SINGLE MOLECULE FLUORESCENCE IMAGING OF BIOSENSORS, RIBOZYMES AND MOLECULAR SPIDERS

by

Chamaree de Silva

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in The University of Michigan
2009

Doctoral Committee:

Associate Professor Nils G. Walter, Chair
Professor Roseanne J. Sension
Associate Professor Hashim M. Al-Hashimi
Assistant Professor Jennifer P. Ogilvie

Do not believe in anything simply because you have heard it.
Do not believe in anything simply because it is spoken and rumored by many.
Do not believe in anything simply because it is found written in your religious books.
Do not believe in anything merely on the authority of your teachers and elders.
Do not believe in traditions because they have been handed down for many generations.
But after observation and analysis,
When you find that anything agrees with reason,
And is conducive to the good and benefit of one and all,
Accept it and live up to it.

- Siddhartha Gautama Buddha
563-483 B.C.

To *Amma* and *Thaththa*
For life and unconditional love

ACKNOWLEDGEMENTS

I thank my mentor Dr. Nils Walter for introducing me to the world of RNA and the supportive and collegial environment of Walter Lab. I appreciate the scientific advice he has given me over the past five years. I am also thankful to the other members of my committee Dr. Hashim Al-Hashimi, Dr. Jennifer Ogilvie and Dr. Roseanne Sension for their support and advice.

I thank Dr. Meredith Newby who was my mentor during my rotation in the Walter Lab. She was a great teacher who helped me fall in love with my work. In addition, she taught me professionalism by example, and was a valuable colleague and friend. I missed her dearly during my last years in the Walter Lab. I am also thankful to Dr. David Rueda who helped me a great deal to get acquainted with the single molecule microscope setup as well as introducing me to the Aptazyme Project.

Working in the Walter Lab gave me the opportunity to meet several other wonderful people. Dr. Jana Sefcikova taught me how to have a positive outlook when times are tough. She was also a great friend and a companion. Dr. Liz McDowell was a friend who looked out for me and walked miles in a foot of snow on Saturdays to help me with biochemistry and cell biology classes. She was my unofficial mentor who always had valuable contributions towards the Aptazyme Project while I was dealing with unruly

single molecules. She was also the person with whom I shared my enthusiasm towards animals and nature, and perspectives on cuisine, gardening and community involvement. I also thank Dr. Anthony Manzo for sharing my experience with the Spider Project. His work tremendously helped the progress of the project. He also has been a valuable colleague whose commitment, inquisitiveness and attention to details helped me look at science from a different perspective. I am also grateful to have met Dr. Jennifer Furchak with whom I shared an office and early morning discussions of science and everyday matters. In addition, I am thankful to Ms. Xiaomu Guan, Ms. Robin Johnson, Dr. Shiamalee Perumal, Dr. Hannah Townsend and Mr. Alex Johnson-Buck for contributing towards my work. I also very much appreciate the input and company of all the other Walter Lab members past and present, especially Mr. Miguel Pereira, Dr. Mark Ditzler and Ms. Afi Rawlings.

I am also fortunate to have met several other wonderful colleagues who are not a part of Walter Lab. Special thanks to Mr. Hao Ding whose advice on which classes to take as well as choosing a mentor greatly helped me during my first year. Dr. John Heish helped me immensely with training on the stopped-flow instrument and with stimulating discussions on the Aptazyme Project. Also thanks to Ms. Qihui Ni, Ms. Amy Payeur and Mr. Colin Jennings for keeping me good company in the shared office space. I must also thank collaborators Dr. Milan Stefanovic and Dr. Steven Taylor from Columbia University for introducing me to the Spider World, sending Spider samples and for helpful suggestions.

I was truly fortunate to have met great friends in Ann Arbor. Ms. Heather Wolf has been an amazing friend and an extraordinary roommate. She introduced me to a

whole new world and subjects outside science. It was refreshing to have a friend with a different professional background. She has always listened when I needed to talk, given me suggestions and courage on how to face certain challenges and has been truly supportive. Mr. Akash Bhattacharya has been a companion with whom I regularly had discussions about cuisines, culture, politics, arts, science and photography. In addition, Mr. Karthik Sathiyamoorthy, Mr. Sethu Pitchaiya, Mr. Krishnan Raghunathan and Ms. Vishalakshi Krishnan have kept me company, given me good food and have shared everyday challenges that graduate school brings. I also enjoyed the company of numerous other wonderful friends I met in Ann Arbor, especially Ms. Courtney Wilkinson and Ms. Alison Beehr.

