Convex Lens Induced Confinement: A Novel Study of Weak and Slow DNA & Protein Interactions

by

Ariana Joy Mann

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Author .....
Department of Physics
May 11, 2013

Certified by...
Leonid Mirny
Associate Professor
Thesis Supervisor

Accepted by .....
Nergis Mavalvala
Senior Thesis Coordinator, Department of Physics

Convex Lens Induced Confinement: A Novel Study of Weak and Slow DNA & Protein Interactions

by

Ariana Joy Mann

Submitted to the Department of Physics
on May 10, 2013, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

The newly developed single molecule microscopy method of flow cell-convex lens induced confinement (FC-CLIC) allows for the previously unobtainable probing of slow and weak DNA and protein interactions. Such measurements are made possible by CLIC's simultaneously improved signal-to-background ratio, observation time, observation volume, and permissible molecule concentration. The 10,000 increase in the observation time is demonstrated with the protein-protein interaction between actin filaments and formin. CLIC is used to generate data of the binding of a HMG protein to double-stranded DNA with a sufficient signal-to-background over long enough times for efficient particle tracking. The protein-DNA complex's diffusion coefficient is also extracted and compared with past measurements to successfully verify the accuracy of CLIC's measurements. Finally, the application of fluorescence cross correlation spectroscopy (FCCS) to the conformation fluctuations of DNA hairpins via Förester resonance energy transfer (FRET) is demonstrated. Progress is made towards the future use of FCCS and CLIC to determine the hairpin conformational reaction rates. The unimolecular hairpin is the preliminary molecular system for the study of electrostatic interactions in the alignment of DNA fragments, a new theory for the alignment of homologous chromosomes.

Thesis Supervisor: Leonid Mirny
Title: Associate Professor

Acknowledgments

No one individual can truly make a difference. There is always a myriad of others who helped, inspired, and allowed for the person to reach their achievements, from the most basic to the most awe inspiring. To start where I began, I would like to exude gratitude to my parents and my entire extended Mann, Rubin, Graubard, and most recently Preskill family. My arrival at MIT's Physics Department owes a thanks to Mr. Enrique Chee whose Socratic method high school Physics class both challenged me and introduced me to my adoration of understanding everything from first principles. Since then I would like to my classmates specifically Rachel Bandler, Adin Schmahmann, and Paul Lazarescu for there past four or more years of friendship and support that got me through MIT.

I would not be writing this if not for Adam E. Cohen's confidence in me and my abilities. From there Sabrina R. Leslie's trust in me not only allowed me to experience what ideal teamwork is like and vastly improve my research skills, she is the ultimate reason that I can say I accomplished something in my undergraduate time. I would also like to thank Leonid Mirny for his continuous support and guidance throughout this research.

I also owe my gratitude to the UROP program that made it possible for me to conduct this research while financing myself through college, especially the Paul E. Grey UROP Fund.

Contents

List of Figures

Chapter 1

Background and Theory

Underlying our genetic diversity is the homologous recombination of genes, a process whereby DNA segments with similar sequences are exchanged between two parental strands of DNA. Despite extensive research in genetics and molecular biology, the exact molecular mechanisms for how two pairs of homologous chromosomes properly align to exchange appropriate pieces of DNA are not well understood.

This research starts where molecular biologists left off. By combining new theoretical and experimental methods, my research examines the microscopic mechanisms underlying the interactions between DNA molecules. Previous research has shown that while a variety of proteins assist with the process of homologous recombination, it is likely that direct interactions between the DNA strands lead to the observed alignment of the chromosomes. The theory has been developed that DNA's sequence-dependent, surface charge distribution is responsible for it being electrostatically desirable for homologous sequences to bind and significantly less desirable for non-homologous stands to do likewise [2]. This theory has yet to be experimentally verified.

In this research, recently invented microscopy techniques are used to measure such weak sequence-dependent interactions that were unattainable with previously existing molecular methods. The imaging capabilities of CLIC are first demonstrated by studying protein-protein and protein-DNA interactions. By combining measurements of DNA interactions with theoretical tools, work is conducted to begin to differentiate

between alternate proposed mechanisms for these DNA interactions, and gain new insights into the process of homologous recombination. For the measurements of the interactions of DNA molecules convex lens induced confinement (CLIC) is used, a newly developed single-molecule microscopy method from the Cohen and Leslie labs. By applying Fluorescence Cross Correlation Spectroscopy (FCCS) techniques that rely upon the temporal and spatial correlations in the movement of interacting molecules as a sensitive probe of their interactions, the expectation is to be able to detect weak sequence-dependent interactions between strands of DNA by analyzing the correlations of their free diffusion in the CLIC device.

1.1 Single Molecule Imaging

Previous experiments have attempted to measure DNA-DNA interactions, but they were heavily constrained by issues of background fluorescence, diffusion-limited observation time, and sample concentrations. These specific issues are key to the success of single-molecule spectroscopy, especially to the observation of slow and weak interactions.

In order to reliably measure single fluorophores, the standard requirement is a maximum average of one fluorophore per detection volume. Such a tight constraint causes additional background fluorescence to quickly becomes an obstacle. Total internal reflection fluorescence (TIRF), now a standard method, has provided notable success in solving issues of background fluorescence. It works to utilize the quality of evanescent waves, which decay exponentially in intensity, to selectively illuminate only the top layer of the detection volume next to the surface [3]. Unfortunately, by shrinking the detection volume it takes less time for molecules to diffuse out of view. The reduction in observation time greatly hinders what TIRF and single molecule microscopy have been able to accomplish.

Long observation times are particularly important for slow molecular processes; often the process is actually longer than the observation time for the molecule or molecular complex. This is an issue not only for methods such as TIRF, but especially

for confocal microscopy. Confocal microscopy, specifically when used for correlation spectroscopy, is designed to measure temporal fluctuations instead of the usual spatial fluctuations [4]. If the time interval per molecule is too small then no meaningful information is captured. However, if the detection volume is increased then the background fluorescence is too high and again no meaningful data is recorded [5].

The limit on fluorophore resolution per volume also imposes a limit on the concentration of molecules. Specifically for weak interactions, concentrations on the order of micro-molars are required. Yet, confocal is only effective for concentrations that do not exceed 10nM, and TIRF only up to 100nM [5].

The standard methodologies used, most prominently total internal reflection fluorescence (TIRF) imaging and confocal microscopy, are limited in the key areas of background fluorescence, observation volume, observation time, and sample concentration. CLIC provides solutions and opportunities for future, previously unobtainable experimental results.

1.1.1 CLIC

Given the increased interest in studying both slow and weak interactions, new methods were developed to attempt to fill the gaps of the previous, constrained ones. Convex lens induced confinement (CLIC) is one of those new systems. CLIC enables direct imaging of a thin fluid film of fluorescently labeled, freely diffusing molecules. The device significantly expands the allowed and experimentally viable range of detection volume, molecule concentration, and observation time. At the same time, it is a remarkably simple setup that can be easily integrated into laboratory research without high costs or complicated technology.

CLIC utilizes the natural curvature of a plano-convex lens to contort the top planar coverslip of a flow-cell to form a well-defined, nanoscale gap. The gap ranges radially in height from several microns down to 2nm where it is limited by the roughness of the coverslips' surfaces at the contact point. The vertical confinement serves several purposes. The reduced height of the imaging field naturally reduces the imaging volume, which in turn minimizes the background fluorescence. Previous experiments with

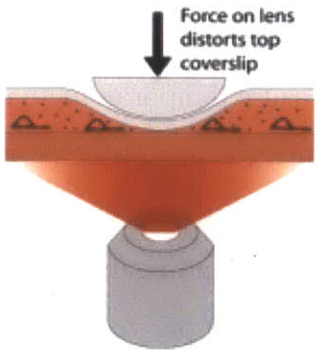

Figure 1-1: The general setup of CLIC is depicted to highlight the main features of the system. The top coverslip is conformed to the shape of the applied lens to form a radially varying, vertical gap in the sample-filled flow cell [6].

CLIC have demonstrated a 20-fold reduction in background fluorescence compared to TIRF and 200-fold compared to confocal microscopy [5]. The reduced background fluorescence permits measurements of weak intermolecular reactions to be performed at the required micro-molar concentrations. CLIC has been shown to be effective for concentrations of up to $2\mu M$[6]. The observation of weak molecular reactions is possible with CLIC's high concentration tolerance caused by the system's vertical confinement.

Another consequence of the restricted gap is that the height of the flow-cell is less than the focal depth of the optical system. The observation time for freely diffusing molecules in CLIC is therefore only limited by lateral diffusion and photobleaching. The elimination of undetectable vertical diffusion results in a 10^4-fold increase in observation times for CLIC in comparison to both TIRF and confocal microscopy [5]. The extended observation time makes it possible to observe slow molecular reactions. Other methods to solve the problem of molecules diffusing out of the field of view, accomplish the goal via the immobilization of the molecules to the surface. However, immobilization easily and dramatically interferes with molecular function.

The point may be raised that CLIC's confinement would likewise interfere with the function of the observed molecules and processes. However, confinement of varying extents is an inherent feature of intracellular processes. CLIC not only solves problems with previous microscopy methods, but it inherently provides a system to study

the biologically relevant confinement effects. The molecules of each sample can be observed at different gap heights, and the effect of the vertical confinement can be observed and analyzed. By recording the interference pattern of the lens system, a precise gap profile can be determined for use in the analysis of confinement.

These four notable advances of CLIC over traditional confocal microscopy make reliable measurements of weak sequence-dependent DNA-DNA interactions plausible for the first time and offer additional improvements for the easier collection of more meaningful data.

CLIC allows for the image capture of a variety of biological systems. It was recently utilized in the study of myosin and has the potential to be applied to many future protein systems [6]. This research is decidingly focused however on DNA-DNA interactions. For the first endeavor of its type with this new system, the most basic of DNA-DNA interactions was selected: the interaction between the two tails of DNA hairpins. For this especially weak interaction dealing with particularly small molecules of around 50 base pairs, fluorescence cross-correlation spectroscopy (FCCS) was utilized as an incredibly useful analysis technique.

FCCS is typically performed via confocal microscopy, where the temporal fluctuations over one point in space are correlated and analyzed. Here again, CLIC's full field view provides a clear improvement. Now both temporal and spatial data can be collected and analyzed. The increased data provides for better statistics.

1.2 Biological Systems

1.2.1 Actin Filaments

Actin is one of the microfilaments that are essential to the eukaryotic cell's cytoskeleton. It is a protein that can exist in two states: G-actin or F-actin. G-actin is the monomer form and is a globular protein. F-actin is the linear polymer of G-actin molecules. Stands of F-actin rotate to properly align and acquire the double-helix structure that serves as a microfilament [7]. Strands of actin are characterized by a

negatively charged, or pointed, end and a positively charged, or barbed, end. G-actin cannot be generally added to the pointed end of a F-actin filament. At the same time there are a series of protein that bind to the barbed end, capping it and prevents further elongation [8].

