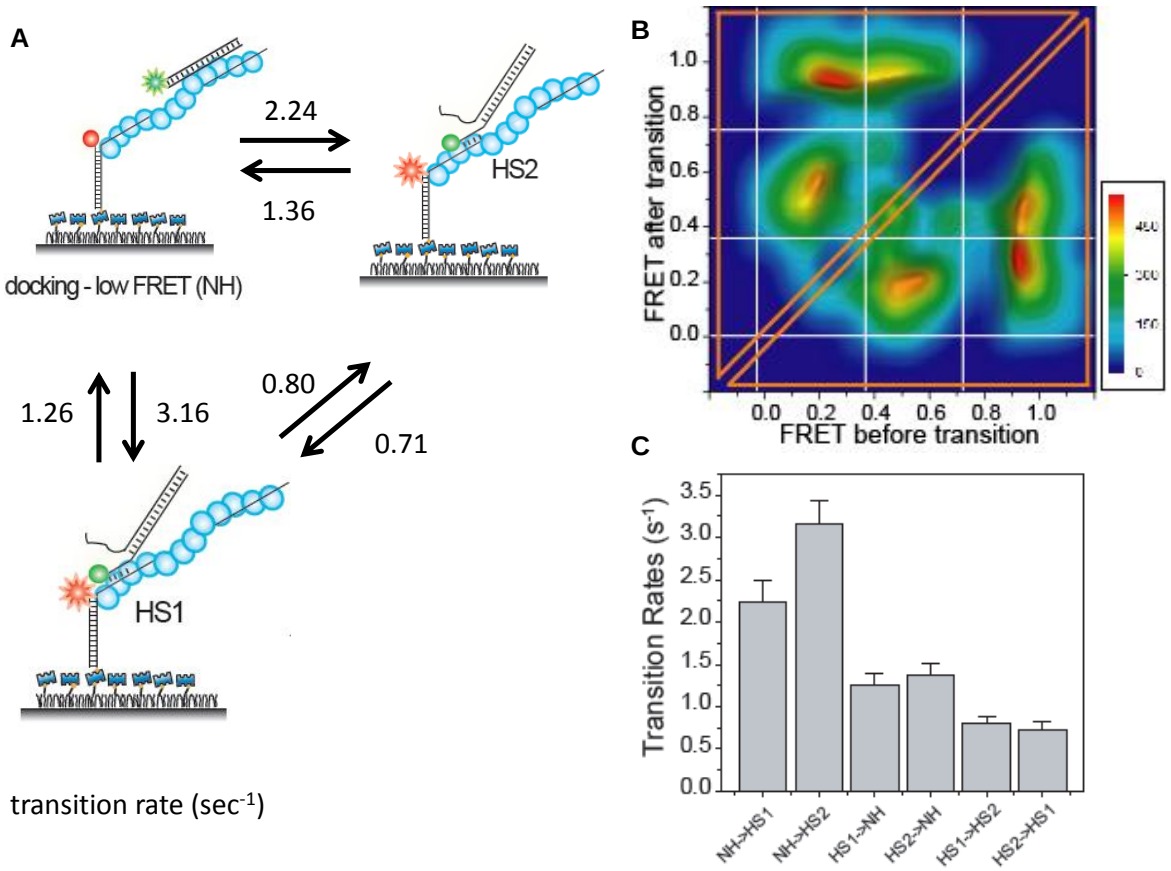
A. A poly T ssDNA ($L_{filament}$ = 50nt) with an acceptor fluorophore (Cy5) at the junction was immobilized. After RecA filament formation in the presence of ATPγS as a co-factor, non-homologous dsDNA with a donor fluorophore (Cy3) was added. Single molecule time trace showing donor (green) and acceptor(red) intensities (top panel) and the corresponding FRET values (bottom panel) **B**. Histogram of the all single molecule time traces exhibiting FRET displays a broad distribution

Figure 6.3 Non-homologous dsDNA interaction with a HS1 and HS2 (L_h= 6nt) coated by RecA



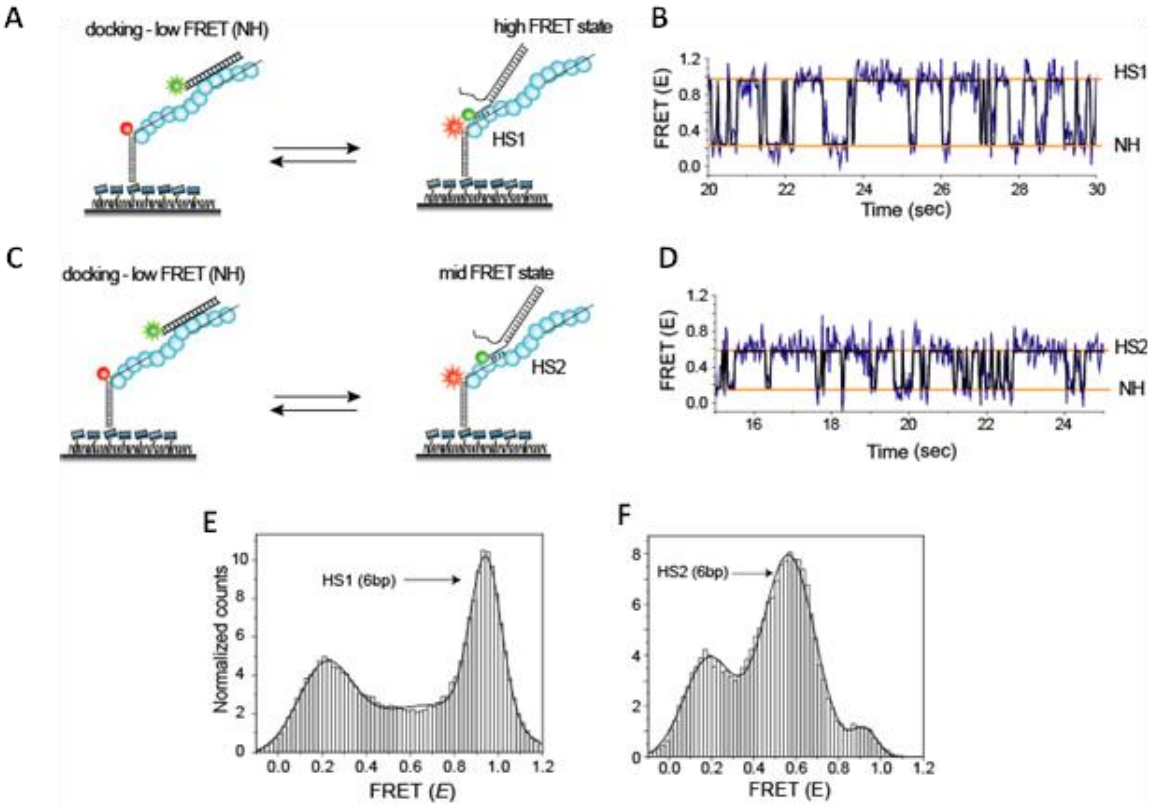
A. Single molecule time traces showing donor (green), acceptor (red) intensities (top panel) and the corresponding FRET trace (blue) (bottom panel) for immobilized ssDNA with two identical 6bp homology sequences at HS1 and HS2 in a poly T sequence background. Idealized time trajectory obtained from HMM analysis is overlaid (black). B. Normalized histograms of time traces exhibiting FRET for docking of dsDNA to immobilized ssDNA with two identical 6bp homology sequences (HS1 and HS2)

Figure 6.4 Testing basepairing during RecA filament sliding



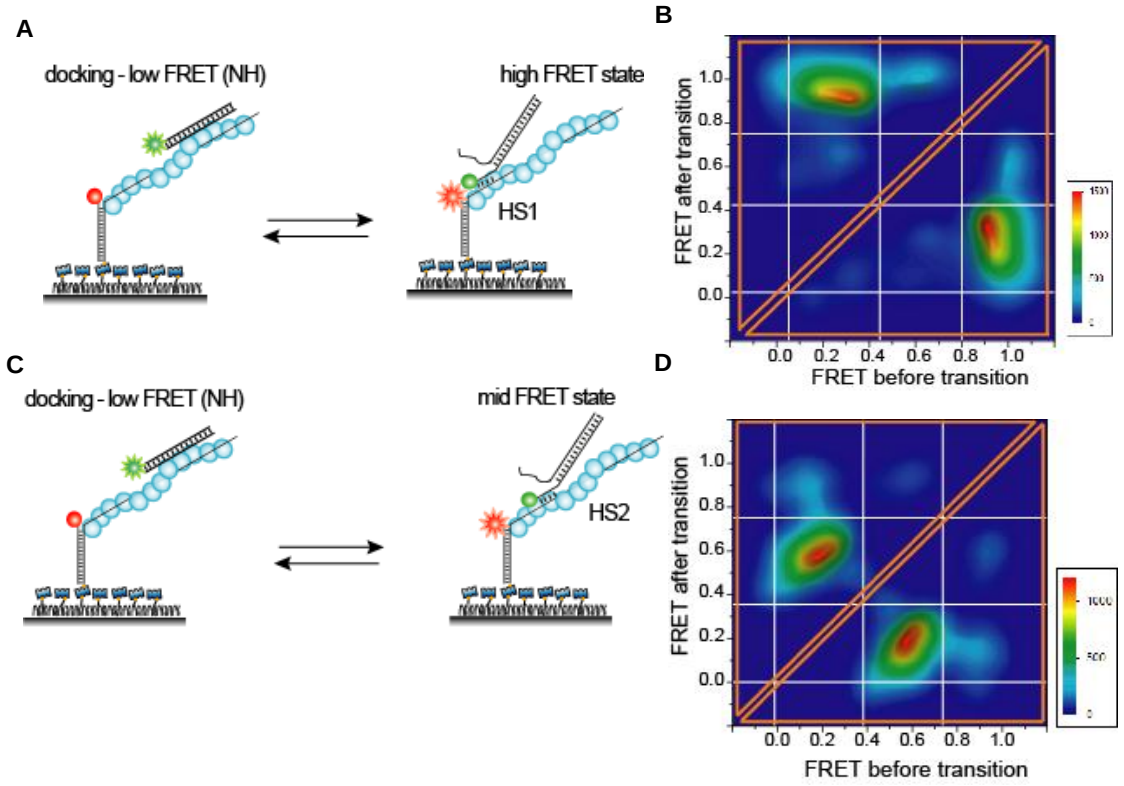
A. Schematic of the transitions between different basepairing states with numbers denoting the transition rates from one state to another **B.** Transition density plot (TDP) of all FRET transitions for immobilized ssDNA with two 6bp repeat sequences (HS1 and HS2) from 236 molecules exhibiting 13020 transitions. **C.** Forward and reverse transition rates between HS1,HS2 and NH states are displayed. Error bars denote standard errors of the mean measured for each set of transitions.