I am happy to have met Mr. Eric Huff at the University of Arizona who has been a constant supporter and a great friend, even when we now live thousands of miles apart. He has been a great listener and given me encouragement in my professional and personal life. I am also grateful for all the other friends I met in Tucson who enriched my life and helped me get accustomed to life in the US. Furthermore, special thanks to Mr. Rachitha Dayaratne who lightened up my mood whenever we talked. I am fortunate to have known him since we were born and to know that I have a lifelong friend.

I must acknowledge professors Dr. William Bickel, Dr. John Beiging and Dr. Adriana Pesci from the University of Arizona who always supported me throughout my undergraduate career. I also have the utmost respect for numerous teachers I had in Sri Lanka. They always challenged my capabilities and were encouraging me to aim higher. My gratitude also goes to all my school friends with whom I shared 13 years of everyday life.

I thank Mr. Dushan Aththidiyavidana for always being there for me. He has always been supportive of my professional career and has been a great partner and a friend. I am fortunate to have found him to walk beside me in life. I am looking forward to our future together.

I am truly fortunate to have the support and love of my uncle Mr. Sumithra de Silva. It was through his immense generosity that I had the opportunity to come to the United States for higher studies and to travel the world. He is always there for me to guide and support the choices that I make, and to encourage me in any possible way. I feel privileged that he considers me his daughter. I also thank all my uncles, aunts, cousins, late grandparents and family friends who have loved me throughout the years.

It is my utmost fortune that I have the encouragement of my parents Mrs. Amita de Silva and Mr. Daya de Silva to pursue higher studies and to do the best in anything I try. My father introduced mathematics and science to me at a very early age and always encouraged me to explore higher grounds. My mother sat beside me while I was studying and helped me with subject matter all throughout my school years. She also taught me kindness and patience by example. Moreover, both my parents always encouraged me to help the less fortunate. They poured all they have and supported me from day one to be the scientist and the person that I am today. I have nothing but immense love and gratitude for them. I am honored to be their daughter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

SINGLE MOLECULE FLUORESCENCE IMAGING OF BIOSENSORS, RIBOZYMES AND MOLECULAR SPIDERS

by

Chamaree de Silva

Chair: Nils G. Walter

Single molecule fluorescence imaging has been developed in recent times to expand our understanding of the heterogeneity and biological mechanisms of molecular ensembles. In this dissertation, such imaging techniques, along with ensemble fluorescence spectroscopy tools have been used to investigate three systems of nucleic acid enzymes.

An engineered biosensor built from a theophylline aptamer and the hammerhead ribozyme (termed an aptazyme) was scrutinized using single molecule fluorescence resonance energy transfer (smFRET) and ensemble fluorescence studies. It was found that a catalytically active state is accessed both in the theophylline-bound and, if less frequently, in the ligand-free state. The resultant residual activity (leakage) in the absence of theophylline contributes to the limited dynamic range (<100 fold) observed for the

aptazyme. In addition, slow conformational rearrangements dampen the speed in which the catalytically active conformation is accessed. In contrast, the only known naturally occurring aptazyme uses a chemical cofactor to instantaneously trigger catalysis (with a 100,000 fold activation range), rather than the slower rearrangement of an inactive into an active structure.

To examine the effects of the U-turn of the hepatitis delta virus (HDV) ribozyme, recently found to be at the heart of the ribozyme's catalytic core, both DNA and RNA ligase mediated methods were evaluated to assemble the ribozyme from chemical synthesized fragments. Upon successful assembly of the ribozyme, preliminary smFRET studies were performed revealing global dynamics and heterogeneity promising to unveil new insight into the functional role of the U-turn.

Recently designed nano-robots called Molecular Spiders use nucleic acid enzymes as "fuel" to traverse on a specific two-dimensional landscape. In this thesis, individual Spider movement was surveyed by single fluorescent particle tracking. Two-dimensional Spider movement was followed in real-time, providing evidence for the previously hypothesized model that Spiders move in a self-repellent autonomous (cybernetic) walk. These nano-walkers represent potential drug delivery vehicles with the ability to understand and follow external cues.

Overall, the work presented in this dissertation has illuminated the suitability of single particle fluorescence techniques to monitor the functional behavior and heterogeneity of single nucleic-acid based molecules ranging from biosensors and small catalytic ribozymes to novel molecular nano-robots.