Actin also serves as an ATPase, an enzyme that hydrolyzes ATP. Bound to each G-actin monomer is a nucleotide of adenosine triphosphate (ATP), the chemical energy source and workhorse of the cell. The molecule of ATP is hydrolyzed to the lower free energy state of adenosine diphosphate (ADP). The hydrolyzed state of the nucleotide also affects the confirmation of the actin protein. It therefore affects the binding affinity of one G-actin to another, and the rate of actin polymerization. This makes the rate equations for actin's polymerization and depolymerization notably complex as derived in [7].

Formins are a different set of homodimer proteins. They are known for their interaction with actin and its polymerization. Formin has been shown to bind and remain continuously bound to the barbed end of filaments to allow for the uninterrupted polymerization of actin. Not only does it prevent the capping of the actin strand, but it has also been shown to cause up to a five time increase in the rate of polymerization. However, the mechanism behind the interaction of formin and actin is not known [8].

Essential for accurate and usable data on the polymerization of F-actin filaments is not only high resolution, but also an extended detection time with a large detection volume. The volume allows for the collection of data for a long filament, and the long detection time allows for the accumulation of data for the polymerization of longer strands. Currently, TIRF imaging has been used since it provides the necessary time lapse. Formin is noted for being particularly difficult for to work with experimentally. In vitro, formin successfully elongate filaments while in bulk solution. However, once moved onto the surface for TIRF imaging the filaments shear. CLIC offers the advantages of a long detection time, large detection volume, and a varying gap height that allows the molecules to be observed away from a surface. The imaging of actin filaments and their polymerization with and without formin with CLIC should

not only provide previously unacquirable data and science, but also demonstrate the capabilities of CLIC system.

1.2.2 HMGB Proteins

High motility group (HMG) proteins are found at levels of 10^6 per eukaryotic cell. In general, they bind to DNA and thereby change its confirmation. HMG boxes are the common structure unit of the protein classification and it is comprised of three helical proteins and a tail that forms hydrophobic bonds with nearby proteins for stabilization [9]. One of the major subgroups is the HMG group B proteins.

HMGB proteins are characterized by their non-sequence specific binding to DNA through one or two HMG boxes. The proteins' physical effect on double-stranded DNA has been characterized with optical tweezers and the corresponding force-extension curves. Typical measures of DNA were taken, namely the persistence and contour lengths. HMGB appears to bind in the DNA's minor groove to cause a measurable increase in contour length, decreases in persistence length, and bends in the DNA. The binding of HMBG to DNA corresponds to an increased flexibility of the DNA [9]. They appear to induce and stabilize DNA loops, and are predicted to contribute to the binding of transcription factors [10]. However, the mechanisms behind HMGB's effect on helical DNA are unknown.

Previous experiments have been hindered by insufficient observation times. Both the loops and flexibility that occur as a result of the HMGB proteins should effect the shape of the corresponding DNA molecules, and affect the diffusion and binding rates of the molecules. These two variables are proposed to be measurable by CLIC, which also possesses the advantage of an extended observation time. The diffusion coefficient D has a stronger dependence on the contour length L of DNA, $D = L^{-v}$ where $v = 0.589$ for the helical form of DNA [11]. By measuring the molecules' diffusion coefficients with CLIC it will also be possible to verify the accuracy of CLIC's measurements in comparison to those of the optical tweezers.

1.2.3 DNA Hairpin

Homologous Recombination

Homologous recombination contributes significantly to the biological diversity of nature. As one of the most important intracellular processes, it is surprising how little it is understood. The only accepted mechanism for the sequence-based recognition of nucleic acids is for complementary single strands of DNA. It is now commonly accepted that the pairing of DNA fragments in the first step of homologous recombination is carried out between complete, double-stranded DNA strands [12].

The theory developed by Kornyshev and Leikin looks to electrostatic dynamics to model the alignment of double stranded DNA such as in homologous recombination. A DNA strand is a sequence constructed from four different nucleobases, which divide into two sets of possible pairs. While each base has a different chemical structure, they each bind to a negatively charged phosphate group thereby form a nucleotide. The phosphate groups bind the nucleotides together and form the backbone of the DNA strand. Due to the helical shape of double-stranded DNA, the location of the phosphate groups also rotates about the azimuthal direction. The rotation leads to negatively charged grooves along the DNA molecule.

The vertical spacing between each base-pair is determined by the particular nucleotides that compose the pair; it is sequence dependent. The location of the phosphate groups directly correlates with the spacing of the bases. Therefore, in DNA's helical form the location of the negative phosphate groups along any vertical line is a function of the sequence. For segments similar over an extended length, the negative grooves of one strand align with the positive grooves of the other strand and vice-versa. Due to the length of sequence similarity the spacing of positive and negative grooves is consistent over a long enough stretch to form an attraction electrostatically favorable enough for the strands to remain aligned. Long fragments of uncorrelated sequences in comparison, do not align because of the repulsion between the mismatched positive and negative grooves [2].

The sequence dependence of double-stranded DNA as a result of electrostatic dy-

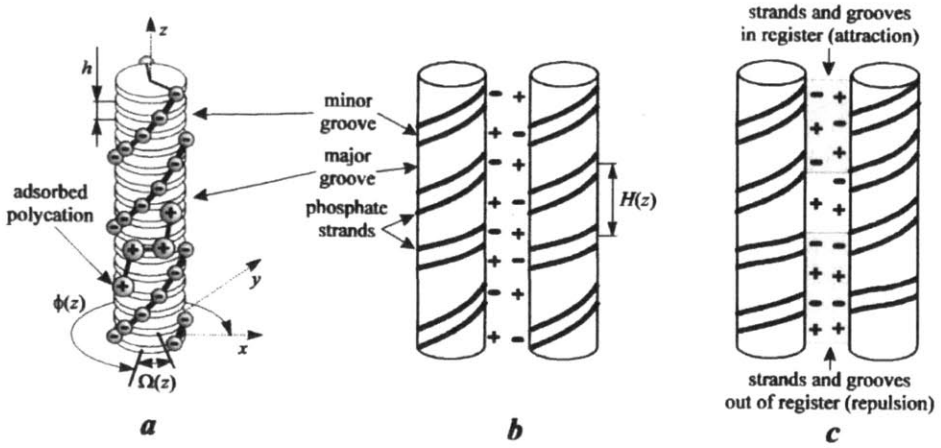

Figure 1-2: (a) Double stranded DNA is depicted here with each base-pair as a disc with vertical separation h. The negatively charged phosphate groups, one per base and two per pair, are shown and highlight the helical orientation of DNA around the azimuthal angle $\phi(z)$. The twist per base pair is $\Omega(z)$, which is on average $34° \pm 5°$. In addition to the negative grooves from the phosphate groups, polycations often bind to DNA and are included in the model. (b) Because of $\Omega(z)$'s sequence dependence, the helical pitch $H(z)$ also varies with the strand's sequence. The result is that only strands with significant stretches of similar sequences can bind by aligning the positive and negative grooves. (c) The strand segments for which is it energetically viable to align are electrostatically attracted, while the strands with mismatched sequences are repulsed and unable to align for homologous recombination [2].

namics can be applied to explain how homologous recombination occurs only between the same genes of each strand surrounded by a generally similar sequence, and not between random genes that would lead to a nonsensical sequence. The energy of the electrostatic interaction is on the order such that the sequences have to be homologous on a scale of 50-200 base-pairs, such that Kornyshev's theory accounts for the initial pairing of bound and twisted DNA strands prior to homologous recombination [2]. The enzymes known to be involved in homologous recombination then step in to match the DNA precisely on the order of 10 base-pair and facilitate the actual exchange of genetic information.

1.2.4 Hairpin

Although DNA is most well known for its double-stranded, helical form, it is also found in different single-strand structures. One such example is the DNA hairpin system, Fig. 1-3. The standard hairpin naturally fluctuates between open and close confirmations. The open confirmation is simply a single-strand of DNA. The closed confirmation consists of an unbound loop and a stem formed by the binding of complementary base pairs in the single-strand. DNA hairpins have been shown to contribute to such biological processes as gene expression, and in line with the focus of this research, DNA recombination. The open and closed confirmations are subject

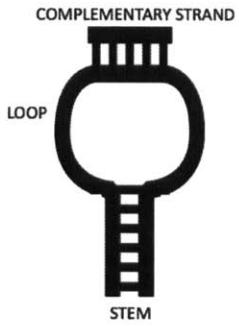

Figure 1-3: DNA hairpin with annealed complementary strand to the loop portion [1].

to different energy constraints. The open state has a high number of possible orientations, Ω. By the standard thermodynamic relation, entropy is proportion to $\ln \Omega$ and therefore the open state has a high entropy. The energy required to transition from open to closed is that required for the bending of the strand into the proper alignment of base pairs, followed by the energy for nucleation and stretching caused by base pairing. In comparison, the closed state due to the bound base pairs has a low enthalpy. To reconfigure into the open form from the closed one, the energy required is that to break the bonds between each of the base pairs.

The dependence of each direction of the conformational transition has been predicted and experimentally verified. The rate k_- to transition to the open form is not significantly affected by the specific hairpin's properties. However, the closing rate is

affected by the loop's length, the dominant type of base, and the salt concentration. The sequence length of the loop decreases the closing rate by a factor of approximately exp 2.6. Likewise, a loop of the adenosine base repeats has an activation energy five times that of a loop with a repetition of the more rigid thymine bases, which also notably decreases the k_+ rate. The one surprising effect, is that of salt concentration. Charges from the salt do not screen charges to increase the likelihood of base pair bonding. Instead, an higher salt concentration, led to a higher activation energy for the closing of the DNA hairpin and thus a higher k_+. Screening did however slightly decrease the unzipping rate k_- of the DNA [13].

One additional conformational change can be applied to the DNA hairpin system. Normally, the loop is unrestrained. If however a complementary sequence is annealed to a section of the loop (Fig. 1-3) the likelihood of the hairpin being in the open verses the closed confirmation will shift. If the sequence is sufficiently short then its presence should not affect the energetic incentives at the stem. When the complementary strand is sufficiently long however, then it will exert a pressure on the stem and increase its likelihood of bonding closed [1].

1.3 Analysis Methods

1.3.1 Particle Tracking

The concept of particle tracking is very simple. There are particles that are visually distinct from the rest of the image and their locations can be algorithmically determined by applying masks to select for their distinct feature. The locations of particles are then extracted for each frame in the movie. On a per particle basis, the location in each successive frame can be strung together to produce the trajectory.