Figure 6.5 Effect of a single 6bp homology site (*HS1or HS2*) in a poly T sequence background



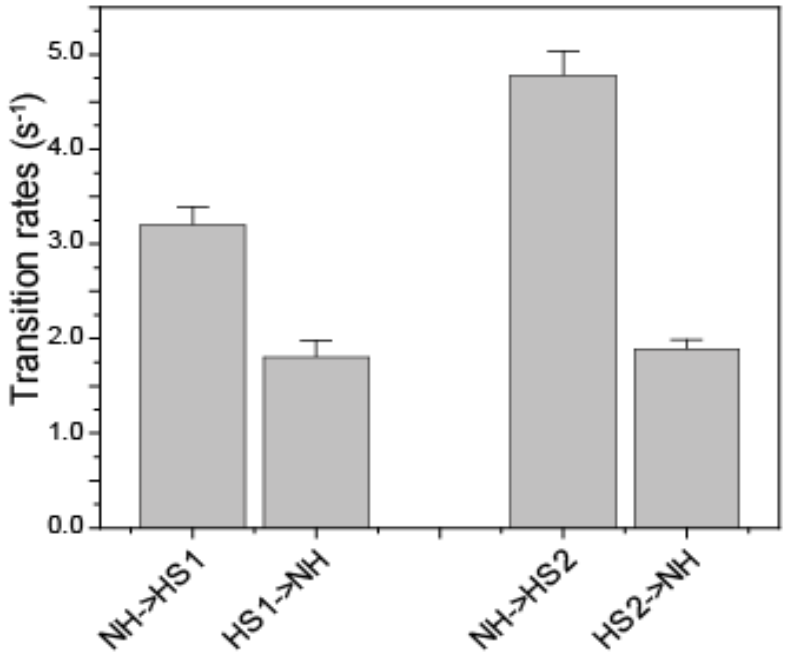
A. A ssDNA with a single 6bp homology site HS1 ($L_{filament}$= 50nt) with an acceptor fluorophore (Cy5) at the junction was immobilized. Homology recognition and basepairing at HS1 would result in high FRET due to the close proximity of the donor and acceptor dyes **B**. Single molecule time indicate FRET transitions (blue) predominantly between high and low FRET states with an overlay of the idealized time trajectories (black) via HMM analysis. **C.** Same as A, except that homology recognition and basepairing at HS2 would result in FRET lower than HS1 due to its distal location from the acceptor fluorophore. **D**. Same as B except that transitions are between medium and low FRET states with an overlay of the idealized time trajectories (black) via HMM analysis. **E**. Histogram of single molecule time traces exhibiting FRET displays the presence of a distinct high FRET state due to the presence of a single 6bp homology site (HS1) in close proximity to the acceptor. **F**. Same as F, with FRET histogram displaying a distinct mid FRET state due to the presence of a single 6bp homology site (HS2) at a position distal to the acceptor.

Figure 6.6 Transition density plot for single 6bp homology ssDNA (either HS1 or HS2)



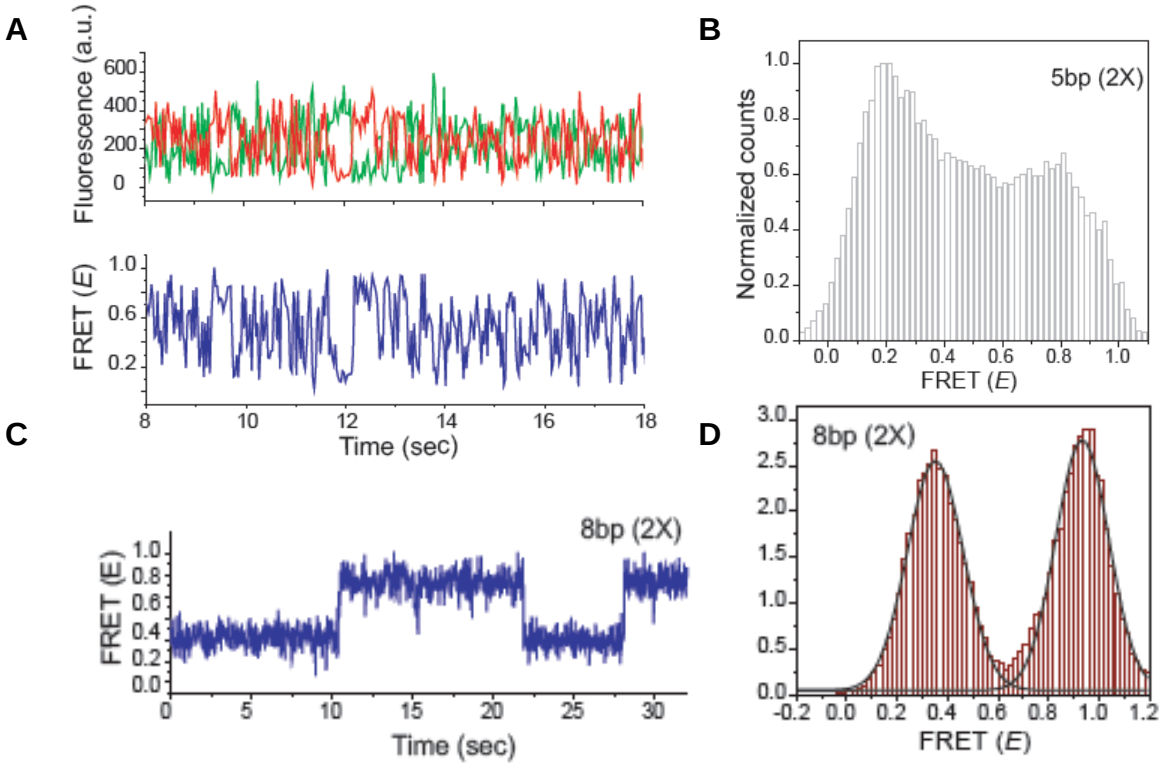
A. A ssDNA with a single 6bp homology site HS1 ($L_{filament}$= 50nt) with an acceptor fluorophore (Cy5) at the junction was immobilized. Homology recognition and basepairing at HS1 would result in high FRET due to the close proximity of the donor and acceptor dyes upon docking of dsDNA. **B.** Transition density plot (TDP) of all FRET transitions for immobilized ssDNA with two 6bp repeat sequences (HS1 and HS2) from 236 molecules exhibiting 13020 transitions. **C.** A ssDNA with a single 6bp homology site HS2 ($L_{filament}$= 50nt) with an acceptor fluorophore (Cy5) at the junction was immobilized. Homology recognition and basepairing at HS2 would result in FRET lower than HS1 due to its distal location from the acceptor fluorophore. **D.** Transition density plot (TDP) of all FRET transitions for immobilized ssDNA with two 6bp repeat sequences (HS1 and HS2) from 236 molecules exhibiting 13020 transitions.

Figure 6.7 Summary of transition rates for single 6bp homology site ssDNA



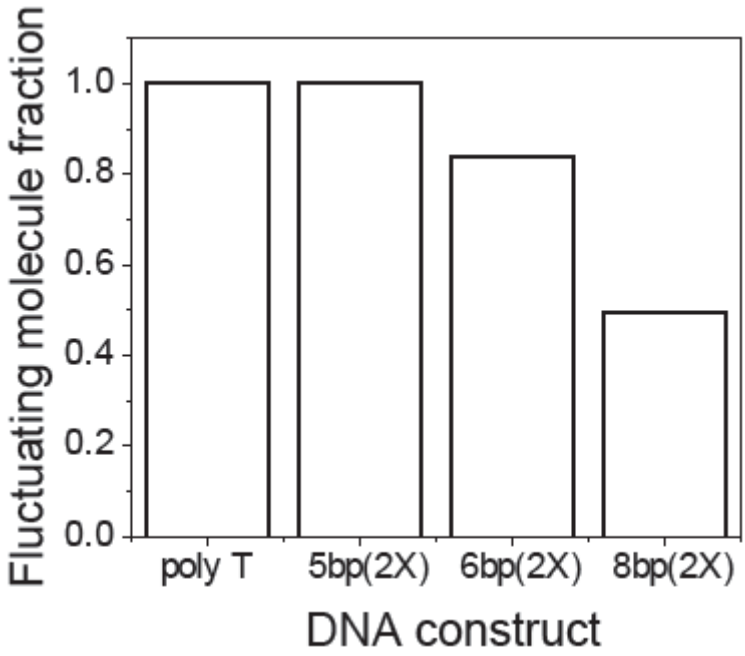
Forward and reverse transition rates between HS1 and NH and also HS2 and NH states are displayed. Error bars denote standard errors of the mean measured for each set of transitions.

Figure 6.8 Dependence of homology length at HS1 and HS2 homology sites



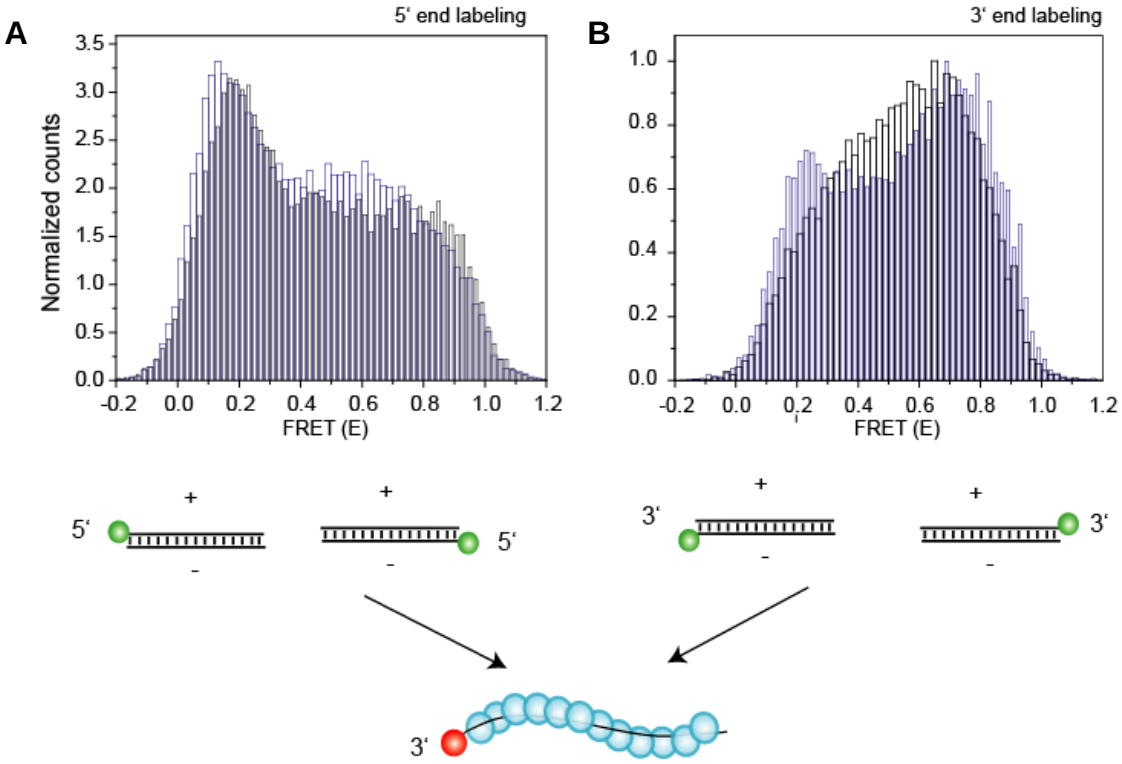
A. Single molecule time traces showing donor (green), acceptor (red) intensities (top panel) and the corresponding FRET trace (blue) (bottom panel) for immobilized ssDNA with two identical 5bp homology sequences at HS1 and HS2 in a poly T sequence background. **B.** Normalized histograms of time traces exhibiting FRET for docking of dsDNA to immobilized ssDNA with two identical 5bp homology sequences (HS1 and HS2) **C.** Single molecule time traces showing donor (green), acceptor (red) intensities (top panel) and the corresponding FRET trace (blue) (bottom panel) for immobilized ssDNA with two identical 8bp homology sequences at HS1 and HS2 in a poly T sequence background. **D.** Normalized histograms of time traces exhibiting FRET for docking of dsDNA to immobilized ssDNA with two identical 8bp homology sequences (HS1 and HS2)

Figure 6.9 Fraction of sliding molecules versus homology length



The fraction of molecules exhibiting at least one FRET fluctuation versus the length of homology at HS1 and HS2 homology sites.