CHAPTER ONE

ADVANCES IN BIOPHYSICAL STUDIES OF NUCLEIC ACIDS

1.1 Introduction

During the last twenty years, Ribonucleic Acids (RNAs) have been recognized as an essential component of the maintenance, transfer and processing of genetic information. For decades, proteins were thought to catalyze all chemical reactions. However, this idea was shattered by the discovery of catalytic RNA, supporting the notion of RNAs in a pre-biotic world (known as the RNA World [1]), where RNAs served as the original macromolecules that encoded both their own genetic information and the catalysts needed for reproduction without the need for any protein cofactors. It was initially difficult to believe that the four nucleotides that constitute RNA catalyze similar reactions as the 20 or more diverse amino acids that make up proteins. However, when additional RNA functions and characteristics were discovered, their ability to serve numerous functional roles became clear.

The central dogma of molecular biology now says that DNA, which is the carrier of genetic information, serves as a template for DNA replication as well as RNA transcription, which in turn directs the formation of protein through translation of a

transcribed mRNA (Figure 1.1) [2]. The coding sequence of the mRNA determines the amino acid sequence of the protein produced. However, many RNAs are not used for protein coding (non-coding RNAs). These include transfer RNA, ribosomal RNA, microRNA and siRNAs. These RNAs are quickly becoming recognized as key elements of the cellular machinery responsible for modification, regulation, and maintenance of genetic information. Riboswitches residing in the 3′ untranslated region of mRNAs are another example of non-coding RNAs. Riboswitches can specifically bind a small target molecule which in turn can effect downstream gene expression by transcription termination or translation inhibition [3]. From these and similar examples, it is now apparent that RNA serves a crucial role in biology.

1.2 Catalytic RNAs and their applications

An important discovery proving the biological importance of RNA was the discovery of the catalytic properties of RNA. Thomas Cech and Sidney Altman were awarded the Nobel Prize in Chemistry in 1989 for this finding. RNA catalysis was independently observed in group I introns from pre-ribosomal RNA and in RNase P [1]. These large catalytic RNAs, quickly termed ribozymes, led to a rapid growth of RNA studies. The recent discovery of small self-cleaving ribozymes further prompted the RNA field to explore RNA function, catalytic activity, as well as structure and dynamics. The small self cleaving ribozymes include the hairpin, hammerhead, hepatitis delta virus (HDV) and the *Neurospora* Varkud Satellite (VS) ribozymes [4]. These ribozymes are

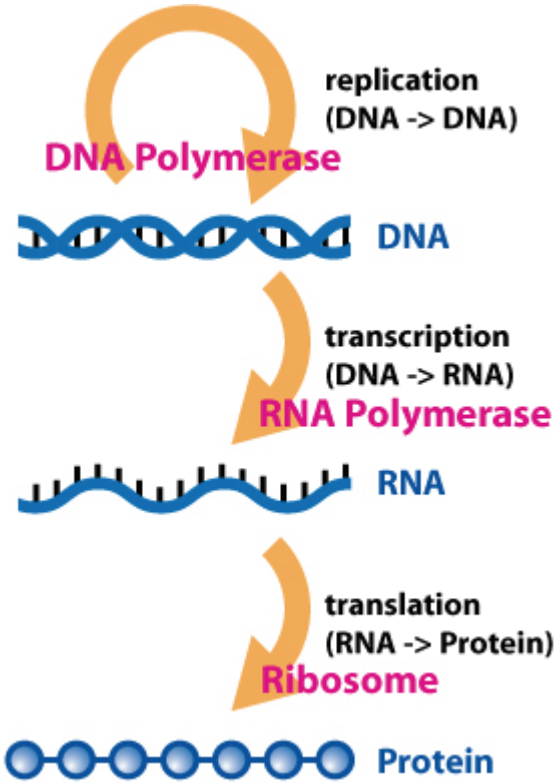


Figure 1.1 Central dogma of molecular biology. The genetic material carrier DNA can direct its own replication and transcription to create RNA, which in turn will direct formation of proteins through translation. (http://en.wikipedia.org/wiki/File:Central_Dogma_of_Molecular_Biochemistry_with_Enzymes.jpg).

less than 200 nucleotides in length, and are required for their respective genome replication.