While conceptually straight-forward there are numerous variables that can affect the ease and accuracy of the trajectories. For that reason, particle tracking code repositories are available for academic use and to encourage collaboration for the improvement of the codes. For this research, we used a MATLAB adaption of the

widely used IDL Particle Tracking software. The MATLAB repository comes from [14] and the original IDL software is available at [15].

Trajectory Acquisition

The variables of brightness and size are used to select the particles from the background image. Standard fluorescently tagged particles appear as a gaussian-like ovular bright region. Several masks are applied to differentiate the bright particles from noise, dirt, and the background. First, the image is normalized, and then the median image is subtracted. A rotationally symmetric Gaussian lowpass filter is applied to the background subtracted image. The mean and standard deviation values of the filtered image are then used to calculate a threshold value for the image. The threshold is equal to the mean plus a factor of the the standard deviations, where the factor can be adjusted based on the particular sample's noise levels. The extent of the gaussian filter can also be adjusted. A mask is created to select the regions where the gaussian filtered image has a greater intensity then the threshold. This process is performed on each image frame of a recorded movie.

Peaks corresponding to particles are extracted from the masked frames. Adjustable input parameters for the size of the particle and the threshold brightness are entered. Pixels are located above the threshold. Using the inputed expected molecule size, any peaks within that size of another peak are removed and the brightest peak is retained. This should hopefully prevent multiple selections of the same particle. The centroid of each of the peaks is calculated next via the selection of a small window around each peak. The total brightness within the window is calculated and used to normalize the region in the x and y directions. This allows for the calculation of the center and the radius that define the centroid peak. The centroid data is accumulated for each frame in order to create the trajectories.

Trajectories are calculated by minimizing displacement between frames. From one frame to the next there are n particles at time $t(i)$ each with an identified position, that move to m particles at time $t(i+1)$ with new positions. The total squared displacement is calculated for each combination of particles peaks between each frame,

and the arrangement that minimizes the calculation is selected. This method should produce the most probable set of particle identifications and trajectories.

Several important restrictions are applied in the post-processing of the accumulated trajectories. A minimum number of frames is set, where any trajectories of peaks that appeared for less than the minimum number of consecutive frames is not included. The possibility of a particle disappearing briefly is also included. Such disappearances may occur if the particle is near the edge of the image and goes out of view for a limited number of frames, or crosses paths with another potentially brighter particle. This consideration will allow the two disjoint trajectories that are in close proximity but do not have a temporal overlap to be combined into one final trajectory. Finally, a minimum trajectory distance can be set to remove the tracking of particles stuck to the surface from the analysis.

Trajectory Analysis

High quality trajectories can be subjected to a number of valuable calculations. A standard, important one featured in this research is for the diffusion coefficient of particles. Most simply, the displacement Δs between each frame of the trajectory of each particle is calculated and the diffusion coefficient is,

$$D = \frac{(\Delta s)^2}{\Delta t} \tag{1.1}$$

after the displacement is converted from pixels to physical spatial units. To determine the characteristic diffusion coefficient for a particular particle type this calculation is improved upon by taking the average displacement between frames for each particle and then the average displacement of all the particles. Finally, the average between particles can be weighted for accuracy by weighting the contribution of the each particle by the number of frames over which it is observed.

Confinement effects on the diffusion coefficient can be studied by comparing each particle's average diffusion coefficient to its position and the corresponding gap height of CLIC. Further applications also include more sophisticated analysis such as reaction

binding coefficients. These are calculated by tracking the binding and unbinding of two differently labeled molecules with a dual-imaging system.

1.3.2 Correlation Spectroscopy

Fluorescence Correlation Spectroscopy (FCS) was developed in the 1970's as a combination of two key methods: fluorescence spectroscopy and relaxation analysis. Fluorescence spectroscopy was itself developed as a highly sensitive, non-invasive single molecule spectroscopy technique. Experimental relaxation methods involve the application of a perturbation to a macroscopic variable and the observation of the system's return to equilibrium. From those observations, information about the kinetic parameters of the system are can be analytically obtained [4].

FCS utilizes the fluctuations in molecules' fluorescence emission corresponding to the spontaneous fluctuations of the molecules' physical parameters. The decay of spontaneous fluctuations is described by the same physical constants that also determine the rate of return to equilibrium that follows the external variable's perturbation. The recorded emission intensity is proportional to the number of molecules contained in the detection volume, and therefore the seemingly random fluorescence fluctuations are in fact caused by changes in the molecule population [16]. Physical constants the correspond to fluorescence fluctuation and are thus measurable by FCS include diffusion coefficients, reactions kinetics, rotational dynamics, and excited state lifetimes [17].

The principles of the fluorescence spectroscopy procedure are as follows. The appropriate wavelength laser is setup and focused on the sample such that the detection volume is the part of the sample contained within the focus of the laser. Molecules in the detection volume are excited by the laser and emit red-shifted fluorescence. The light emitted from the excited molecules is separated from the excitation light via a dichroic filter, travels through a pinhole that selects for light from the focal region, and is recorded with a photodetector [17]. The fluctuations in molecule population due to diffusion or reactions are transmitted as fluctuations in the emission intensity. The fact that because different molecules are excited by and emit at different wave-

lengths, means that fluorescence microscopy has the particularly useful characteristic of chemical specificity.

A noteworthy point of consideration is that since the fluctuations are a measurement of molecule population, to collect precise data the number or concentration of molecules must be low enough that a shift in the population by one molecule produces a significant, observable fluctuation. FCS was developed into an effective method when it was combined with confocal microscopy. Confocal microscopy records temporal fluctuations over a diffraction limited spatial region. This minimization of the detection volume delivers a low enough molecule number to maintain good signal-to-noise ratios and meaningful data [4].

CLIC offers a unique opportunity for FCS. The confinement of the CLIC system is able to sufficiently limit the detection volume to produce good signal-to-noise ratios over a wide field view. The opportunity for the analysis of spatial as well as temporal fluctuations provides the opportunity for more data that will lead to more significant conclusions.

Autocorrelation

For normal fluorescence spectroscopy, the fluorescence fluctuations that are the focus of FCS are instead categorized as undesired noise in the system. In FCS the proper analysis tools are utilized to extract important information from the so-called noise. The trick of FCS is the autocorrelation function (1.2), from which FCS earns its name. The correlation of the fluorescence intensity with itself, or the autocorrelation, measures the degree of temporal similarity in the signal. In this case, that measurement corresponds to the strength and duration of the fluctuations [4].

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t+\tau) \rangle}{\langle I \rangle^2}, \tag{1.2}$$

where $F(t)$ is the temporal average of the signal, and $\delta F(t) = F(t) - \langle F(t) \rangle$. A sample autocorrelation curve and its construction is displayed in Fig. 1-4. For example, if two molecules attract one another, the detection of one molecule increases the chance

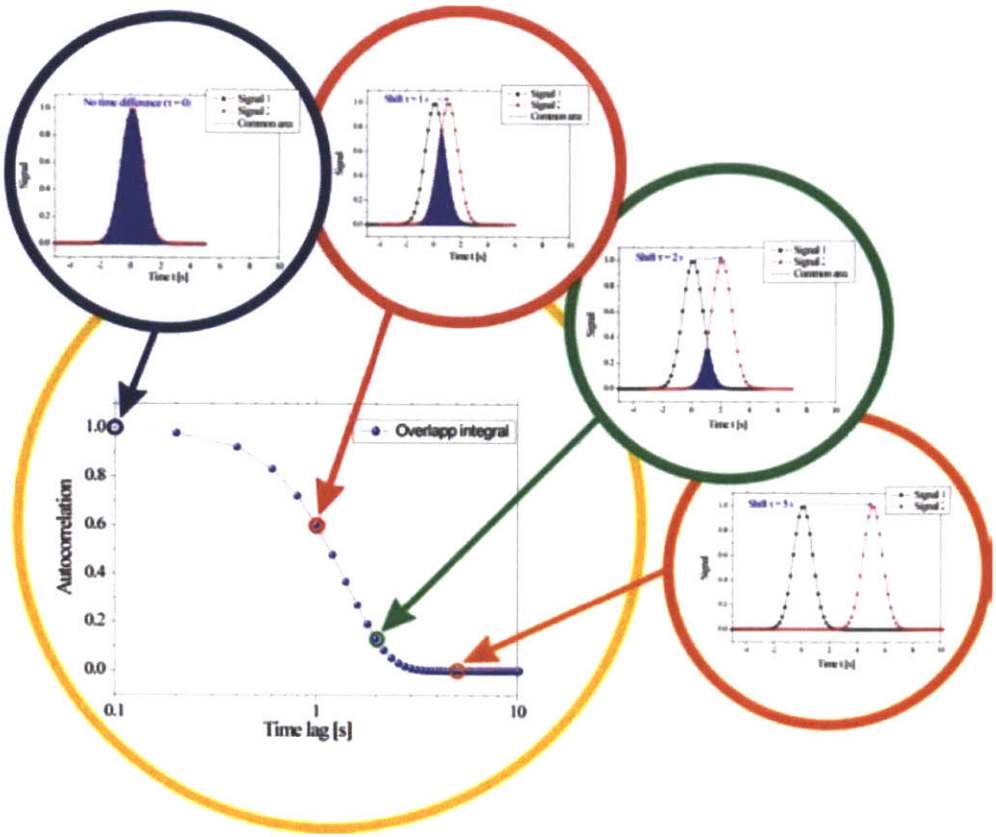

Figure 1-4: The generation of an autocorrelation curve is illustrated. The signals at different times are plotted and the overlap integral calculated as the autocorrelation that represents the similarity between the signal at the two times. This is repeated across all time to accumulate the complete autocorrelation curve [4].

of detecting another molecule a short time later, at the same location. This will be reflected by a higher value in the autocorrelation function at those times. For fluctuations caused only by changes in molecule concentration

$$\delta F(t) = \kappa \int_V I_{ex}(r) \cdot S(r) \cdot \delta(\sigma \cdot q \cdot C(r,t)) \cot dV. \tag{1.3}$$

variable meanings... $I_{ex}(r) \cdot S(r)$ are approximated by the three-dimensional Gaussian

$$W(r) = \exp -2\frac{x^2+y^2}{r_0^2} \cdot \exp -2\frac{z^2}{z_0} \tag{1.4}$$

that characterizes the emitted light's spatial distribution, assuming a decay of $1/e^2$ at the characteristic lengths r_0 and z_0 in the appropriate directions.