Figure 6.10 RecA filament exhibits preference for binding parallel strand



A. Normalized histogram of FRET events observed upon binding of 5' donor labeled dsDNA. DNA is labeled at the 5' end of the plus strand or its complementary minus strand. **B.** Normalized histogram of FRET events observed upon binding of 3' donor labeled dsDNA. DNA is labeled at the 3' end of the plus strand or its complementary minus strand

Chapter 7

Four color FRET study of RecA strand exchange[4]

The real voyage of discovery consists
of not in seeking new landscapes
but in having new eyes.
Marcel Proust (1913)

7.1 Introduction

In Chapter 3 of the thesis, the kinetics of strand exchange reaction were measured using a two color FRET approach. However, the measurement of the reaction kinetics was based on the assumption that strand exchange reactions primarily initiated from the ends where partial melting of the dsDNA by thermal fluctuations would bias initiation events to primarily initiating from the dsDNA ends. We attributed the absence of linearity in the time for longer DNA lengths to complete strand exchange to the existence of multiple initiation sites with variable reaction kinetics that skew the observed rate distribution. To test whether there is in fact an intrinsic preference for an initiation site and measure the correlation of in strand exchange reaction between the initiating end of the DNA and the end where the reaction proceeds to completion, we performed a four color FRET measurement (see also Chapter 2 for detailed description of Experimental Methods).

7.2 Results

7.2.1 Correlating reaction completion at both ends of the duplex

In conventional two color FRET experiments, it was not possible to correlate the kinetics of events at the two ends of the dsDNA undergoing strand exchange as they are spatially separated by a distance which is well beyond the FRET range (1-10nm). Hence, we implemented a “dual FRET” scheme, where two independent FRET pairs measure the correlation between two spatially separated events. For our

[4] Parts of Chapter 7 have been published
• Lee, J., S. Lee, **K. Ragunathan**, C. Joo, T. Ha and S. Hohng, "Single-molecule four-color FRET", Angew. Chem. Int. Ed. 49(51), 9922-9925 (2010)

measurement, a partial duplex DNA labelled with two acceptors (Cy5 at the junction and Cy7 at the terminal end of the ssDNA) of a single-strand region was immobilized on a polymer-coated quartz surface. After incubating the sample for 15 minutes with RecA and ATPγS to form a RecA filament, we injected a DNA duplex, whose homologous strand is labelled with two donor dyes (Alexa488 and Cy3) at both ends. Alexa488/Cy5 and Cy3/ Cy7 FRET pairs are expected to report on the completion of strand exchange events at both ends (Figure 7.1A). The appearance of high FRET (high acceptor signal) at either end indicates the completion of strand exchange at the labelled end.

Molecules that undergo strand exchange could be classified into those which exhibit a delay at one or both ends of the DNA and those that exhibit no delay at either end. We could broadly classify the molecules we observed into four categories**:** (Figure 7.1B) Type I (46 % of molecules) which did not exhibit any delay between docking and completion of strand exchange at both ends, Type II (18 %) which showed a delay at both ends, and Type III (18 %) and IV (18 %) which showed a delay only at one of the two ends. The distribution of dwell times of the delay at either or both ends of the dsDNA demonstrates that strand exchange can be initiated either from the middle or the ends of the duplex molecule (Figure 7.2A).

We also plotted the relative distribution of delays at both ends. Events with no delay time were arbitrarily assigned to 0.01 s in this plot for displaying the data on a log-log plot. Events of type II (molecules which have delays in completing strand exchange at both ends) exhibit delay times prior to completion of the reaction. An important aspect of this data is the fact that the delay times are correlated with respect to the time it takes for the completion of strand exchange at the two DNA ends. It is also noteworthy that strand exchange is completed more slowly when it is initiated from the middle than when it is initiated from the ends (Figure 7.2B). Our results demonstrate that the initiation of strand exchange can occur from the ends of DNA (Types II and III), or from the middle (Type IV). This conclusion could not be reached from two-color FRET measurements since it was not possible to monitor events at the two ends of single RecA synaptic complexes. The existence of Type I species where both ends completed strand exchange upon docking was also unexpected.

7.3 Discussion

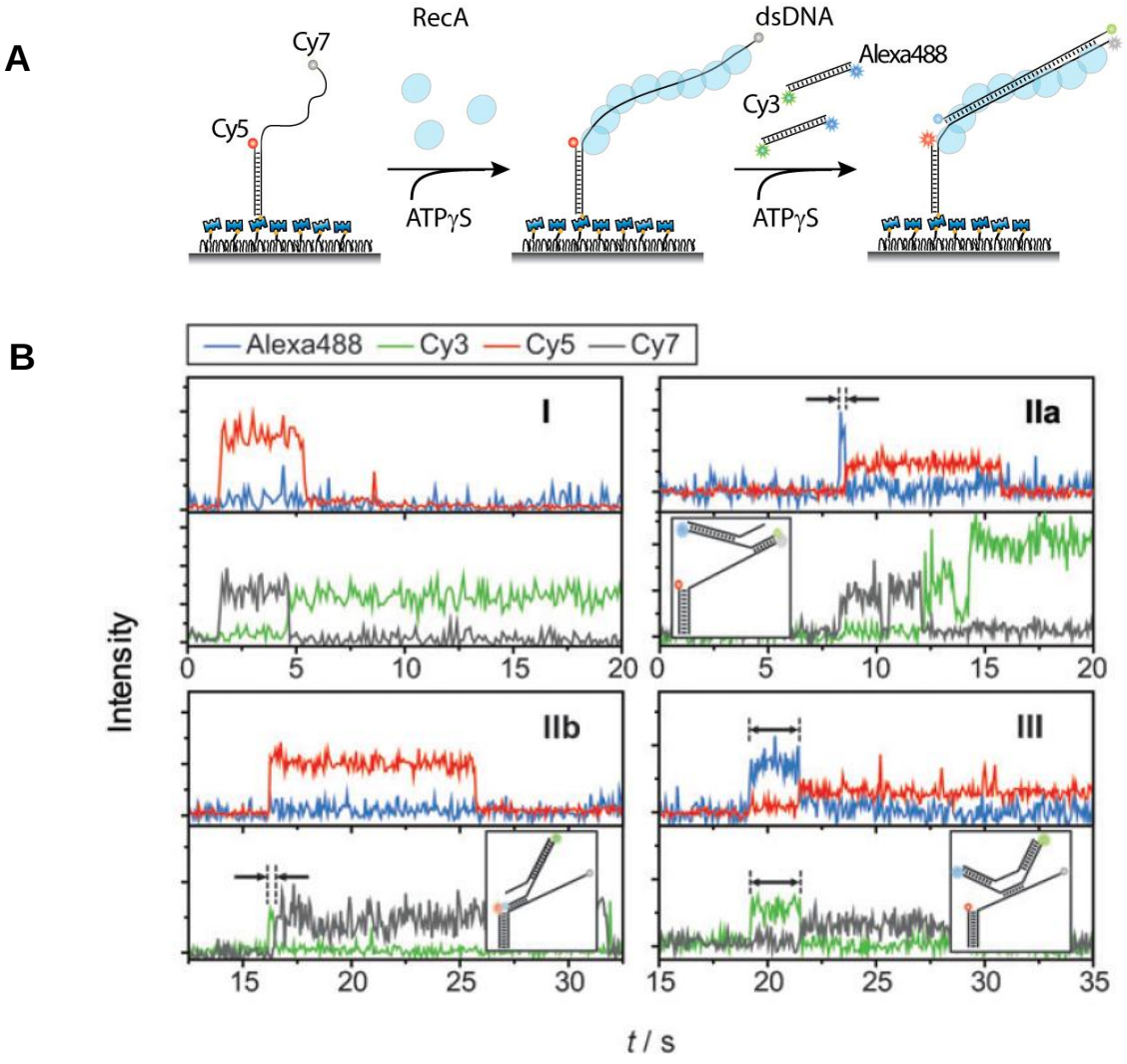
Dual FRET four color measurements reveal a number of surprising observations which two-color FRET cannot provide (Ragunathan et al., 2011) : 1) Events are completed more slowly compared to those which initiated from the ends (Figure 7.2A). 2) Strand exchange events that initiated from the middle showed a strong correlation between the two delay times (Figure 7.2B). It is clear that more systematic studies varying DNA sequence, length and labelling positions would be needed for a complete understanding of this complex process. Our measurements demonstrate the power of multi-dimensional analysis offered by four-color measurements in probing the complexity of biological processes.

7.4 Future questions

Four color measurements provide a unique tool to measure the correlation between the various processes that occur during strand exchange reaction. 1) Labelling the double strand DNA at internal positions can allow for correlation of basepair exchange at the ends and then allow us to draw a time line of events of the reaction progressing towards the middle.2) Labelling of the outgoing strand can allow us to visualize its release while simultaneously monitoring reaction completion. This measurement will provide direct conformation for the model proposed in Chapter 3 where outgoing strand release is uncorrelated from heteroduplex formation 3) The same labelling scheme for monitoring the outgoing strand may also help visualize the role of SSB in the strand exchange reaction by monitoring the FRET change resulting from SSB binding to the outgoing ssDNA.

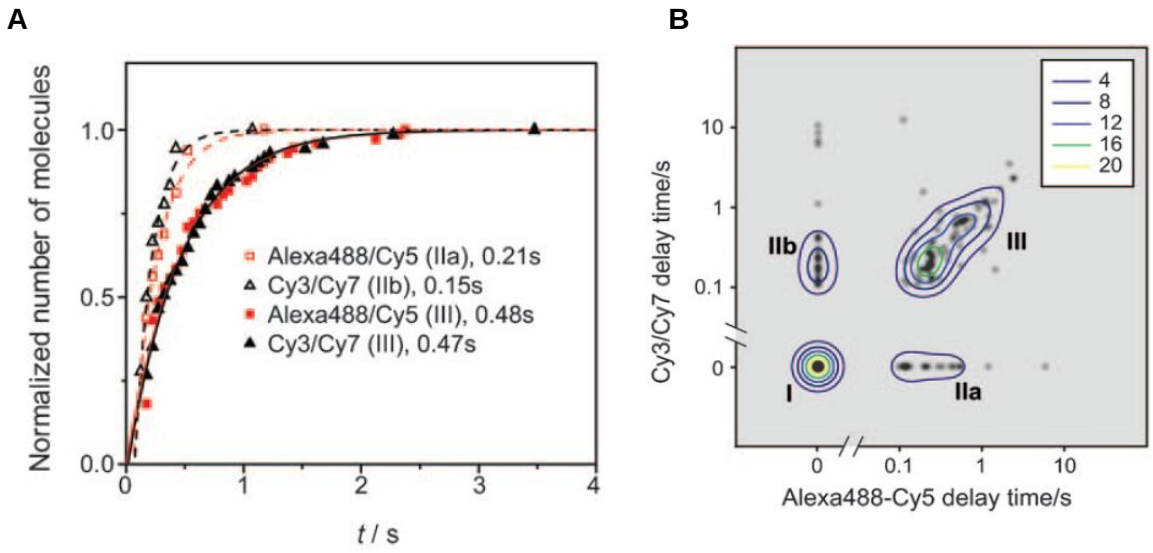
7.5 Figures

Figure 7.1 RecA strand exchange via four color FRET



A. Schematic diagram of the experiments. **B.** Representative fluorescence intensity time traces of strand exchange. In each plot, top panel shows the intensity trace of Alexa488 (blue lines), and that of Cy5 (red lines) at 473-nm excitation and bottom panel shows Cy3 (green lines), and Cy7 (gray lines) signals at 532-nm excitation.