RNA genomes of viroids, virusoids and satellite RNAs are replicated through a double-rolling circle replication mechanism. In this process, a polymerase will continuously move around a circular genomic template synthesizing a complementary antigenomic strand. When one circle is complete, the 5′ end of the copy will be displaced. The copy presents an oligomeric linear antigenomic strand that is site-specifically self-cleaving through the action of the embedded ribozyme motif. The monomeric antigenomic strands self-ligate into circular monomers that serve as templates for the polymerase to also synthesize genomic strands in a similar fashion [5].

Hammerhead ribozyme

The hammerhead ribozyme is the smallest known natural ribozyme. Its secondary structure is comprised of a conserved core flanked by three stems (stems I, II and III) with a "Y shape" and resembles a carpenter's hammer, giving rise to its name. The hammerhead ribozyme has been found in many organisms including the avocado sunblotch viroid, the peach latent mosaic viroid, the cave cricket, and the genome of rodents [6-10]. The hammerhead ribozyme carries out a phosphodiester transfer reaction that site-specifically cleaves the RNA at the C17 nucleotide. The catalytic core is comprised of 13 conserved nucleotides, most of which are not Watson-Crick base paired. The presence of divalent cations, such as Mg^{2+}, is required for self-cleavage of the hammerhead ribozyme, although high concentrations of monovalent cations can also

induce catalysis by promoting the correct folding of the ribozyme [11]. The minimal and the extended hammerhead ribozymes have been extensively studied for their structural, dynamic and kinetic properties [12-15]. In a recent study, the crystal structure of the full length hammerhead showed that nucleotide C17 is positioned for in-line attack, while residues C3, G5, G8 and G12 are all involved in interactions vital for catalysis (Figure 1.2) [16].

Therapeutic applications of the Hammerhead ribozyme

In recent years, hammerhead ribozyme motifs have been applied as therapeutic agents for rendering T-cell lines HIV resistant and against amyotrophic lateral sclerosis [17, 18]. Moreover, an *in-vitro* selected aptamer domain specific to the bronchodilator drug theophylline has been incorporated into stem II of the ribozyme to monitor theophylline concentration dependent cleavage activity and quantify it by radioactive and fluorescence assays [19, 20]. This aptamer integrating ribozyme, also called an aptazyme, shows a high level of specificity for theophylline, which is essential to distinguish between theophylline and its close chemical analog, caffeine. The hammerhead based aptazyme shows discrimination against caffeine, and in Chapter Two its global and local conformational changes has been probed using fluorescence techniques. In addition, its shortcomings in comparison to a naturally occurring aptazyme are presented.