For free diffusion with a diffusion coefficient D, the characteristic decay time of $G(\tau)$ is $\tau_D = \frac{r_0^2}{4D}$. Also important to the fluctuation and concentration analysis is the focal volume and effective detection volume,

$$V_{eff} = \frac{(\int W(r)dV)^2}{\int W^2(r)dV}. \tag{1.5}$$

The common assumption is that the molecular reaction time-scale is significantly shorter than the time-scale of diffusion, such that separation of the dynamics in (1.6) is plausible [4].

$$G_{total}(\tau) = G_{motion}(\tau) \cdot X_{kinetics}(\tau). \tag{1.6}$$

FRET

Fluorescence or Förster resonance energy transfer (FRET) is the electronic excitation transfer from a donor to an acceptor molecule. It is a nonradiative transfer that does not include the emission or re-absorption of light. The interaction between the weakly coupled donor and acceptor is characterized as a dipole-dipole interaction. There are four electronic states at play, namely the ground and excited states of both the donor and the acceptor.

The typical case involves the excitation of the donor's ground state. Once excited, the donor coherently oscillates in its excited state energy gap. FRET earns its name from the required resonance condition for the energy transfer. Specifically, this donor excited state oscillation is resonant with the acceptor's ground state electronic energy gap. Upon the transfer of energy, the donor vibrationally relaxes, and the acceptor emits fluorescence before also vibrationally relaxing. The fluorescence of the acceptor is shifted from the donor fluorescence. Inherent in the expression for the energy transfer rate of FRET are correlation functions for both the donor emission and the acceptor absorption. They are derived from the action of dipole temporal correlation function on the donor's excited state and the acceptor's ground state. For the donor

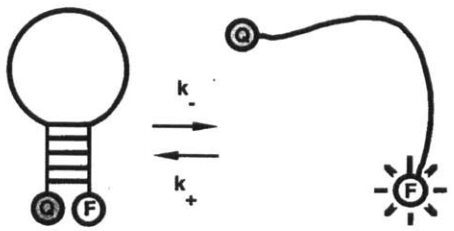

Figure 1-5: A common and basic FRET system is the DNA hairpin. The complementary stem alternates between the closed and open configurations with the characteristic forward and backwards rates, k_- and k_+ respectively. In the open state the fluorophore on one end emits fluorescence. When the fluorophore and quencher within a short enough proximity for energy transfer then the fluorophore's emission is quenched and the hairpin is in the closed state [13].

state $|D\rangle$, the acceptor state $|A\rangle$, and the corresponding dipole operator μ, the energy transfer rate is proportionally

$$w_{ET}\alpha \int dt \frac{1}{r^6}\langle D^*|\mu_D(t)\mu_D(0)|D^*\rangle\langle A|\mu_A(t)\mu_A(0)|A\rangle. \tag{1.7}$$

A key feature in this equation, is the energy transfer rate's dependence on r^{-6}, where r is the distance between the donor and acceptor molecules. It is often utilized in spectroscopy as a measurement of the distance between the donor and acceptor, or more specifically in this research as a measure of the macromolecules' confirmations [18].

Cross Correlation

Fluorescence cross-correlation spectroscopy (FCCS) utilizes the specificity of fluorescence spectroscopy to extract information about molecular dynamics. The conformational shifts of a molecules can be linked to changes in emission wavelengths using methods such as FRET as explained in Section 1.3.2. For the appropriate system, the fluorescence fluctuations will therefore be caused not only by diffusion, but also by changes in the molecule's population of its different conformational states. The reaction rate of the conformational change can then be measured with a two-channel imaging setup and an analysis that cross-correlates the signals [4].

Conceptually, cross-correlation evaluates the similarity in signals from different volumes or spatial locations. It is a generalization of FCS's autocorrelation. FCCS is particularly effective at measuring characteristic movement because the average time for a molecule to transfer from one measured location to another is naturally highlighted in the cross-correlated signals.

1.3.3 Unimolecular System

For a unimolecular system the alternates between two conformations A and B with a forward rate k_-, a backwards rate k_+, a diffusion coefficients D_A and D_B, wave vector q the concentration fluctuation expressions are derived as in [16] from the matrix,

$$\begin{pmatrix} -(D_A q^2 + k_-) & k_+ \\ k_- & -(D_b q^2 + k_+ \end{pmatrix}. \tag{1.8}$$

The resulting eigenvalues are,

$$\lambda_1 = -\frac{1}{2}[q^2(D_A + D_B) + R] - \frac{1}{2}[q^4\Delta^2 + 2q^2(k_+ - k_-] + R^2)^{1/2} \tag{1.9}$$

$$\lambda_2 = -\frac{1}{2}[q^2(D_A + D_B) + R] + \frac{1}{2}[q^4\Delta^2 + 2q^2(k_+ - k_-] + R^2)^{1/2} \tag{1.10}$$

where $\Delta = D_A - D_B$ and $R = k_+ + k_-$. The corresponding eigenfunctions are,

$$F_1 = (q^2 D_A + k_- + \lambda_1)\frac{1}{k_+} \tag{1.11}$$

$$F_2 = (q^2 D_A + k_- + \lambda_2)\frac{1}{k_+ s} \tag{1.12}$$

$$\tag{1.13}$$

. When the assumption that $D = D_A = D_B$ is made the resulting two autocorrelation functions and the cross correlation function are,

$$G_{AA} = \frac{1 + ke^{-Rt}}{1 + k} e^{-q^2 Dt} \tag{1.14}$$

$$G_{BB} = \frac{k + e^{-Rt}}{1 + k} e^{-q^2 Dt} \tag{1.15}$$

$$G_{AB} = \frac{1 - e^{-Rt}}{1 + k} e^{-q^2 Dt} \tag{1.16}$$

$$\tag{1.17}$$

for $k = k_+/k_-$ [19]. From these functions characteristic constants that describe the dynamics of the unimolecular system can be extracted. The diffusion coefficient D is equal to the slope of $\log G_{ij}$ verses q^2 for $i, j = \{A, B\}$.

The reaction rates can be determined by extracting both k and r to then solve for k_+ and k_-. By examining only the $q = 0$ component of G_{ij}, the exponentially factor is removed. It is important to note that for $q = 0$, $\log G_{AA}$ and $\log G_{BB}$ decrease with time while $\log G_{AB}$ increases with time before reaching a plateau. The ratios between the $G_{ij}(q = 0)$ functions can be taken, resulting in the following equations in the long time limit,

$$\frac{G_{AA}}{G_{BB}} = \frac{1}{1 + k} \tag{1.18}$$

$$\frac{G_{AB}}{G_{BB}} = k \tag{1.19}$$

$$G_{AB} = \frac{k}{1 + k}. \tag{1.20}$$

$$\tag{1.21}$$

Finally, $\log(k + 1)G_{AB} - k$ simplifies to $-Rt$. Therefore, R is equal to the slope of $\log(k+1)G_{AB} - k$ verses t. The forward and backwards reaction rates are then solved for from the system of equations for k and R.

By applying FCCS to a unimolecular system such as the DNA hairpin it has been derived how to extract the constants that characterize the dynamics of the system, namely its diffusion and rates of conformational shifts.

Chapter 2

CLIC System and Characterization

This research uses different biological single molecule systems to study the ability of the newly developed convex lens induced confinement (CLIC) imaging system to measure fast and weak reactions over long observation times. The experiments conducted with CLIC contain two distinct levels of setup: the Flow Cell-CLIC system and the molecular system.

2.1 System Components

The CLIC apparatus is as seen in Fig. 2-1. The hinge system and cantilever controls the position of the lens (Thorlabs LA4966) as the lens is mounted in the cantilever. The flow-cell is placed and centered under where the lens will lie when the cantilever is lowered. Once in place and the cantilever lowered, the finely $1/4-80$ threaded rod (Thorlabs F25US200) is placed through a hole in the cantilever. Around the rod is a spring (Small Parts CSXX-0160-05) and then a finely threaded lock-nut.

The general Flow Cell-CLIC device is used here with an inverted epifluorescence microscope (Olympus IX71) and a 60x oil-immersion objective (N.A. 1.45, Olympus 1-U2B616). A XYZ translation stage (Thorlabs DT12) is incorporated into the microscope for precise, automated position control. Three lasers were aligned for use with the microscope and CLIC: a blue 647nm, a green 488nm, and a red 633nm HeNe laser. The alignment is such that the dual imaging of the green and either of the other

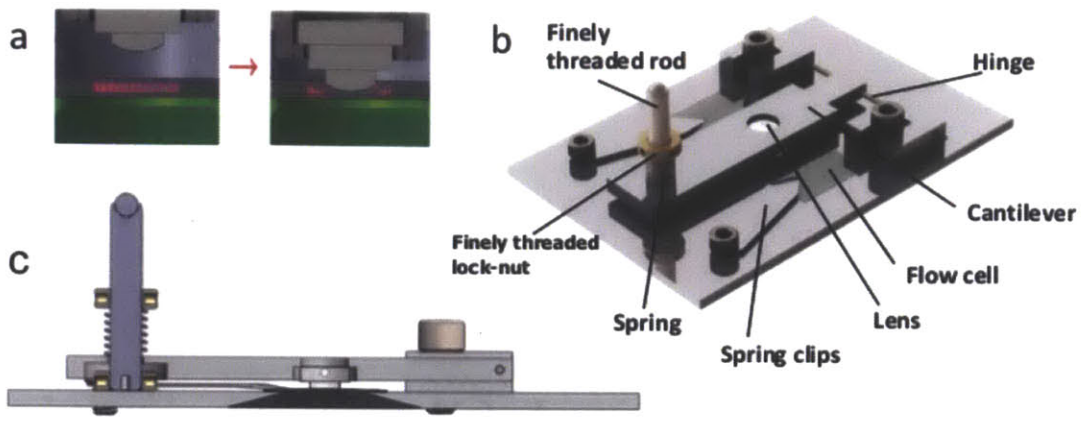

Figure 2-1: (a) The general CLIC system is seen with the lens lifted (left) above the flow-cell containing a sample of fluorophores (pink) excited from below by the green laser. The lens is lowered (right) until the top coverslip of the flow-cell makes contact with the bottom coverslip to form CLIC's characteristic gap for improved imaging and data collection. (b) A top-view of CLIC is presented, which includes the entire system that is placed in an inverse microscope. (c) A side view of the complete system with the lens in its lowered position. The darkened section of the underside illustrates the section of the plate carved out for microscope's objective [6].

two lasers can occur. The dual imaging system is constructed by splitting the output from the microscope into two identical beams that are sent through the appropriate emission filters to select for the desired wavelength. The two filtered images are simultaneously recorded at the cost of decreasing the field of view by half its area. A EMCCD camera (Andor iXon3 897) with 0.45μm/pixel was used to capture all the data movies.