Figure 7.2 Distinct kinetic pathways with different reaction rates



A. Cumulative probabilities of delay times of each type with each delay time being fitted by single exponential function. **B.** Relative distribution of delay times at both ends of the synaptic complex. Each data point is represented by Gaussian peak, each line is represented as the population distribution of data point.

References

Aathavan, K., Politzer, A.T., Kaplan, A., Moffitt, J.R., Chemla, Y.R., Grimes, S., Jardine, P.J., Anderson, D.L., and Bustamante, C. (2009). Substrate interactions and promiscuity in a viral DNA packaging motor. Nature *461*, 669-673.

Adzuma, K. (1992). Stable synapsis of homologous DNA molecules mediated by the Escherichia coli RecA protein involves local exchange of DNA strands. Genes & development *6*, 1679-1694.

Adzuma, K. (1998). No sliding during homology search by RecA protein. J Biol Chem *273*, 31565-31573.

Ameres, S.L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. Cell *130*, 101-112.

Baliga, R., Singleton, J.W., and Dervan, P.B. (1995). RecA.oligonucleotide filaments bind in the minor groove of double-stranded DNA. Proc Natl Acad Sci U S A *92*, 10393-10397.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bazemore, L.R., Folta-Stogniew, E., Takahashi, M., and Radding, C.M. (1997). RecA tests homology at both pairing and strand exchange. Proc Natl Acad Sci U S A *94*, 11863-11868.

Bianchi, M., Riboli, B., and Magni, G. (1985). E. coli recA protein possesses a strand separating activity on short duplex DNAs. EMBO J *4*, 3025-3030.

Bianco, P.R., Tracy, R.B., and Kowalczykowski, S.C. (1998). DNA strand exchange proteins: a biochemical and physical comparison. Front Biosci *3*, 570-603.

Blainey, P.C., van Oijen, A.M., Banerjee, A., Verdine, G.L., and Xie, X.S. (2006). A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. Proc Natl Acad Sci U S A *103*, 5752-5757.

Bonnet, I., Biebricher, A., Porte, P.L., Loverdo, C., Benichou, O., Voituriez, R., Escude, C., Wende, W., Pingoud, A., and Desbiolles, P. (2008). Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. Nucleic Acids Res *36*, 4118-4127.

Camerini-Otero, R.D., and Hsieh, P. (1993). Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. Cell *73*, 217-223.

Cazaux, C., Blanchet, J.S., Dupuis, D., Villani, G., Defais, M., and Johnson, N.P. (1998). Investigation of the secondary DNA-binding site of the bacterial recombinase RecA. J Biol Chem *273*, 28799-28804.

Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.

Chiu, S.K., Rao, B.J., Story, R.M., and Radding, C.M. (1993). Interactions of three strands in joints made by RecA protein. Biochemistry *32*, 13146-13155.

Cirz, R.T., Chin, J.K., Andes, D.R., de Crecy-Lagard, V., Craig, W.A., and Romesberg, F.E. (2005). Inhibition of mutation and combating the evolution of antibiotic resistance. PLoS Biol *3*, e176.

Clark, A.J., and Margulies, A.D. (1965). Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12. Proc Natl Acad Sci U S A *53*, 451-459.

Cox, M.M. (1999). Recombinational DNA repair in bacteria and the RecA protein. Prog Nucleic Acid Res Mol Biol *63*, 311-366.

Cox, M.M. (2007). Motoring along with the bacterial RecA protein. Nat Rev Mol Cell Biol *8*, 127-138.

Cox, M.M., Goodman, M.F., Kreuzer, K.N., Sherratt, D.J., Sandler, S.J., and Marians, K.J. (2000). The importance of repairing stalled replication forks. Nature *404*, 37-41.

Davidsen, T., and Tonjum, T. (2006). Meningococcal genome dynamics. Nat Rev Microbiol *4*, 11-22.

Drees, J.C., Lusetti, S.L., Chitteni-Pattu, S., Inman, R.B., and Cox, M.M. (2004). A RecA filament capping mechanism for RecX protein. Mol Cell *15*, 789-798.

Dunn, K., Chrysogelos, S., and Griffith, J. (1982). Electron microscopic visualization of recA-DNA filaments: evidence for a cyclic extension of duplex DNA. Cell *28*, 757-765.

Folta-Stogniew, E., O'Malley, S., Gupta, R., Anderson, K.S., and Radding, C.M. (2004). Exchange of DNA base pairs that coincides with recognition of homology promoted by E. coli RecA protein. Mol Cell *15*, 965-975.

Forget, A.L., and Kowalczykowski, S.C. (2012). Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. Nature.

Fulconis, R., Mine, J., Bancaud, A., Dutreix, M., and Viovy, J.L. (2006). Mechanism of RecA-mediated homologous recombination revisited by single molecule nanomanipulation. Embo J *25*, 4293-4304.

Galletto, R., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2006). Direct observation of individual RecA filaments assembling on single DNA molecules. Nature *443*, 875-878.

Gonda, D.K., and Radding, C.M. (1986). The mechanism of the search for homology promoted by recA protein. Facilitated diffusion within nucleoprotein networks. J Biol Chem *261*, 13087-13096.

Goodman, M.F. (2000). Coping with replication 'train wrecks' in Escherichia coli using Pol V, Pol II and RecA proteins. Trends Biochem Sci *25*, 189-195.

Gorman, J., Chowdhury, A., Surtees, J.A., Shimada, J., Reichman, D.R., Alani, E., and Greene, E.C. (2007). Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. Mol Cell *28*, 359-370.

Gorman, J., and Greene, E.C. (2008). Visualizing one-dimensional diffusion of proteins along DNA. Nat Struct Mol Biol *15*, 768-774.

Gourves, A.S., Defais, M., and Johnson, N.P. (2001). Equilibrium binding of single-stranded DNA to the secondary DNA binding site of the bacterial recombinase RecA. J Biol Chem *276*, 9613-9619.

Graneli, A., Yeykal, C.C., Robertson, R.B., and Greene, E.C. (2006). Long-distance lateral diffusion of human Rad51 on double-stranded DNA. Proc Natl Acad Sci U S A *103*, 1221-1226.

Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proc Natl Acad Sci U S A *93*, 6264-6268.

Ha, T., Rasnik, I., Cheng, W., Babcock, H.P., Gauss, G.H., Lohman, T.M., and Chu, S. (2002). Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. Nature *419*, 638-641.

Halazonetis, T.D., Gorgoulis, V.G., and Bartek, J. (2008). An oncogene-induced DNA damage model for cancer development. Science *319*, 1352-1355.

Hegner, M., Smith, S.B., and Bustamante, C. (1999). Polymerization and mechanical properties of single RecA-DNA filaments. Proc Natl Acad Sci U S A *96*, 10109-10114.

Hohng, S., Joo, C., and Ha, T. (2004). Single-molecule three-color FRET. Biophys J *87*, 1328-1337.

Honda, M., Park, J., Pugh, R.A., Ha, T., and Spies, M. (2009). Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. Mol Cell *35*, 694-703.

Honigberg, S.M., and Radding, C.M. (1988). The mechanics of winding and unwinding helices in recombination: torsional stress associated with strand transfer promoted by RecA protein. Cell *54*, 525-532.

Hsieh, P., Camerini-Otero, C.S., and Camerini-Otero, R.D. (1990). Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA. Genes Dev *4*, 1951-1963.

Hsieh, P., Camerini-Otero, C.S., and Camerini-Otero, R.D. (1992). The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. Proc Natl Acad Sci U S A *89*, 6492-6496.

Jain, A., Liu, R., Ramani, B., Arauz, E., Ishitsuka, Y., Ragunathan, K., Park, J., Chen, J., Xiang, Y.K., and Ha, T. (2011). Probing cellular protein complexes using single-molecule pull-down. Nature *473*, 484-488.

Jain, S.K., Cox, M.M., and Inman, R.B. (1995). Occurrence of three-stranded DNA within a RecA protein filament. J Biol Chem *270*, 4943-4949.

Joo, C., McKinney, S.A., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006). Real-time observation of RecA filament dynamics with single monomer resolution. Cell *126*, 515-527.

Kowalczykowski, S.C. (2000). Initiation of genetic recombination and recombination-dependent replication. Trends Biochem Sci *25*, 156-165.

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D., and Rehrauer, W.M. (1994). Biochemistry of homologous recombination in Escherichia coli. Microbiol Rev *58*, 401-465.

Kowalczykowski, S.C., and Krupp, R.A. (1995). DNA-strand exchange promoted by RecA protein in the absence of ATP: implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions. Proc Natl Acad Sci U S A *92*, 3478-3482.

Kurumizaka, H., Aihara, H., Ikawa, S., Kashima, T., Bazemore, L.R., Kawasaki, K., Sarai, A., Radding, C.M., and Shibata, T. (1996). A possible role of the C-terminal domain of the RecA protein. A gateway model for double-stranded DNA binding. J Biol Chem *271*, 33515-33524.

Kurumizaka, H., Ikawa, S., Sarai, A., and Shibata, T. (1999). The mutant RecA proteins, RecAR243Q and RecAK245N, exhibit defective DNA binding in homologous pairing. Arch Biochem Biophys *365*, 83-91.

Kuzminov, A. (1999). Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. Microbiol Mol Biol Rev *63*, 751-813, table of contents.

Lee, A.M., Xiao, J., and Singleton, S.F. (2006). Origins of sequence selectivity in homologous genetic recombination: insights from rapid kinetic probing of RecA-mediated DNA strand exchange. J Mol Biol *360*, 343-359.

Lee, J., Lee, S., Ragunathan, K., Joo, C., Ha, T., and Hohng, S. (2010a). Single-molecule four-color FRET. Angew Chem Int Ed Engl *49*, 9922-9925.

Lee, S., Lee, J., and Hohng, S. (2010b). Single-molecule three-color FRET with both negligible spectral overlap and long observation time. PLoS One *5*, e12270.

Liu, S., Abbondanzieri, E.A., Rausch, J.W., Le Grice, S.F., and Zhuang, X. (2008). Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. Science *322*, 1092-1097.

Lohman, T.M., and Overman, L.B. (1985). Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. J Biol Chem *260*, 3594-3603.

Lusetti, S.L., Shaw, J.J., and Cox, M.M. (2003). Magnesium ion-dependent activation of the RecA protein involves the C terminus. J Biol Chem *278*, 16381-16388.

Malkov, V.A., Panyutin, I.G., Neumann, R.D., Zhurkin, V.B., and Camerini-Otero, R.D. (2000). Radioprobing of a RecA-three-stranded DNA complex with iodine 125: evidence for recognition of homology in the major groove of the target duplex. J Mol Biol *299*, 629-640.

Mani, A., Braslavsky, I., Arbel-Goren, R., and Stavans, J. (2009). Caught in the act: the lifetime of synaptic intermediates during the search for homology on DNA. Nucleic Acids Res *38*, 2036-2043.

Mazin, A.V., and Kowalczykowski, S.C. (1996). The specificity of the secondary DNA binding site of RecA protein defines its role in DNA strand exchange. Proc Natl Acad Sci U S A *93*, 10673-10678.

Mazin, A.V., and Kowalczykowski, S.C. (1998). The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. Embo J *17*, 1161-1168.

McKinney, S.A., Joo, C., and Ha, T. (2006). Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys J *91*, 1941-1951.

Menetski, J.P., Bear, D.G., and Kowalczykowski, S.C. (1990). Stable DNA heteroduplex formation catalyzed by the Escherichia coli RecA protein in the absence of ATP hydrolysis. Proc Natl Acad Sci U S A *87*, 21-25.

Michel, B. (2005). After 30 years of study, the bacterial SOS response still surprises us. PLoS Biol *3*, e255.