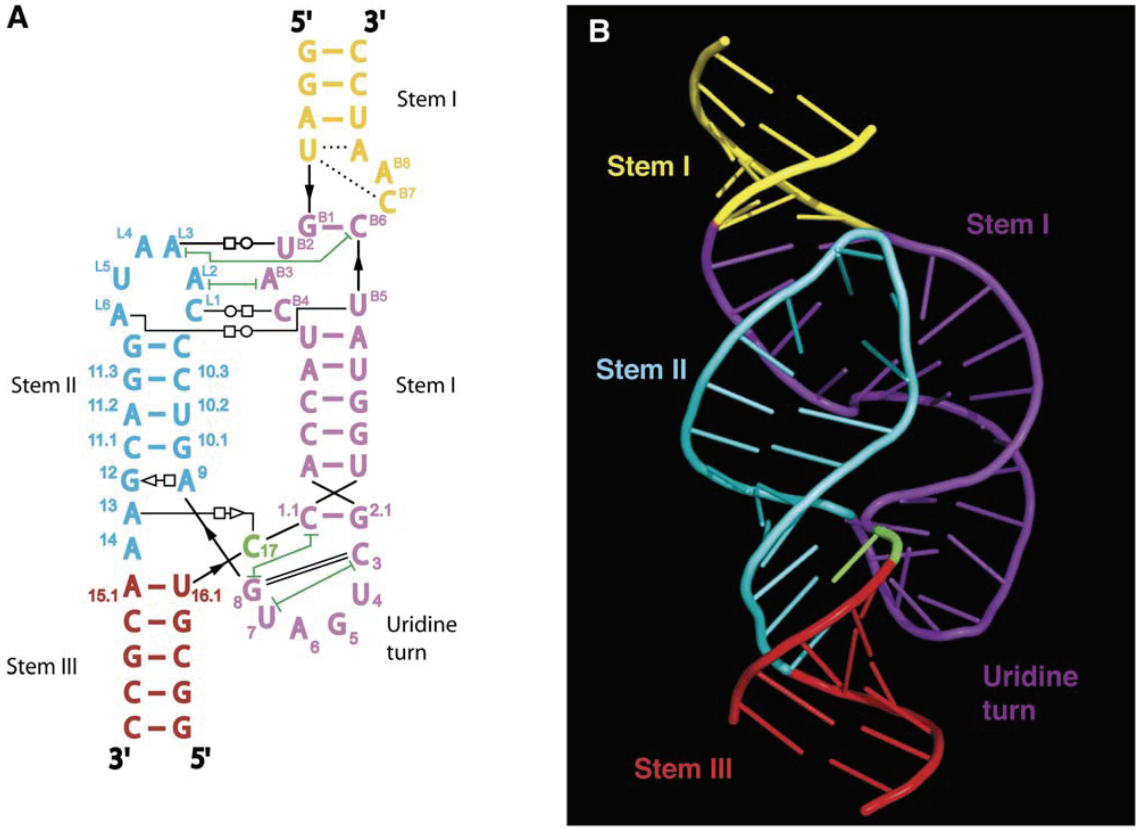


Figure 1.2 The full length hammerhead ribozyme from *S. mansoni*. Color coded (A) secondary and (B) tertiary structures. Reproduced from Martick, M. and Scott, W.G. Cell, 2006, 126(2): p 309-320, with permission from Elsevier.

Hepatitis Delta Virus ribozyme

The hepatitis delta virus (HDV) ribozyme is a single stranded RNA satellite of the hepatitis B virus (HBV). HBV, in the presence and absence of HDV, can lead to liver failure and liver cancer in humans [21, 22], however co-infection with HDV leads to more severe hepatitis. HDV was first discovered in late 1970s in the Mediterranean region, and is still prevalent in this area along with sub-Saharan Africa, South America and parts of the Middle East [22, 23]. It is also the only known human pathogen to encode a ribozyme in its genome. In the viral particle, the RNA is associated with multiple copies of the HDV encoded delta antigen, and this complex is enveloped by lipoprotein containing HBV surface antigens [22].

The HDV RNA genome is 1700 nucleotides in length and has ~70% self complementarity. The latter feature gives it an unbranched rod-like double stranded structure [24]. Both genomic and antigenomic HDV ribozymes are about 85-nucleotide long RNA motifs that have about 75% sequence homology. Moreover, they share a common secondary and tertiary structure. The HDV ribozyme folds into a structure comprising five helical segments (P1, P1.1, P2, P3 and P4) connected as a nested double pseudoknot. These helices form two parallel stacks, and are joined side-by-side by five strand crossovers, located at the joining regions J1/2 and J4/2 [25, 26] (Figure 1.3).

Both the genomic and antigenomic forms of the HDV ribozyme catalyze the self cleavage of a specific phosphodiester bond by a transesterification reaction. The cleavage site of the genomic HDV ribozyme lies next to the G1:U37 wobble pair at the 5′ end of

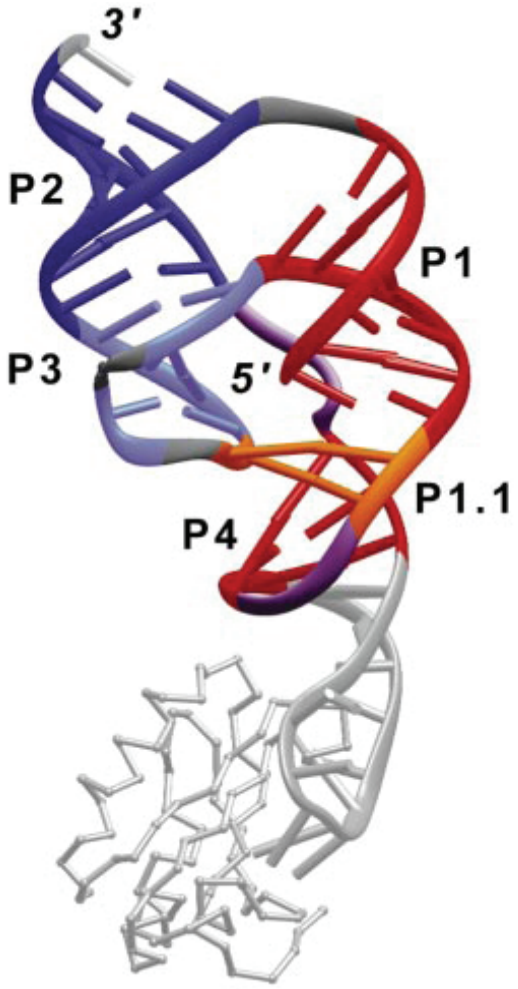


Figure 1.3 Crystal structure of the HDV ribozyme. Adapted by permission from Macmillan Publishers Ltd: Nature [14], copyright 1998.