Flow cells were constructed either from glass or from quartz coverslips. Prior to each experiment, the coverslips were cleaned with a piranha solution at a 1:2 ratio of 30% hydrogen peroxide to 95% sulfuric acid for an hour. The cleaned coverslips were stored in milliQ water until constructed into flow cells, which was within 24 hours of being cleaned. In the Terahood the coverslips were then rinsed with water, dried, and immediately constructed into a flow cell. Double-sided adhesive (30 or 50μm x7mm, Nitto Denko) in rectangular strips were applied to fasten the two coverslips together and set the flow cell channel width to 8mm. For the flow cells with glass coverslips

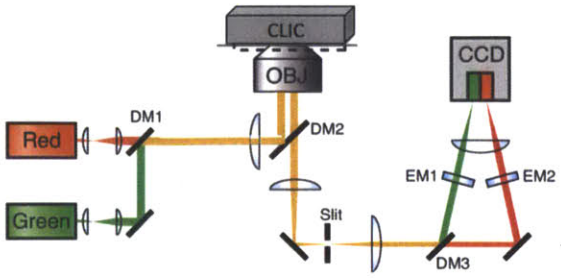

Figure 2-2: The general optical setup to obtain a dual-imaging setup is diagramed, where OBJ = objective, DM = dichroic mirror, EM = emission filter and CCD = charge-coupled device [20].

the top coverslip was a No. 0 (22x22mm), the bottom was a No. 1.5 (22x40mm), and tape 50μm thick were used. When quartz coverslips were used, both were 22x40mm and the tape was 30μm thick. Since both coverslips were the same size, they were aligned to be slightly offset and not overlap at the top and bottom to enable the insertion of the sample solution. Quartz, while more costly has significantly lower levels of autofluorescence and was used for the more sensitive measurements.

2.2 Experimental Method

Buffer solution is carefully flowed into a freshly constructed flow-cell lying horizontally in the CLIC device to wet the flow cell. A Kim-wipe can be held at the non-loading flow cell end to create suction. The flow cell is aligned such that the inner edge of the tape is across the center of the objective. Enough solution is loaded to cover the inlet of the flow cell to account for evaporation. The objective is raised until it makes contact with the bottom coverslip. Using white light, the microscope is adjusted to focus on the bottom of the tape. The flow cell can then be shifted so its center is over the objective and it is then secured with clips.

Freshly prepared sample is loaded into the flow cell in its secure position on the CLIC device. It is important to use the Kim-wipe here to absorb the old buffer as the sample solution is flowed in. Bubbles should not be introduced while loading, which can be prevented by ensuring the inlet is kept full. Since the solution is replacing the

buffer it is important to load at least 100μL to fully remove the buffer.

Immersion oil is applied to the top of the coverslip to match the index of reflection between the coverslip and the lens, thereby preventing reflections. Initially, the height of the cantilever is such that the lens will just come into contact with the top coverslip of the flow cell. The lens is lowered by by turning the $1/4 - 80$ threaded rod until it cannot rotate further. With the lens at its current height, the focus is finely adjusted until the fluorescence of the sample is in focus. The nut can then be turned until it reaches the top of the spring. Image settings that produce a reasonable signal are selected at this point for alignment, normally an attenuation of OD1.

To form the nano-gap, the nut is rotated further, which causes the cantilever arm to apply a force to the top coverslip. The force required to compress the spring with the nut is approximately equivalent to the force required for the coverslip to deform under the lens. The spring can be compressed to varying extents in order to create gap profiles of varying depths. As the gap height decreases, the fluorescence counts will like-wise diminish. These next stages require steady adjustments in order not to induce cracks in the coverslips. As the nut is turned more, the focus should be continuously adjusted to maintain the focus on the flow cell's bottom surface.

Once the fluorescence has significantly decreased the translation stage should be dialed in both the X and Y directions to find the minimum fluorescence region that corresponds to the contact point between the top and bottom coverslip and the minimum gap center. The emission filter can be briefly removed and the interference rings examined to select the minimum gap with a greater precision. Once centered, the nut is lowered more until full contact of the coverslips occurs, which is generally visibly evident with stuck molecules blinking in a homogeneous sample. At this point a humidifier box should be placed over the sample and stage to minimize evaporation, and the focus should be adjusted if needed.

LabView code was developed to control the translation stage. The stage is zeroed at its current central location. LabView is then used to adjust the position to various radial distances and gap heights. Movies of the sample at each position are recorded. At each field of view spanning 512x512 pixels, separate data collections are taken

with each of the appropriate lasers. The exposure times of 2, 5, 10, 20, 40 and 80 are collected to capture the different dynamics on different timescales. The gain is adjusted depending on the molecular system to avoid saturation.

Once the location specific data is collected, the LabView code is run to conduct scans that the span an adjustable number of fields of view across the flow cell. It is set so that each translation of the motor N_{steps} corresponds to 80% of a field of view. The number of locations N_{loc} in the X and Y directions are entered, as well as the total time for a movie at each location. The stage is then translated in each direction to the corner of the rectangular scan that is to be performed, which is equal to $-\frac{1}{2}N_{steps}N_{loc}$ in each direction. The system is re-zeroed there. A scan is run of the fluorescence profiles across the flow cell and another is taken after the emission filter is removed is capture the interference fringes. Four scans should be captured in total, collecting one of each type for each of the two laser wavelengths.

2.3 Calibrations

The calibration measurements described below are collected to improve the accuracy of the analysis. Some are more general characterizations taken for each sample type, and others are daily calibrations.

Each experiment utilizes certain fluorescent dyes. The standard fluorescence of the dyes is measured by loading a flow cell with free samples of the dyes used and conducting the entire standard CLIC process on the dyes alone. The data can then be analyzed analogous to the molecular system data as a control experiment to demonstrate any effects are due to the molecules and not the dye system.

Previously prepared uniform polyvinyl alcohol (PVA) film is imaged with the inverted microscope under both lasers used in the sample data collect with OD0, 0.1s exposure, 100 gain, and for 10 frames. These images are used to normalize the illumination profile of each laser in each of the dual fields. As it is specific to the precise alignment of the lasers, this calibration is recorded and applied to the analysis for each data independent session.

The second daily calibration uses a sample of dense fluorescent beads. Images are taken with each of the relevant lasers, producing images such as that seen in Fig 2-3. The images are analyzed to correct the alignment of the dual fields of view based on the relative locations of the beads in each window. The corrects are applied to each data set to ensure correlations across the dual views are spatially accurate.

Figure 2-3: An image of the fluorescent beads for dual-imaging alignment. The precise location of the beads on each side are determined and the one side is translated and rotated to align with the other side to within one pixel accuracy.

2.4 Nano-Gap Profile Characterization

The confinement feature of CLIC not only improves the imaging qualities, but also provides a useful variable to study molecular systems. Biological molecules such as DNA and protein exist and operate in cells, which themselves provide confinement. Therefore, the CLIC systems contains the possibility for the study of confinement effects in intracellular reactions. The first step in achieving that is the accurate characterization of the CLIC's nano-gap profile.

2.4.1 Fluorescence Profile

The varying radial gap height of the flow cell when CLIC's lens is lowered until contact is achieved between the top and bottom coverslip causes the exclusion of molecules or limitations on the number of molecules at different locations in the gap profile. These restrictions on the fluorescent molecules' locations are reflected in the intensities of

the acquired images. By analyzing the varying fluorescence intensities the gap height and molecular density can be characterized.

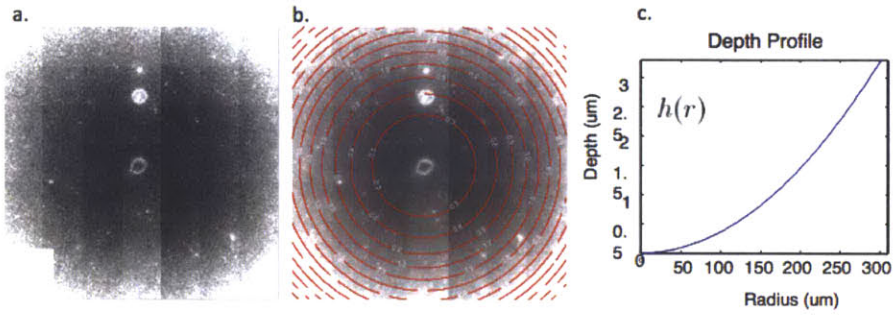

Figure 2-4: (a.) The fluorescent supermat. (b.) A gradient obtained as the result of a least squares regression to a quartic model for the intensity variation of the fluorescence across the supermat. (c.) The radial depth profile obtained from the regression fit.

As described in earlier, scans are taken across the flow cell-CLIC system. Code was developed to precisely insert the collected scan images into one image referred to as a supermatrix of the images. The supermatrices of the fluorescence data display the fluorescence profile across the compressed flow cell. A two-dimensional quartic least squares fit is applied to the intensity distribution of the supermatrix to fit the fluorescence gradient and extract the radial gap profile as seen in Fig. 2-4.

2.4.2 Interference Characterization

The emission filters in the dual-imaging system normally block the reflected excitation light so that only the light from the sample is recorded. However, when they are removed the reflected light is also recorded by the CCD camera. The continuous gap height profile means that at certain radii their will be constructive and destructive interference that will create rings of interference fringes. The radii of the constructive interference rings are analytically selected by fitting to extract the highest radial intensities. They are also theoretically located at $m\lambda/2n$ where m in the order of the ring, λ is the light's wavelength, and $n = 1.33$ is the refractive index of the coverslips.

By fitting supermat interference rings produced by each of the two laser wavelengths the minimum gap height offset can be analytically determined.

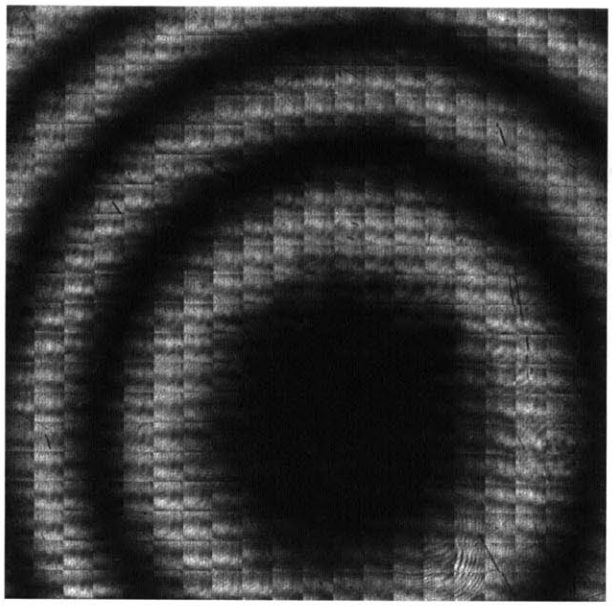

Figure 2-5: Interference Fringe Supermatrix

The results from the interference gap height profile can be applied to improve that of the fluorescence profile. The interference result produces very accurate results for the more narrow constructive fringes. The fluorescence result provides a continuous gradient of gap height, but requires the set values from the interference analysis to properly scale the profile and provide a reliable gap height profile for across the flow cell.