Modesti, M., Ristic, D., van der Heijden, T., Dekker, C., van Mameren, J., Peterman, E.J., Wuite, G.J., Kanaar, R., and Wyman, C. (2007). Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. Structure *15*, 599-609.

Moffitt, J.R., Chemla, Y.R., and Bustamante, C. (2010). Methods in statistical kinetics. Methods Enzymol *475*, 221-257.

Myong, S., Bruno, M.M., Pyle, A.M., and Ha, T. (2007). Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. Science *317*, 513-516.

Myong, S., Rasnik, I., Joo, C., Lohman, T.M., and Ha, T. (2005). Repetitive shuttling of a motor protein on DNA. Nature *437*, 1321-1325.

Nishinaka, T., Ito, Y., Yokoyama, S., and Shibata, T. (1997). An extended DNA structure through deoxyribose-base stacking induced by RecA protein. Proc Natl Acad Sci U S A *94*, 6623-6628.

Nishinaka, T., Shinohara, A., Ito, Y., Yokoyama, S., and Shibata, T. (1998). Base pair switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination. Proc Natl Acad Sci U S A *95*, 11071-11076.

Park, J., Myong, S., Niedziela-Majka, A., Lee, K.S., Yu, J., Lohman, T.M., and Ha, T. (2010). PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. Cell *142*, 544-555.

Podyminogin, M.A., Meyer, R.B., and Gamper, H.B. (1995). Sequence-specific covalent modification of DNA by cross-linking oligonucleotides. Catalysis by RecA and implication for the mechanism of synaptic joint formation. Biochemistry *34*, 13098-13108.

Potter, H., and Dressler, D. (1976). On the mechanism of genetic recombination: electron microscopic observation of recombination intermediates. Proc Natl Acad Sci U S A *73*, 3000-3004.

Ragone, S., Maman, J.D., Furnham, N., and Pellegrini, L. (2008). Structural basis for inhibition of homologous recombination by the RecX protein. Embo J *27*, 2259-2269.

Ragunathan, K., Joo, C., and Ha, T. (2011). Real-time observation of strand exchange reaction with high spatiotemporal resolution. Structure *19*, 1064-1073.

Rasnik, I., Myong, S., Cheng, W., Lohman, T.M., and Ha, T. (2004). 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. J Mol Biol *336*, 395-408.

Reymer, A., Frykholm, K., Morimatsu, K., Takahashi, M., and Nordén, B. (2009). Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy. Proc Natl Acad Sci U S A *106*, 13248-13253.

Rosselli, W., and Stasiak, A. (1990). Energetics of RecA-mediated recombination reactions. Without ATP hydrolysis RecA can mediate polar strand exchange but is unable to recycle. J Mol Biol *216*, 335-352.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. Nat Methods *5*, 507-516.

Roy, R., Kozlov, A.G., Lohman, T.M., and Ha, T. (2009). SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. Nature *461*, 1092-1097.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods *3*, 793-795.

Shibata, T., Nishinaka, T., Mikawa, T., Aihara, H., Kurumizaka, H., Yokoyama, S., and Ito, Y. (2001). Homologous genetic recombination as an intrinsic dynamic property of a DNA structure induced by RecA/Rad51-family proteins: a possible advantage of DNA over RNA as genomic material. Proc Natl Acad Sci U S A *98*, 8425-8432.

Shivashankar, G.V., Feingold, M., Krichevsky, O., and Libchaber, A. (1999). RecA polymerization on double-stranded DNA by using single-molecule manipulation: the role of ATP hydrolysis. Proc Natl Acad Sci U S A *96*, 7916-7921.

Spies, M., and Kowalczykowski, S.C. (2005). Homologous recombination by RecBCD and RecF pathways. In The bacterial chromosome, N.P. Higgins, ed. (Washington, D.C., ASM Press), pp. 389-403.

Stasiak, A., Di Capua, E., and Koller, T. (1981). Elongation of duplex DNA by recA protein. J Mol Biol *151*, 557-564.

Stasiak, A., and Egelman, E.H. (1986). RecA-DNA helical complexes in genetic recombination. Biochem Soc Trans *14*, 218-220.

Story, R.M., and Steitz, T.A. (1992). Structure of the recA protein-ADP complex. Nature *355*, 374-376.

Story, R.M., Weber, I.T., and Steitz, T.A. (1992). The structure of the E. coli recA protein monomer and polymer. Nature *355*, 318-325.

Sutton, M.D., Smith, B.T., Godoy, V.G., and Walker, G.C. (2000). The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. Annu Rev Genet *34*, 479-497.

Uphoff, S., Holden, S.J., Le Reste, L., Periz, J., van de Linde, S., Heilemann, M., and Kapanidis, A.N. (2010). Monitoring multiple distances within a single molecule using switchable FRET. Nat Methods *7*, 831-836.

van der Heijden, T., Modesti, M., Hage, S., Kanaar, R., Wyman, C., and Dekker, C. (2008). Homologous recombination in real time: DNA strand exchange by RecA. Mol Cell *30*, 530-538.

van Loenhout, M.T., van der Heijden, T., Kanaar, R., Wyman, C., and Dekker, C. (2009). Dynamics of RecA filaments on single-stranded DNA. Nucleic Acids Res *37*, 4089-4099.

Voloshin, O.N., and Camerini-Otero, R.D. (2004). Synaptic complex revisited; a homologous recombinase flips and switches bases. Mol Cell *15*, 846-847.

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T., and Patel, D.J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature *461*, 754-761.

Wang, Y.M., Austin, R.H., and Cox, E.C. (2006). Single molecule measurements of repressor protein 1D diffusion on DNA. Phys Rev Lett *97*, 048302.

Wyman, C. (2011). Mechanistic insight from chaos: how RecA mediates DNA strand exchange. Structure *19*, 1031-1032.

Xiao, J., and Singleton, S.F. (2002). Elucidating a key intermediate in homologous DNA strand exchange: structural characterization of the RecA-triple-stranded DNA complex using fluorescence resonance energy transfer. J Mol Biol *320*, 529-558.

Yeom, K.H., Heo, I., Lee, J., Hohng, S., Kim, V.N., and Joo, C. (2011). Single-molecule approach to immunoprecipitated protein complexes: insights into miRNA uridylation. EMBO Rep *12*, 690-696.

Yeykal, C.C., and Greene, E.C. (2006). Visualizing the behavior of human Rad51 at the single-molecule level. Cell Cycle *5*, 1033-1038.

Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y.E., and Selvin, P.R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science *300*, 2061-2065.

Yildiz, A., Tomishige, M., Vale, R.D., and Selvin, P.R. (2004). Kinesin walks hand-over-hand. Science *303*, 676-678.

Zhou, R., Kozlov, A.G., Roy, R., Zhang, J., Korolev, S., Lohman, T.M., and Ha, T. (2011). SSB functions as a sliding platform that migrates on DNA via reptation. Cell *146*, 222-232.

Zhou, X., and Adzuma, K. (1997). DNA strand exchange mediated by the Escherichia coli RecA protein initiates in the minor groove of double-stranded DNA. Biochemistry *36*, 4650-4661.

Appendix A

Experimental procedures

Listed below are experimental procedures used during the course of this thesis.

A.1 Strand exchange measurements

Acceptor labeled reference ssDNA molecules were immobilized on the passivated surface by means of a biotin-neutravidin interaction. After washing away excess of acceptor molecules, the reference ssDNA was incubated with 1μM RecA (New England Biolabs) and 1mM ATPγS (Calbiochem) in an incubation buffer containing 25mM Tris Acetate pH 7.5, 100mM Sodium Acetate and 1mM Magnesium Acetate. In some cases, 1mM ATP was used instead of ATPγS. After incubation for 15 minutes to ensure complete filament formation on the reference ssDNA molecules, the buffer in the chamber was exchanged with a solution of homologous 500pM dsDNA and 1mM ATPγS in a strand exchange buffer (25mM Tris Acetate pH 7.5, 100mM Sodium Acetate, 10mM Magnesium Acetate) supplemented with an oxygen scavenging system (1mg/mL glucose oxidase, 0.8% glucose, 0.04mg/mL catalase and 3mM Trolox). Imaging was initiated as soon as the buffer exchange was complete. All measurements were carried out at room temperature (23±1°C). In addition to the above components, we also added SSB (a generous gift from Dr. T.M. Lohman, Washington University) depending on the experimental scheme.

A.2 Restriction enzyme assay to test for heteroduplex formation

After joint molecule formation (L_h= 39nt), we removed RecA from the DNA by exchanging the solution in the imaging chamber with a buffer containing no magnesium (10mM Tris-Cl, 50mM NaCl, pH 8.0). The restriction enzyme DdeI (New England Biolabs) was suspended in the vendor supplied reaction buffer and flowed into the channel containing the DNA. After an incubation period of 30 mins at 37ºC, we counted the number of molecules that remained bound to the surface. Appropriate controls were carried out to ensure that there was no non- specific cleavage activity under the conditions used.

A.3 DNA preparation

DNA oligos used in our measurements were purchased from Integrated DNA Technologies (IDT). The oligos were suspended in T50 buffer (10mM Tris-Cl, 50mM NaCl, pH 8.0). Double strand DNA was prepared by mixing complementary DNA molecules and heating to 90^0C followed by slow cooling to room temperature over a period of 2 hours. dsDNA was purified from free ssDNA using a 12% native PAGE gel to ensure the absence of free ssDNA in our DNA preparations. The partial duplex DNA molecules for immobilization were prepared by annealing with the biotin_DNA sequence All DNA molecules were labeled at the terminal ends (5' or 3') with Cy3 or Cy5 (labeling performed by IDT) as specified in the experimental scheme. For internally labeled DNA oligos, Cy3 N-hydroxysuccinimido (NHS) ester and Cy5 NHS ester (GE Healthcare) were internally labeled to dT of ssDNA via a C6 amino linker (IDT).

A.4 Cleaning and passivation of quartz and glass slides

Visualizing single molecules in a reliable way requires the preparation of slides which are free of auto fluorescent impurities and also are passivated to repel the non-specific binding of labeled protein and nucleic acids.

Methods for passivation of slides involve the use of Bovine Serum Albumin (BSA) (Figure A.1A) or Polyethylene glycol (PEG) (Figure A.1B). In general, PEG surfaces are more hydrophobic and suppress non-specific binding to a larger degree compared to BSA coated surfaces (A.6.1A). However, the preparation of BSA coated surfaces is convenient and may be used in the case of experiments involving the visualization of nucleic acids which in general do not non-specifically adsorb to the surface as much as proteins.

Cleaning slides involves the use of acetone and a high molar solution (>1M) of potassium hydroxide to strip the quartz surface of impurities and also activate functional moieties on the glass surface for subsequent chemistry with PEG for surface passivation. There are several variations of this protocol primarily arising from differences in the molarity of KOH used and at least one procedure to my knowledge recommends the use of saturated KOH solution in ethanol as a cleaning agent.

The procedure to clean slides is as follows:

1. Sonicate slides in a Teflon container with acetone for 20 mins.

2. Take the container out. Dispose of acetone. Rinse with MilliQ water

3. Fill containers with 1M KOH and sonicate for 20 mins

4. Rinse slides with MilliQ water and sonicate with water for 10 mins.

5. Burn slides thoroughly for about one minute to ensure that residual impurities are completely eliminated.

6. Cool it down by blowing nitrogen in or let it sit at room temperature for about 10 mins.

The same sequence of steps needs to be followed to clean glass coverslips.

If using BSA to passivated slides, the slides can be assembled and BSA labeled with biotin can be flowed into imaging chambers and it will non-specifically coat the slide surface.