Chapter 3

Actin Filament Polymerization

3.1 Experimental Method

Data was collected for two different actin samples[1]: one with actin and the second with actin and native formin. The actin monomers were labeled with Alexa-488 dye and the formin was labeled with Alexa-647 via a SNAP-tag molecule (New England BioLabs). The actin molecules were diluted in solution to a micro-molar concentration with detergent and oxygen scavenger agents to respectively prevent sticking and improve the dye's photo-stability. Formin has included in the sample solutions at a nano-molar concentration.

Once the dilute solution of molecules was inserted into the flow cell and the CLIC device's lens lowered into its standard setup, data was recorded over a range of exposure times and frame rates with a magnification of 0.3μm/pixel.

The samples that contained only actin molecules were imaged with a 488nm (blue) laser at 0.3mW of power. Three different conditions were observed: freely diffusing actin polymers, the shearing and annealing of actin polymers, and the polymerization of actin from a solution of monomers.

The samples that contained both actin and formin were observed with a dual-imaging system with both a 488nm and a 647nm (red) laser. Data was recorded

[1]This set of experiments were carried out in collaboration with Zvnonimir Dogic of Brandeis's Physics Department and his graduate student Kyoko Okada. They provided the biological samples in order to conduct measurements with the CLIC system that were not previously possible.

for the following two conditions at different gap-heights: the growth of long actin filaments bound to formin at short times (high frame rate), and the growth of short actin filaments not bound to formin at long times (low frame rate). The gap-height variance allowed for the dependence of the growth on the gap size to be observed.

3.2 Results and Discussion

The freely diffusing actin molecules were recorded for the purpose of assessing the imaging parameters. Data was acquired and maintained with a good signal-to-background ratio with exposures of 0.2ms for observations times exceeding tens of seconds.

Freshly sheared filaments, following physical tapping agitation, were observed as they began as short, rapidly diffusing molecules (Fig 3-1.a.) and annealed to form longer filaments (Fig 3-1.d.) in the same region. By observing the growth over a long time, the effect of the gap-height is seen. As shown in Fig 3-1.d., once the filaments reach a certain size they are excluded from smaller gap-height regions. This can be used to either characterize the gap height or measure the size of molecules.

The polymerization of actin from a solution of fresh monomers was then recorded over a span of 1000 frames, equivalent to several minutes, which produced filaments such as in Fig 3-1.b. The CLIC's extraordinary detection time was displayed and confirmed as it was for the most part not limited by diffusion out of the field of view like in other single molecule imaging methods.

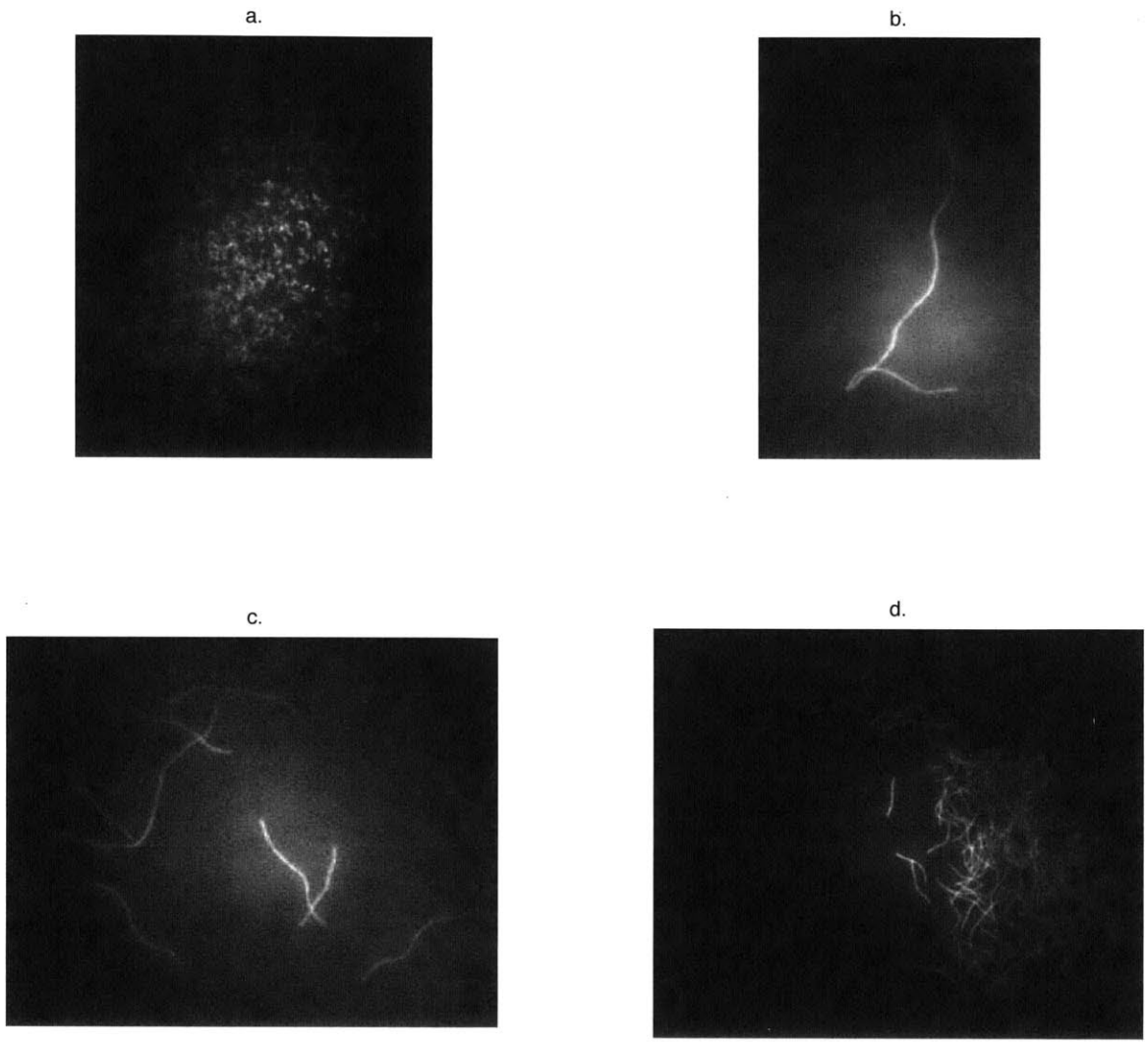

Figure 3-1: (a.) Short actin filaments recently sheared by tapping the flow cell -CLIC device as shown here at a high concentration. (b.) One long actin filament whose polymerization was observed continuously for in total greater than ten minutes. (c.) Multiple mid-length actin filaments examined over a span of several minutes. (d.) Taken at the same location as a., but at a later time. The exclusion of large enough filaments from smaller gap-heights is clearly shown.

The polymerization results were compared for those obtained with formin included in the samples. The formation of actin filament where formin was present appeared to be highly dependent on the concentration of the actin monomers, which was mea-

sured by the levels of background fluorescence. However, the concentrations varied significantly between different flow cells. In the future, methods for samples and flow cells with consistent actin concentrations should be developed and executed. This will also allow for the reliable quantitative analysis of the actin-formin biological system.

3.3 Conclusion

The recorded data, while lacking quantitative value, shows qualitative that there is a significant difference in the length of the actin filaments polymerized after 100 frames or approximately eight minutes. Additional gap-height effects were also observed in the presence of formin. At greater gap-heights the filaments formed notably faster than in smaller gaps. This is potentially a result of the space required by the polypeptide structure of formin to aid the assembly of actin filaments.

The particle tracking analysis of the data is currently underway in the Dogic Lab. Additional data to collect are movies with mutant formin molecules to determine the significance and effects of different sections of the protein.

Chapter 4

HMGB

4.1 Experimental Methods

Imaging of samples containing DNA and HMGB protein[1] were conducted. The protein used was NHP6Ap, found in yeast and a member of the HMGB protein family. It was labeled with Alexa-647 dye. DNA $\phi X154$ strands labeled with YOYO-1 dye were used. Where the persistence length of the DNA is 50nm, a ratio of one protein molecule per DNA persistence length was used. This lead to final concentrations of 40nM for NHP6Ap protein and 1nM for the $\phi X154$ DNA.

Samples were prepared with only DNA, only NHP6Ap, and with both DNA and NHP6Ap. The samples were inserted into a glass flow cell, and placed in the CLIC device. The freely diffusing NHP6Ap proteins and DNA were observed via a dual-imaging system with 488nm and 647nm lasers over a range of exposure times and frame rates.

4.2 Results

The HMGB system was studied for two reasons. The first was to demonstrate the dramatically increased observation times of the CLIC system, and particularly the

[1]This research was performed in collaboration with Jim Maher's group at the Mayo Clinic. They provided the biological samples.

quality over the complete observation as seen in the particle tracking capabilities. Second, was to probe the weak protein-DNA interactions of NHP6Ap and $\Phi X154$ by observing a change in DNA size.

The particle trapping code that was developed was applied to track the freely diffusing trajectories of the labeled protein and DNA molecules both when bound and unbound. An example is displayed in Fig. 4-1. where the red imaging side shows the observed labeled protein and its corresponding trajectories over a two second time interval.

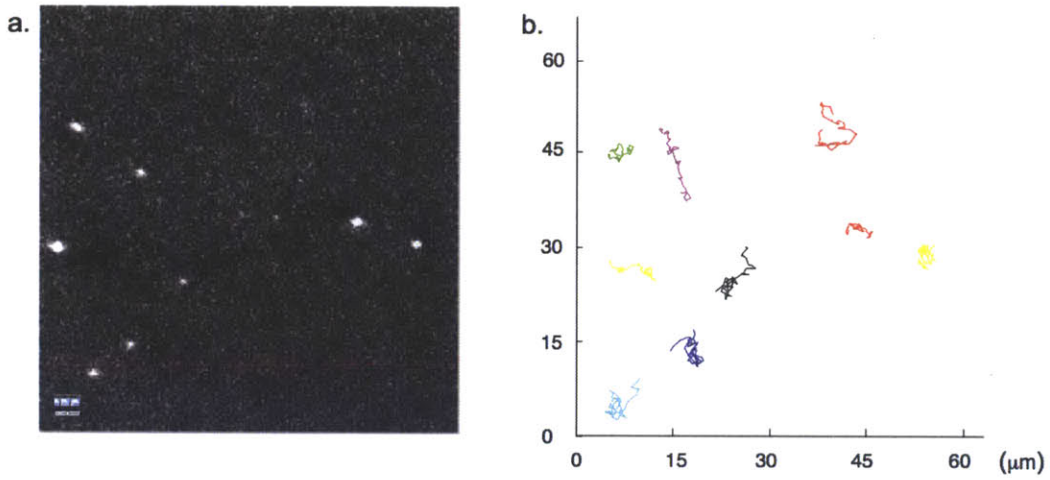

Figure 4-1: (a.) One frame of a red illuminated protein-DNA sample is shown with eight easily visible Alexa-647 labeled NHP6Ap protein molecules. The frame is taken from a movie collected with a frame rate of 40ms over a total of 2s. (b.) The trajectories obtained from the same movie are plotted over the full 2s. Each of the nine trajectories corresponds to a different NHP6Ap protein-$\Phi X154$ DNA molecular complex. The data from each of the trajectories was analyzed to calculate the corresponding diffusion coefficient.

The samples with only DNA produced trajectories in the blue-imaging side. These were analyzed to determine the unbound DNA's diffusion coefficient. The red-imaging side of sample with only protein were analyzed to extract the diffusion coefficient of the bare protein. Both imaging side were used to examine the sample containing

both protein and DNA. The blue DNA-visible side served to verify that the proteins were in fact bound to the DNA. However, the red protein-visible imaging data was analyzed to extract the diffusion coefficient for the bound protein-DNA molecular system. The ratio of average diffusion coefficients between the protein-DNA and the YOYO-DNA system was found to be 0.7 with data accumulated from five different samples.

A key concern when working with labeled molecules is how the dye affects the properties of the molecule. Previous experimental work has determined that YOYO-1 dye changes the spacing DNA base-pairs, which has the eventual effect of an increased DNA contour length and a decreased diffusion coefficient in comparison to bare DNA. For one dye molecule to every 10 base-pairs, the ratio of the contour lengths for YOYO-1 labeled DNA to bare DNA was measured to be 1.12 [21]. Therefore, the diffusion coefficient measurements obtained in this experiment for the YOYO-$\Phi X154$ complex will be lower than those for the bare $\Phi X154$ DNA. If the correction from [21] is applied, then the ratio in contour lengths of the DNA-protein complex to the bare DNA is 1.2. This can be converted back to a ratio of diffusion coefficients of 0.9. The ratio of contour lengths is in good agreement with the results obtained for the optical tweezer experiments at similar salt and protein concentrations of 1.16 [9].

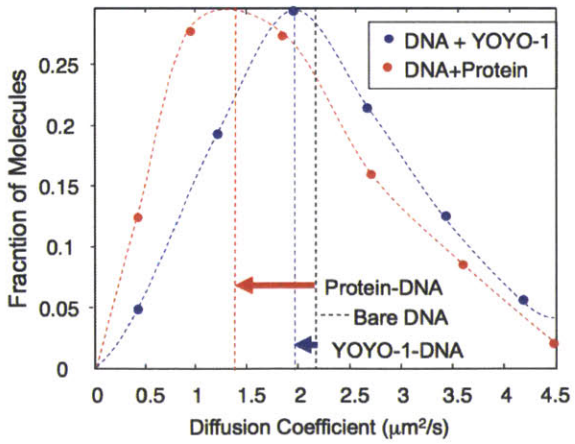

Figure 4-2: The diffusion coefficients for several hundred molecules, both DNA-protein and DNA-YOYO complexes, are plotted and the shift in diffusion between the two molecular systems is visually apparent. The corrected diffusion coefficient for bare DNA is also shown for reference.

4.3 Conclusion

The measurements of the NHP6Ap protein and $\Phi X154$ DNA molecular systems successfully demonstrated the ability of CLIC to capture data for particle tracking. They also verified the accuracy of CLIC's measurement and particle tracking by the very close agreement with optical tweezer results for the extract ratio of persistence lengths between the protein-DNA and bare DNA systems. However, the methods used by CLIC are more direct as they do not require the fit of a model to force extension curve such as in the case of optical tweezer measurements. Future research aims to measure per molecule binding and unbinding rate of this molecular system, currently unexplored by other techniques

Chapter 5

DNA Hairpin

Measurements of the DNA hairpin system and the corresponding analysis with fluorescence cross correlation spectroscopy (FCCS) make the most of CLIC's ability to probe and characterize previously inaccessible slow and weak interactions. This experiment is also the first step towards the long term goals of the Leslie Labs at McGill University to study the physical mechanism for the initial step of homologous chromosome alignment.

5.1 Experimental Methods

The selected hairpin[1] for the experiments had a stem consisting of A and T complimentary base pairs and a complementary strand of 38 bases. The strand length was selected for its high FRET level as seen in Fig. 5-1 [1]. The FRET system for the hairpin includes a fluorophore that emits green light at 532nm and a quencher fluorophore that emits at 647nm. When the hairpin is in the closed conformation, the red quencher fluorophore will continue to emit unlike the green whose energy is transfer via resonance to the quencher.

The dual-imaging is used for green and red wavelength lasers in this experiment. Due to the special FRET system the red viewing field will continuously capture the

[1]The hairpin was constructed by Alex Fields of the Cohen Lab for his research on DNA mechanics. He collaborated with Dr. Sabrina Leslie and myself on this research in order to provide addition insights into the physics of this unpredictably complicated unimolecular system.

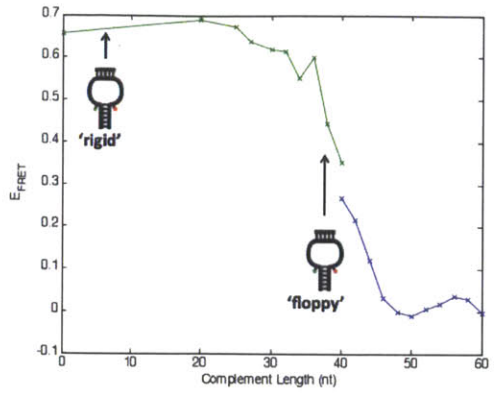

Figure 5-1: [1]

location of the hairpins in all conformational states. The green field of view will record the location and times of the hairpins in the open conformation.

Samples are constructed and observed both with and without the 38 base-pair complementary strand annealed to the hairpin's loop. The solution of the hairpin sample consists of 1x HEPES 10mM buffer with 50mM NaCl (88μL), the oxygen scavengers PCA (10μL) and PCD (2μL), and the hairpin (10μL) for a final volume of 100μL. The sample is imaged following the standard CLIC experimental methods are described in Section 3.2.

5.2 Results and Discussion

After the threshold masks, illumination normalization, and alignment calibrations were applied the correlation analysis was conducted. For this system the two autocorrelation functions correspond to the two signals from the separate fields of view. The emitted red signal measures the presence of the hairpin in either conformation with the corresponding autocorrelation function G_{RR}. The emitted green signal measures the presence of the hairpin in the open confirmation when no quenching occurs via the autocorrelation function G_{GG}. Finally, the cross correlation between the collected red and green signals G_{RG} is applied to measure the presence of the hairpin in the closed conformation via the determination of which red signal corresponds to the

open conformation.

One flow cell experiment's analysis results are highlighted in the following figures. Each graph compares the results for both the same hairpin both with and without the 38 base-pair loop complement. The theory predicts that at short time the diffusion term in the correlation functions will dominate, and at long time it is the reaction rates. In comparing the hairpin with and without the complement strand, the purpose is to visually characterize the time scales of the dynamics. The different exposure rates were collected to determine which rate would most clearly capture the hairpin dynamics. With both hairpin types overlaid the difference in their dynamics account is clear for exposure times not greater than the conformation shifts and diffusion themselves.

Figure 5-2 displays the temporal-spatial autocorrelation curve for the red field of view for each of the six collected exposure times. Likewise, Fig. 5-3 displays the same curves but for the green field of view and Fig. 5-4 for the cross correlation of the two signals. The correlation values as plotted verses the spatial separation δr for select temporal separations τ as indicated in each plot's legend. An exposure of 5ms best captures the dynamics as longer exposures show less and less of a difference between the hairpins with and without the complement.

The point at which the complement hairpin curves cross that of the non-complement hairpin curves are compared for the different correlation curves. Examining the 5ms exposure, there is a clear difference. For G_{RR} the crossover occurs at a separation of 2 pixels for $\tau \leq 10$ sec. For G_{GG} the crossover occurs at a separation of 5 pixels for $\tau \leq 15$ sec. This is consistent with the results from other samples. It indicates that while the difference between the separate red and green signals is significant, the FRET signal of the hairpin is not strong enough to accurately measure the dynamics of the closed hairpin.

This is further seen in Fig. 5-5. Here the three correlation functions of the 5ms exposure data as plotted as a function of τ for select wave vectors $k = [1, 5]$ in the left-most graphs. Fits to the theory for the diffusion coefficients developed in Section 1.3.3 are attempted. The middle column of plots are the logarithms of the correlation

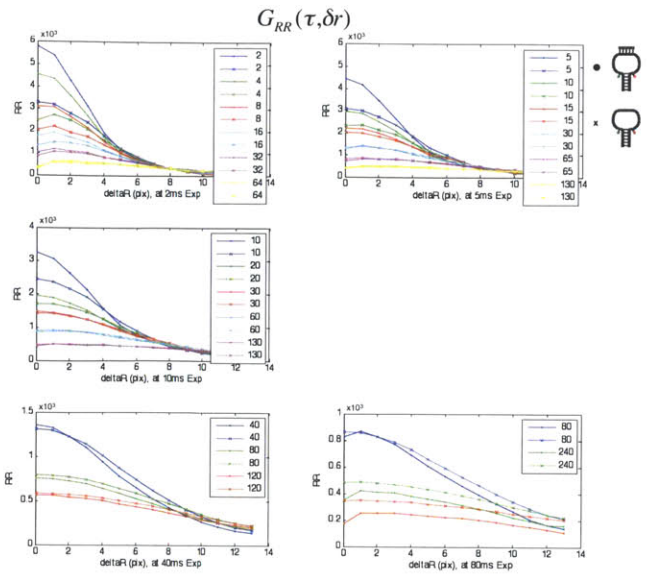

Figure 5-2: $G_{RR}(\delta r)$ over exposure range for select legend values of τ

functions with the least square regression fits to a linear model also displayed. These graphs emphasis that for the cross correlation function that measures the FRET signal, the noise levels in relation to the FRET signal remain to high for precision measurements. The right-most graphs are plots of the slopes extracted from the middle plots as a function of k^2. A least squares linear fit to applied to the final graphs in order to obtain estimates for the diffusion coefficients. While values were obtained within a very close range, they were not well selected for. Attempts to analyze the data and fit for the reaction rates were even more dominated by noise such that no reasonable values could be extracted.