If using PEG to passivated slides, subsequent procedures to activate the quartz surface for subsequent reaction with a PEG-NHS ester need to be followed.

1. Sonicate cleaned quartz slides and coverslips in methanol for ~ 20 mins

2. Incubate with a 1% (v/v) solution of amino silane in methanol and acetic acid which serves to functionalize the quartz surface with amine groups.

3. After 20 mins, rinse the surface with water and methanol.

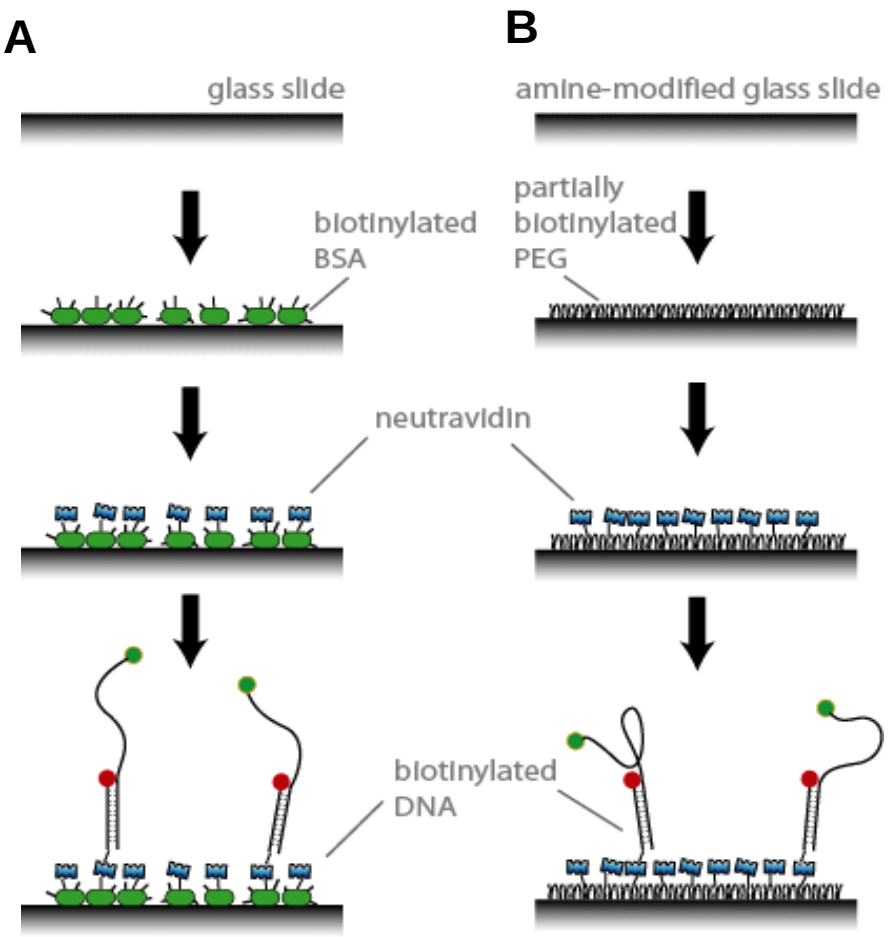
4. Prepare a solution of mPEG and biotin PEG which are both modified with reactive NHS ester groups that form a covalent linkage with the amine coated quartz surface. The ratio of biotin PEG is typically 1-3% of the mPEG used.

5. Coat the desired area of the quartz imaging surface slide with the solution prepared in step 4 and incubate for 3-4 hours

6. Wash slides and store in vacuum at -20ºC. Assemble slide and coverslip prior to use.

A.5 Figures and Tables

Figure A.1 Passivation of single molecule imaging surface



A. Glass slide after cleaning is coated with BSA tagged with biotin. This is followed by addition of neutravidin and immobilization of the biotin tagged molecule of interest **B.** In the case of slides passivated with polyethylene glycol (PEG), slides are activated with an amine modifier following which a reaction with NHS ester modified PEG, leads to passivation of the imaging surface. Addition of neutravidin is followed by immobilization of biotin tagged molecules of interest

Table A.1 DNA sequences used for strand exchange experiments

Name	DNA sequence (5' -> 3')	Experiment
biotin_DNA	/Biotin/TGG CGA CGG CAG CGA GGC/Cy5/	For Reference ssDNA immobilization
21+_ reference	CTT TTC ATC ACG TTG TTA GAT GCC TCG CTG CCG TCG CCA	L_h= 21
21-	ATC TAA CAA CGT GAT GAA AAG	Incoming DNA
21+	CTT TTC ATC ACG TTG TTA GAT	Outgoing DNA
31+_ reference	ATG AGC GCC ACT TTT CAT CAC GTT GTT AGA TGC CTC GCT GCC GTC GCC A	L_h= 31
31+	ATG AGC GCC ACT TTT CAT CAC GTT GTT AGA T	Incoming DNA
31-	ATC TAA CAA CGT GAT GAA AAG TGG CGC TCA T	Outgoing DNA
C39+_reference	TTT ACT TGT ACT TCA TTC ATT CAC ATT CCT ATC ATG TTT GCC TCG CTG CCG TCG CCA	L_h= 39 (sequence 1)
T30 C39+_reference	T30 TTT ACT TGT ACT TCA TTC ATT CAC ATT CCT ATC ATG TTT GCC TCG CTG CCG TCG CCA	For template DNA length effect, 69nt
T60 C39+_reference	T60 TTT ACT TGT ACT TCA TTC ATT CAC ATT CCT ATC ATG TTT GCC TCG CTG CCG TCG CCA	For template DNA length effect, 99nt
172 C39+_reference	T78 GAG ACT ACG TAC CAG GTA GTT ACG T31 TTT ACT TGT ACT TCA TTC ATT CAC ATT CCT ATC ATG TTT GCC TCG CTG CCG TCG CCA	For template DNA length effect, 172nt
C39+	TTT ACT TGT ACT TCA TTC ATT CAC ATT CCT ATC ATG TTT	Incoming DNA
C39-	AAA CAT GAT AGG AAT GTG AAT GAA TGA AGT ACA AGT AAA	Outgoing DNA
K39+_reference	ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT GCC TCG CTG CCG TCG CCA	L_h= 39 (sequence 2)
K39+	ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT	Incoming DNA
K39-	AGC TTA GAG ATT GAT ATT CAG ATC ACC CGC GAC GTT CAT	Outgoing DNA
K39+ dissociation	TGG CGA CGG CAG CGA GGC T20 ATG AAC G/Cy5-dT/C GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT T5	ATP mediated RecA dissociation
K39- dissociation	AGC TTA GAG ATT GAT ATT CAG A/Cy3-dT/C ACC CGC GAC GTT CAT	ATP mediated RecA dissociation

Table A.1 cont.

45+_reference	GCA TAC ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT GCC TCG CTG CCG TCG CCA	L_h= 45
45+	GCA TAC ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT	Incoming DNA
45-	AGC TTA GAG ATT GAT ATT CAG ATC ACC CGC GAC GTT CAT GTA TGC	Outgoing DNA
60+_reference	ATC TAA CAA CCT GAT GAA AAG ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT GCC TCG CTG CCG TCG CCA	L_h= 60
60+	ATC TAA CAA CCT GAT GAA AAG ATG AAC GTC GCG GGT GAT CTG AAT ATC AAT CTC TAA GCT	Incoming DNA
60-	AGC TTA GAG ATT GAT ATT CAG ATC ACC CGC GAC GTT CAT CTT TTC ATC AGG TTG TTA GAT	Outgoing DNA
70+_reference	CTT ACC GAA CTG AAC TCT GGC TAA TGT CTA AAT GAA CGT CGC GGG TGA TCT GAA TAT CAA TCT CTA AGC TGC CTC GCT GCC GTC GCC A	L_h= 70
70-	AGC TTA GAG ATT GAT ATT CAG ATC ACC CGC GAC GTT CAT TTA GAC ATT AGC CAG AGT TCA GTT CGG TAA G	Incoming DNA
70+	CTT ACC GAA CTG AAC TCT GGC TAA TGT CTA AAT GAA CGT CGC GGG TGA TCT GAA TAT CAA TCT CTA AGC T	Outgoing DNA
80+_reference	TTC ACC TTT AAT CTT TAT ACA CGT TCT TCT CAC CAG CCA GTT CTT ATC CGT CAA TAT TTA CTT CTC TAT CAA CTA CCA ATG CCT CGC TGC CGT CGC CA	L_h= 80
80-	ATT GGT AGT TGA TAG AGA AGT AAA TAT TGA CGG ATA AGA ACT GGC TGG TGA GAA GAA CGT GTA TAA AGA TTA AAG GTG AA	Incoming DNA
80+	TTC ACC TTT AAT CTT TAT ACA CGT TCT TCT CAC CAG CCA GTT CTT ATC CGT CAA TAT TTA CTT CTC TAT CAA CTA CCA AT	Outgoing DNA
2X 6bp pdT	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GTT CAT TTT TTG TTC ATG CCT CGC TGC CGT CGC CA	HS1 and HS2 6bp homology
1X 6bp end pdT	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTG TTC ATG CCT CGCTGC CGT CGC CA	HS1 6bp homolology

Table A.1 cont.

1X 6bp mid pdT	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GTT CAT TTT TTT TTT TTG CCT CGCTGC CGT CGC CA	HS2 6bp homololology
2X 5bp pdT	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC ATT TTT TTC ATG CCT CGCTGC CGT CGC CA	HS1 and HS2 5bp homology
2X 8bp pdT	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTA CGT TCA TTT TTT ACGTTC ATG CCT CGC TGC CGT CGC CA	HS1 and HS2 8bp homology
poly T50	T50 GCC TCG CTG CCG TCG CCA	poly T DNA

Appendix B

Analysis methods

B.1 FRET calculation

FRET between two fluorophores serves as a spectroscopic ruler to measure relative distance changes within or between two molecules. The efficiency of energy transfer (E) displays a distance dependence given by,

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

where *R* is the separation between the donor and the acceptor, and R_0 is known as Förster radius which depends on the fluorophore pair which is used. This value can be independently calculated by knowledge of the fluorophore properties .

$$R_0 = \left[(8.785\times10^{17})\Phi^D\kappa^2 n^{-4} J(\upsilon)\right]^{\frac{1}{6}} \quad nm$$

where Φ^D is the donor quantum yield in the absence of the acceptor, *n* is the refractive index of the medium and κ is the orientation factor for the dipole-dipole interaction. κ^2 is estimated to be 2/3 for freely rotating dipoles. $J(\nu)$ is the integral of the spectral overlap between the donor emission and the acceptor absorption.

For the typical fluorophores we use, Cy3 and Cy5, the value of R_0 is ~ 54Å.

In our single molecule measurements however, the information which we have is the donor and acceptor intensities and their ratio provides an approximation for the energy transfer efficiency,

$$E = \frac{I_A}{I_A + I_D}$$

where, I_A and I_D represent the intensities of the donor and acceptor fluorophores respectively.

B.2 Corrections for single molecule FRET data

When we measure the intensities of the donor and acceptor fluorophores there are different contributions which we need to account for.

B.2.1 Leakage correction from donor to acceptor channel

Depending on the dichroic used, there may be leakage from the donor channel to the acceptor channel which needs to be subtracted. The leakage can be measured as the remaining signal in the acceptor channel after acceptor photobleaching. Donor leakage correction factor ($\Delta I_{D->A}$),

$$\Delta I_{D->A} = \frac{I_A}{I_D}$$

We can then correct the raw acceptor intensity obtained through the measurement in the following manner,

$$\Delta I_{A(corrected)} = I_A - \Delta I_{D->A} \quad * I_D$$

For the typical fluorophores, Cy3 and Cy5 and the dichroics we use, the leakage correction factor is ~ 15-18%.