5.3 Conclusion

The collected hairpin data, analysis, and fits serve to show the potential but not accomplished ability for CLIC in combination with FCCS to measure slow and weak DNA interactions. It was recently discovered that the hairpin systems will seemingly the simplest and logical first step actually undergo very strange dynamics, especially when there is the annealed complementary strand [1]. A binary system with two short

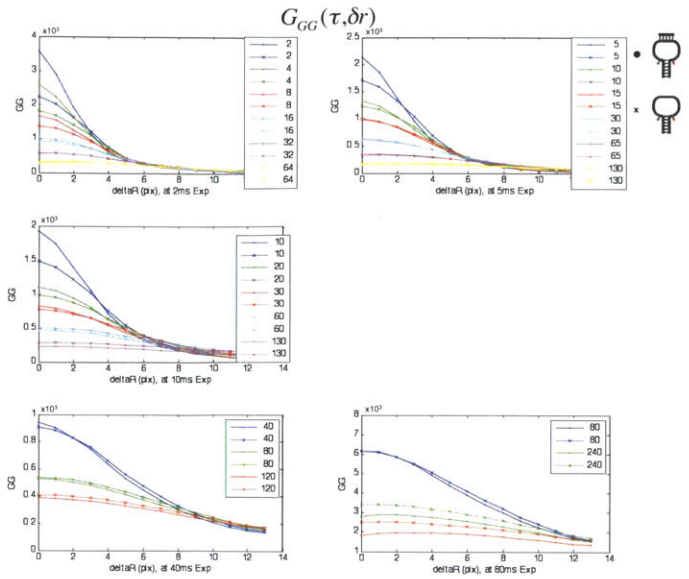

Figure 5-3: $G_{GG}(\delta r)$ over exposure range for select legend values of τ

DNA segments has already been designed and experiments are beginning to be executed in the Leslie Lab at McGill. The FRET signal will hopefully be strong enough and the dynamics stable enough to successfully extract the constants of diffusion and reaction rate that were attempted with the hairpin system.

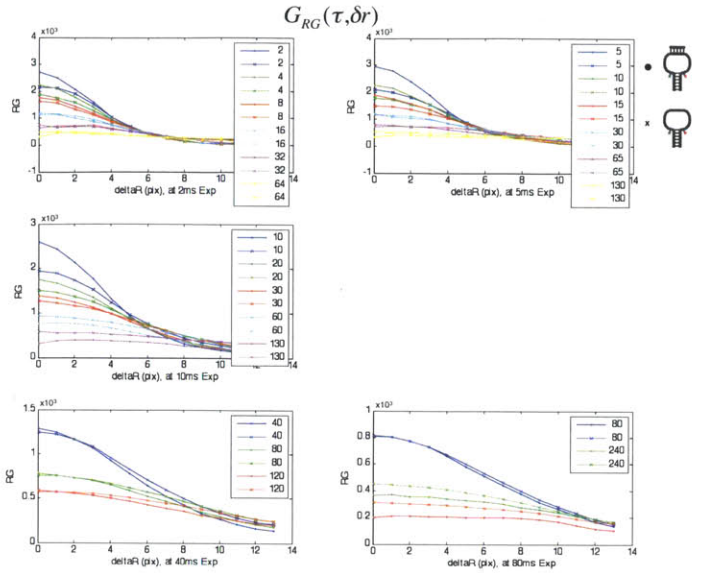

Figure 5-4: $G_{RG}(\delta r)$ over exposure range for select legend values of τ

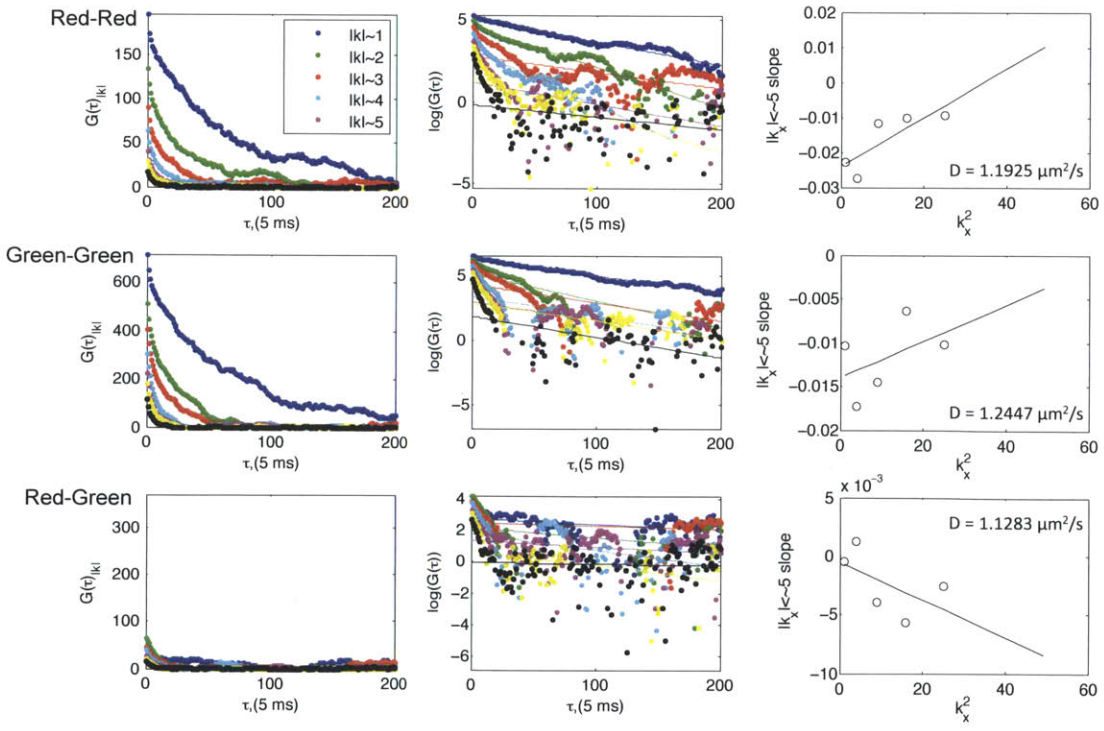

Figure 5-5: Fits to correlation curves for diffusion coefficient

Chapter 6

Conclusion

CLIC's expansion of the possible experimental parameter range is opening the door to new single molecule measurements. Specifically with the allowable higher molecule concentrations, observation volume, and observation time.

In this research, the increased observation time was successfully demonstrated for up to several minutes for protein-protein interactions such as the polymerization of actin filaments with formin. The larger observation volume played a key role in the valid and effective study of HMGB protein and DNA binding, causing the measurements to not be diffusion limited. Higher concentrations permit the collection of meaningful data of weak DNA-DNA interactions such as the DNA hairpin.

The observation volume and ease of the setup allowed for full scans to be conducted across the flow cell-CLIC system. The gap height profile induced by the lens were characterized. In the future, this ability will lend itself to the study of confinement effects on single molecule interactions. The relation between molecule dynamics and confinement size can be determined, and intracellular mechanisms such as the alignment of homologous chromosomes better modeled and understood.

Bibliography

[1] Alex Fields. Dna mechanics. Consultation and Cohen Lab Group Meeting Presentations.

[2] A.A. Kornyshev and S. Leikin. Sequence recognition in the pairing of dna duplexes. *Physical Review Letters*, 86(16):3666, April 2001.

[3] Daniel Axelrod. Cell-substrate contacts illuminated by total internal reflection fluorescence. *The Journal of Cell Biology*, 89(1):141, April 1981.

[4] Petra Schwille and Elke Haustein. Fluorescence correlation spectroscopy: An introduction to its concepts and applications. Technical report, Experimental Biophysics Group, Max-Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany.

[5] Sabrina R. Leslie, Alexander P. Fields, and Adam E. Cohen. Convex lens-induced confinement for imaging single molecules. *Analytical Chemistry*, 62:6224, 2010.

[6] Mary W. Elting, Sabrina R. Leslie, et al. Single-molecule fluorescence imaging of processive myosin with enhanced background suppression using linear zero-mode waveguides (zmws) and convex lens induced confinement (clic). *Optics Express*, 21(1):1189, January 2013.

[7] Frank J. Brooks and Anders E. Carlsson. Nonequilibrium actin polymerization treated by a truncated rate-equation method. *Physical Review E*, 79(3), 2009.

[8] David R. Kovar. Molecular details of formin-mediated actin assembly. *Current Opinion in Cell Biology*, 18(1):11, February 2006.

[9] Micah McCauley, Philip R. Hardwidge, III L. James Maher, and Mark C. Williams. Dual binding modes for an hmg domain from human hmgb2 on dna. *Biophysical Journal*, 89(1):353, 2005.

[10] Micah J. McCauley, Jeff Zimmerman, III L. James Maher, and Mark C.Williams. Hmgb binding to dna: Single and double box motifs. *Journal of Molecular Biology*, 374(4):993, December 2007.

[11] Rae M. Robertson, Stephan Laib, and Douglas E. Smith. Diffusion of isolated dna molecules: Dependence on length and topology. *PNAS*, 103(19):7310, March 2006.

[12] Adi Barzel and Martin Kupiec. Finding a match: how do homologous sequences get together for recombination? *Nature Review*, 9:27, January 2008.

[13] Grégoire Bonnet, Oleg Krichevsky, and Albert Libchaber. Kinetcs of conformational fluctuations in dna hairpin-loops. *Proceedings of the National Academy of Sciences*, 95:8602, 1998.

[14] Daniel Blair and Eric Dufresne. The matlab particle tracking code repository. http://physics.georgetown.edu/matlab/.

[15] John C. Crocker and Eric R. Weeks. Particle tracking using IDL. http://www.physics.emory.edu/ weeks/idl/.

[16] Elliot L. Elson and Douglas Magde. Fluorescence correlation spectroscopy. i. conceptual basis and theory. *Biopolymers*, 13:1, 1974.

[17] Sabrina R. Leslie. Fluorescence correlation spectroscopy. Phys / Biol 319, Lecture 6, McGill University, January 2012.

[18] Andrei Takmakoff. Förester resonance energy transfer, time-dependent quantum mechanics and spectroscopy. MIT Department of Chemistry, March 2008.

[19] Sabrina R. Leslie. Concentration correlation spectroscopy: Clic implementation with wide-field imaging. April 2011.

[20] Min Ju Shon and Adam E. Cohen. Mass action at the single-molecule level. *JACS*, 134:14618, August 2012.

[21] Katrin Gunther, Michael Mertig, and Ralf Seidel. Mechanical and structural properties of yoyo-1 complexed dna. *Nucleic Acids Research*, 38(19):6526, October 2010.