B.2.2 Gamma factor correction

In order to correct for differences in quantum yields between fluorophores, differences in transmission efficiencies in the optics and difference in detection efficiency of the emissions from the various dyes, we use a scaling factor referred to as gamma correction factor.

$$\gamma_{ij} = \frac{\Delta I_i}{\Delta I_j}$$

ΔI_D and ΔI_A represents the change in donor and acceptor intensity when the molecule undergoes photobleaching or a change in conformation which produces a change in FRET.

B.2.3 Measuring FRET for more than two fluorophores

The approximation mentioned above for calculating FRET is valid when there is only one donor and one acceptor. However, in the event that there are three fluorophores in close proximity with three possible FRET interactions (see Chapter 2 for more details), we need to modify our scheme for FRET calculation (Lee et al., 2010a). This situation involves using an alternating laser excitation scheme to independently excite different FRET pairs.

We can obtain E_{34} as follows:

$$E_{34} = \frac{1}{I_3/I_4 + 1} = \frac{I_4}{I_3 + I_4}$$

At green excitation, we can measure the fluorescence intensities of Cy3, Cy5, and Cy7 and using E_{34} obtained at 633-nm laser excitation, we obtain E_{23} and E_{24} as follows:

$$E_{23} = \frac{I_3}{I_2(1 - E_{34}) + I_3}$$

$$E_{24} = \frac{I_4 - E_{34}(I_3 + I_4)}{I_2 + I_4 - E_{34}(I_2 + I_3 + I_4)}$$

Further modification of the above equation is required in the case where there are four fluorophores. In the event that Cy5 and Cy7 dyes are far away from each other with no significant FRET interaction (E_{34} ~0), then E_{24} and E_{23} can be simply calculated as the ratio of acceptor intensity for each fluorophore divided by the total intensity.

B.3 Gamma distribution fitting

One of the analysis methods used for single molecule data to describe the distribution of rates from kinetics measurements is the gamma distribution fit (Moffitt et al., 2010). This method of analysis has been useful to extract information regarding the fundamental step size for different translocases (Myong et al., 2007; Yildiz et al., 2003; Yildiz et al., 2004).

This equation can be used under the assumption that the reaction consists of *N* steps from start to finish of the form:

$$\text{begin} \xrightarrow{\overbrace{\xrightarrow{k}\xrightarrow{k}\ldots\ldots\xrightarrow{k}\xrightarrow{k}}^{\text{N steps}}} \text{end}$$

and the distribution of rates for the N irreversible steps in this process can be described by the equation,

$$y = \Delta t_{delay}{}^{N-1} e^{-k\Delta t_{delay}}$$

This equation is constrained by two free parameters, *N* and *k* which provide information regarding the number of steps in the reaction and the rate for each step (which by definition must be identical).

B.4 Dwell time analysis and Hidden Markov Models

Proteins and nucleic acids undergo conformational changes which translate into changes in FRET efficiency. We can observe the transitions between the different FRET states in real time. To learn about mechanistic details of a system we can measure the time that a molecule spends in each state and measure how the dwell times in each state change as a function of different parameters such as salt, temperature etc. For example, if a molecule undergoes transitions between two states A and B we can measure the lifetime of a molecule in state A and state B which provides us with information about the transition rate from A to B ($k_{A->B}$) and the reverse transition rate from B to A ($k_{B->A}$). While it is possible to measure the kinetics for a two state system by manually measuring the lifetime in each state or by some sort of a threshold analysis, biological systems can be more complicated and can exhibit multiple states.

In the case of a multi-state system, it is difficult to use the aforementioned methods to analyze the data. In this case, we can use Hidden Markov Model (HMM) analysis which uses a probabilistic framework to make FRET state assignments and calculate the dwell times of each state in an unbiased manner (Joo et al., 2006; McKinney et al., 2006). The data from the HMM analysis can be represented in the form of a transition density plot (TDP) which provides a visual cue of the possible transitions

between states and their relative probability. HMM can be applied in cases where the data exhibits discrete FRET states (see Chapter 6 for implementation).

B.5 Cross correlation analysis

In many cases, proteins or nucleic acids may undergo rapid changes in conformation which results in fast changes in FRET between different states that cannot be easily characterized using a simple dwell time analysis (see Chapter 5 for implementation). In such a situation, it is still possible to quantify the data by obtaining a measure of the time scale of fluctuations. The cross correlation of two signals (donor and acceptor):

$$CC(\tau) = \int I_D(t) I_A(t + \Delta\tau) dt$$

The cross correlation can be fit by a single exponential curve providing a characteristic time scale for the fluctuations. This curve typically decays from a negative value to zero. Given a two state system with states A and B, then the inverse of the characteristic time obtained via exponential fitting is equal to the sum of transition rates between states A and B i.e. $k_{B->A}$ + $k_{A->B}$. Also, the relative population of state A versus state B is given by the ratio of the two transition rates. The two equations can be used to compute the individual transition rates between the two states.

B.6 Histogram analysis

While single molecule time traces provide interesting insights into the behavior of individual molecules, it is difficult to objectively gauge the existence of distinct FRET states between molecules and also measure their FRET values. To eliminate bias in this process and also to obtain an ensemble type view of the system under study, we accumulate brief movies of several thousands of molecules to build a histogram which provides an unbiased representation of the state of the molecules within a particular time window. The histogram provides a rapid way to sample trends in the behavior of molecules both as a function of time and as a function of change in solution conditions.

Appendix C

RecA filament stability and photophysical effects

C.1 Filament stability during solution exchange

Strand exchange measurements were carried out by first forming RecA filaments in the presence of ATPɣS followed by which a solution containing dsDNA and ATPɣS and no RecA in solution was added. To test that the absence of RecA in solution does not induce filament disassembly we utilized a ssDNA capture assay where the presence of ssDNA can be rapidly detected by the presence of a protein such as SSB which binds with high affinity to ssDNA. We tested the effect of adding SSB to a 39nt ssDNA with a donor at the ssDNA end and an acceptor at the ssDNA/dsDNA junction (Figure C.1). ssDNA alone produces a FRET peak which is lower in value compared to SSB bound ssDNA which induces tight wrapping of the ssDNA thus resulting in a high FRET state (Figure C.1).

For the same DNA we added RecA and ATPɣS which resulted in a shift towards the low FRET state due to the stretched conformation of ssDNA induced by RecA binding (Figure C.2). During our procedure for solution exchange, we added ATPɣS and SSB. There was no change in FRET upon exchanging the solution thus validating our experimental procedure for ensuring the presence of stable RecA filaments on the immobilized DNA (Figure C.2). This procedure of eliminating RecA in solution when adding dsDNA ensures that the dsDNA is capable of interacting only with the immobilized RecA filament.

C.2 Photophysical effects during RecA strand exchange

In Chapter 3, single molecule measurements indicate the existence of a low FRET state upon dsDNA docking followed by an abrupt transition to the high FRET state. The abrupt nature of the FRET transition arising from the propagation of strand exchange prompted us to question possible photophysical origins for the initial low FRET state (Figure C.3). The reason for raising such a concern lies in the fact that Cy5 has an increased propensity for transitioning to a dark non-emitting state. However, the presence of a donor can quickly trigger its transition to an emitting state resulting in red fluorescence emission. Our concern was related to the fact that

our assay of bringing the donor in close proximity to the acceptor may recapitulate such a scenario. This principle has been used to make the dyes blink stochastically and with subsequent localization and detection leading to the reconstruction of a high resolution image (Rust et al., 2006).

We first formed RecA filaments following which we counted the number of acceptor molecules before and after RecA filament formation. However, the number of Cy5 molecules upon direct excitation remains relatively unchanged pointing to the fact that the presence of a RecA filament does not lead to Cy5 quenching.

Given that blinking can be modulated by changing solution conditions or by changing laser power (532nm laser), we tested for both effects by first changing the reductant in our imaging buffer from β-mercaptoethanol (BME) to Trolox. Trolox has been shown to strongly suppress the effect of fluorophore blinking. Using BME or Trolox did not significantly affect the kinetics of strand exchange. Similarly, changing laser power did not affect the lifetime of the low FRET state indicating that the low to high FRET change was not a consequence of donor mediated activation of acceptor emission.

C.3 Using alternating excitation to detect acceptor status

We used alternating laser excitation (ALEX) to continuously monitor the intensity of Cy5 during the course of the strand exchange reaction thus enabling us to precisely assign whether the acceptor is in the active emitting state prior to strand exchange reaction. Figure C.4A and C.4B present examples of molecules where the top panel represents the typical single molecule time traces observed upon docking and subsequent pairing of homologous dsDNA with the RecA-ssDNA complex. The bottom panel presents the intensity of the acceptor (Cy5) and in both examples, it is clear that Cy5 is in the emitting state at the moment of reaction. It is interesting to note that the completion of strand exchange which brings Cy3 and Cy5 in close proximity results in the quenching of Cy5 dye as seen from a decrease in Cy5 emission (which is different from photobleaching of the acceptor fluorophore where the intensity level reaches that of the background).

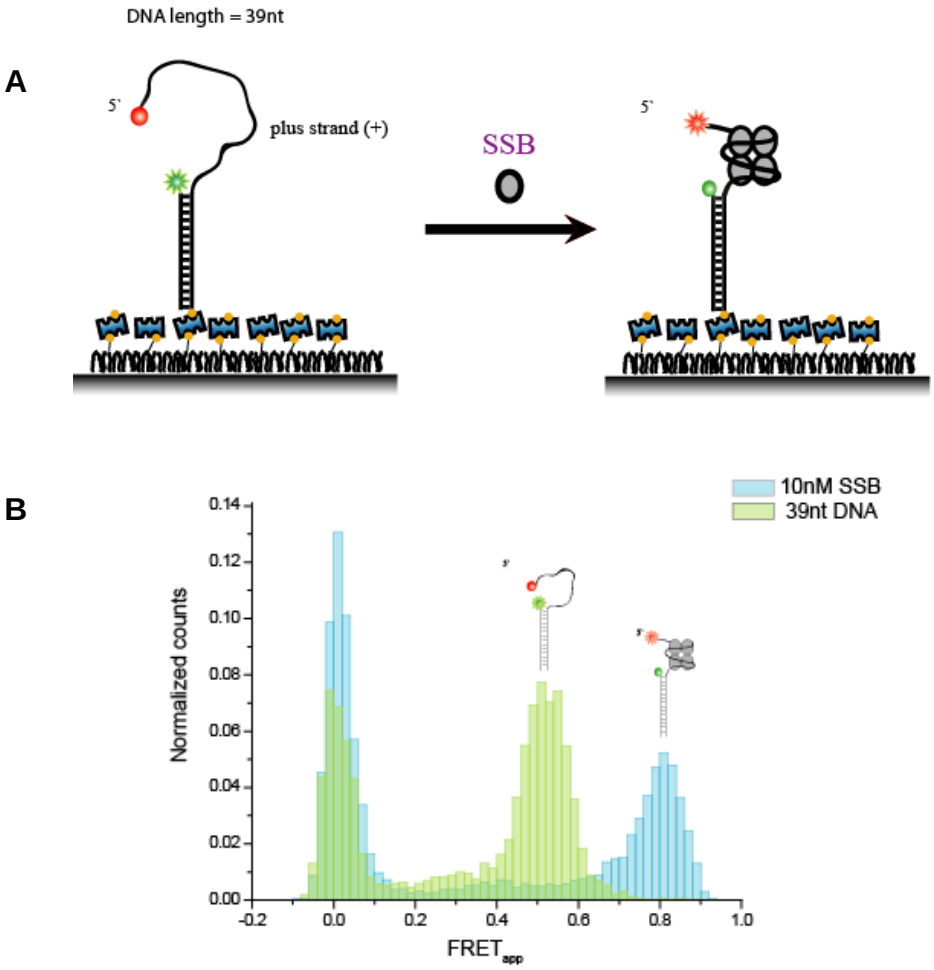
C.4 Effect of dye photobleaching on dissociation times

In Chapter 3, we estimated that the outgoing ssDNA remains bound for ~3s prior to

its release from the RecA filament. To test the contribution of dye photobleaching to the estimation of dissociation rates, we performed concurrent measurements of the binding time for DNA (39nt) at the secondary binding site in one imaging chamber and that of photobleaching time for a RecA filament coated DNA labeled with a donor fluorophore in a second chamber. The measurement was performed under identical conditions where the laser power remains unaltered. The photobleaching time for the donor under our solution conditions was ~ 23s which was 8 times slower than the dissociation time of ssDNA from the RecA secondary binding site (Figure C.5). A note of caution does emerge from this measurement. Estimation of binding time for longer DNA (80nt) bound to the RecA filament after completion of strand exchange may be underestimated by dye photobleaching prior to DNA dissociation.

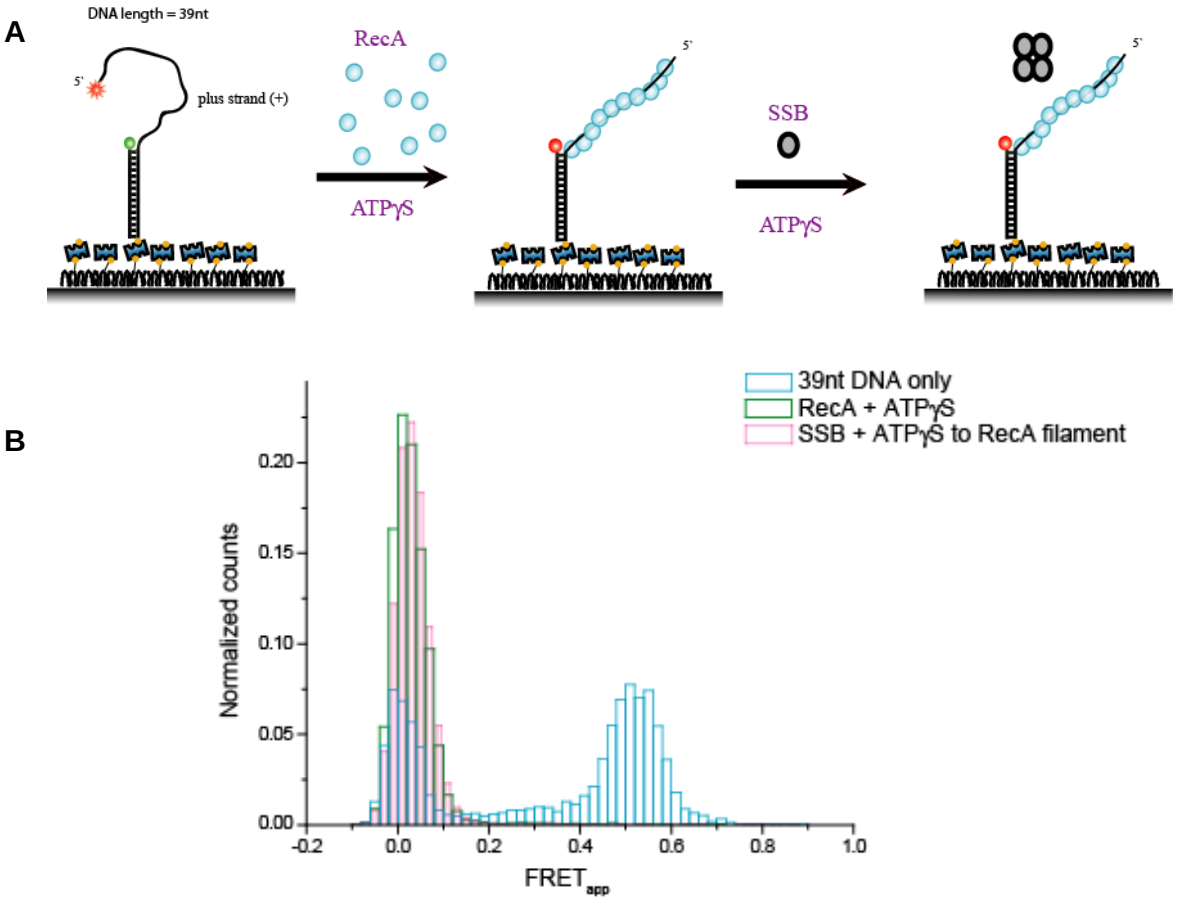
C.5 Figures

Figure C.1 SSB interaction with ssDNA



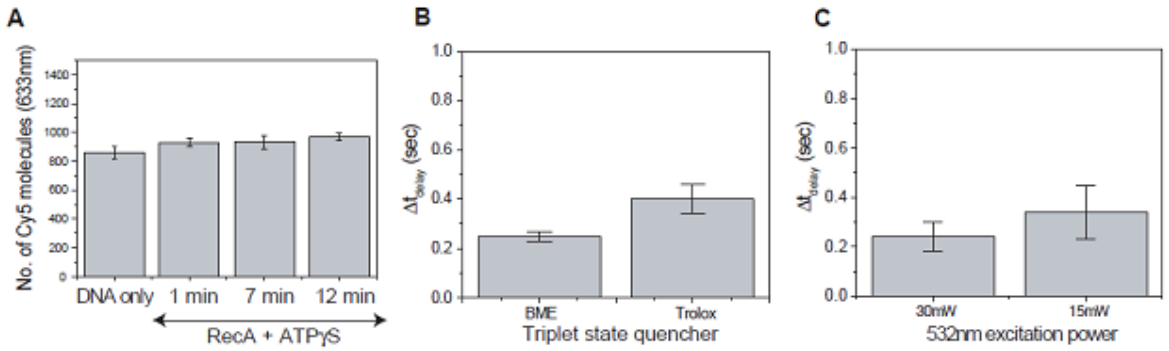
A. Schematic showing a 39nt ssDNA overhang with Cy3 at the terminal end of the ssDNA overhang and Cy5 present at the ssDNA/dsDNA junction position. SSB binding induces a tightly wrapped state of the ssDNA **B.** FRET histogram with and without SSB indicating FRET change upon SSB binding

Figure C.2 Stability of RecA filament during solution exchange



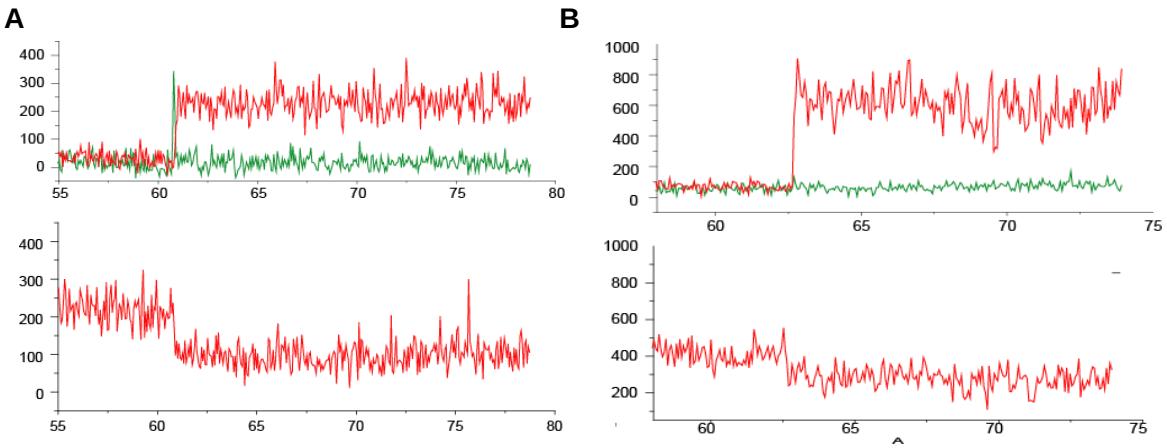
A. Schematic showing a 39nt ssDNA overhang with Cy3 at the terminal end of the ssDNA overhang and Cy5 present at the ssDNA/dsDNA junction position. RecA binding induces a stretched conformation of the ssDNA. SSB was added to challenge the RecA filament wherein SSB could compete for free ssDNA if RecA were to dissociate during solution exchange **B.** FRET histogram demonstrating the shift in FRET induced upon RecA binding relative to protein free ssDNA. The histogram remains unchanged upon altering solution conditions.

Figure C.3 Photophysical considerations during RecA strand exchange



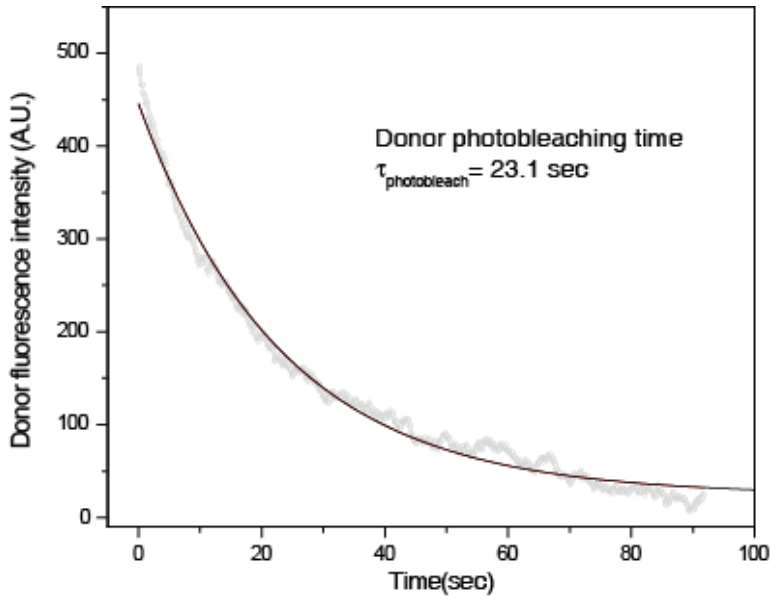
A. Number of acceptor labeled DNA (at 633nm excitation) vs. time after addition of RecA and ATPγS. Error bars are the standard deviation of the number of molecules from multiple CCD images. **B.** Comparison of mean initial low FRET delay time ($<\Delta t_{delay}>$) for strand exchange measurements with BME or Trolox as the triplet state quencher. Error bars are the standard error of the mean determined by bootstrapping. **C.** Comparison of initial low FRET delay time ($<\Delta t_{delay}>$) for strand exchange measurements at different laser excitation intensities (excitation wavelength= 532nm). Error bars are the standard error of the mean determined by bootstrapping.

Figure C.4 Alternating excitation allows simultaneous Cy5 detection during strand exchange reaction



A. Donor and acceptor intensities indicating the docking and subsequent completion of strand exchange reaction with a resulting high FRET state (top panel- 532nm excitation) Acceptor (Cy5) intensity upon direct excitation indicating the active state of Cy5 dye prior to and after strand exchange (bottom panel- 633nm excitation). **B.** Same as in A.

Figure C.5 Comparing the Cy3 photobleaching time to DNA dissociation from the heteroduplex product



Measurement of the photobleaching time of an immobilized donor fluorophore (Cy3) in the presence of a RecA filament bound to a reference ssDNA. The photobleaching time was 8 times slower than the dissociation of DNA from the heteroduplex product.