the sequence. A single dangling nucleotide 5′ of the cleavage site (N-1) is sufficient for catalysis. At the cleavage site, the 2′-OH of N-1 attacks the scissile phosphate to release the 5'-OH. Cleavage occurs with the 2′-OH losing a proton and the 5′ oxygen accepting one during the reaction. It has also been shown that the nucleotide C75 (in the genomic form) is essential for catalysis. X-ray crystal structure studies propose that the U-1 undergoes a rotation bringing its 2′-OH within 2.7 Å, which is in hydrogen bonding distance of N3 of C75. This model suggests C75 acts as a general base to activate the 2′-OH [26, 27]. However, using a hyperactivated RNA substrate, it has also been suggested that C75 may function as a general acid by mediating proton transfer to the leaving group through a protonated N3-imino nitrogen [27, 28]. Additionally, both of these models imply the use of a hydrated magnesium ion for a complementary role in catalysis (Figure 1.4).

The U-turn of the HDV ribozyme

A previously unknown U-turn motif was recently found in the HDV ribozyme similar to that of the Hammerhead ribozyme and t-RNAs [29] (Figure 1.5). It was shown that the U-1 which is essential for fast cleavage of the ribozyme (activity reduction of more than 25 fold U-1 > C-1 > A-1 > G-1) leads to the most stable P1.1 stem. MD simulations have revealed that a kink around the scissile phosphate connecting the U-1 and G1 exposes the cleavage site to the C75 nucleotide in the form of a U-turn [29]. Nevertheless, how this U-turn is involved in the conformational changes and dynamics of the HDV ribozyme remains unknown. In Chapter Three, assembly of the HDV ribozyme

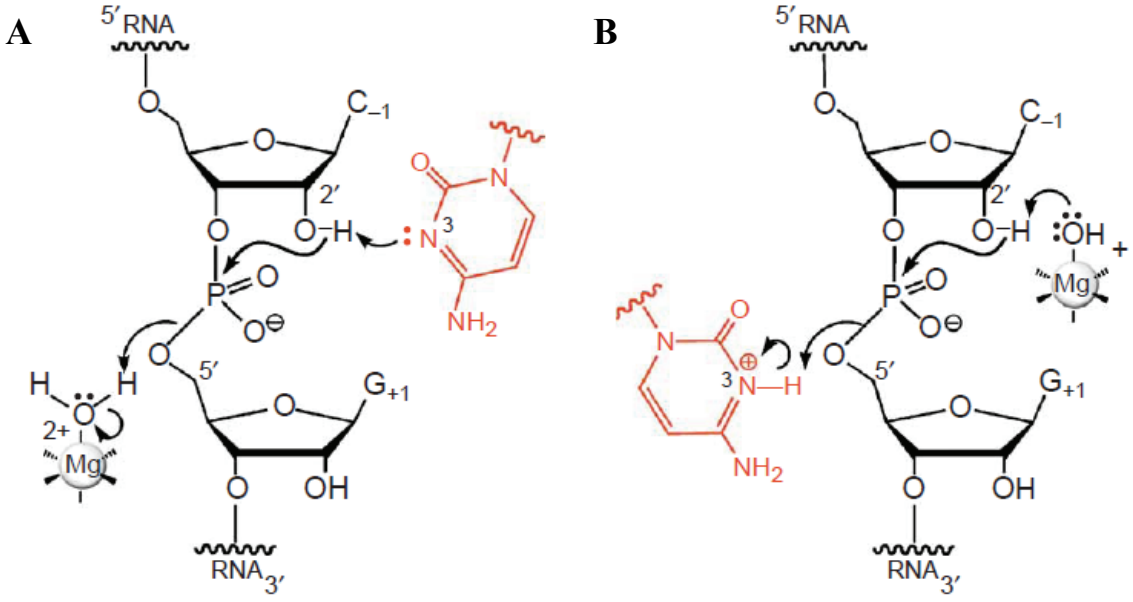


Figure 1.4 Cytosine 75 nucleotide in the HDV ribozyme. (A) General acid mechanism (B) General base mechanism. Adapted by permission from Macmillan Publishers Ltd: Nature Chemical Biology [30], copyright 2005.

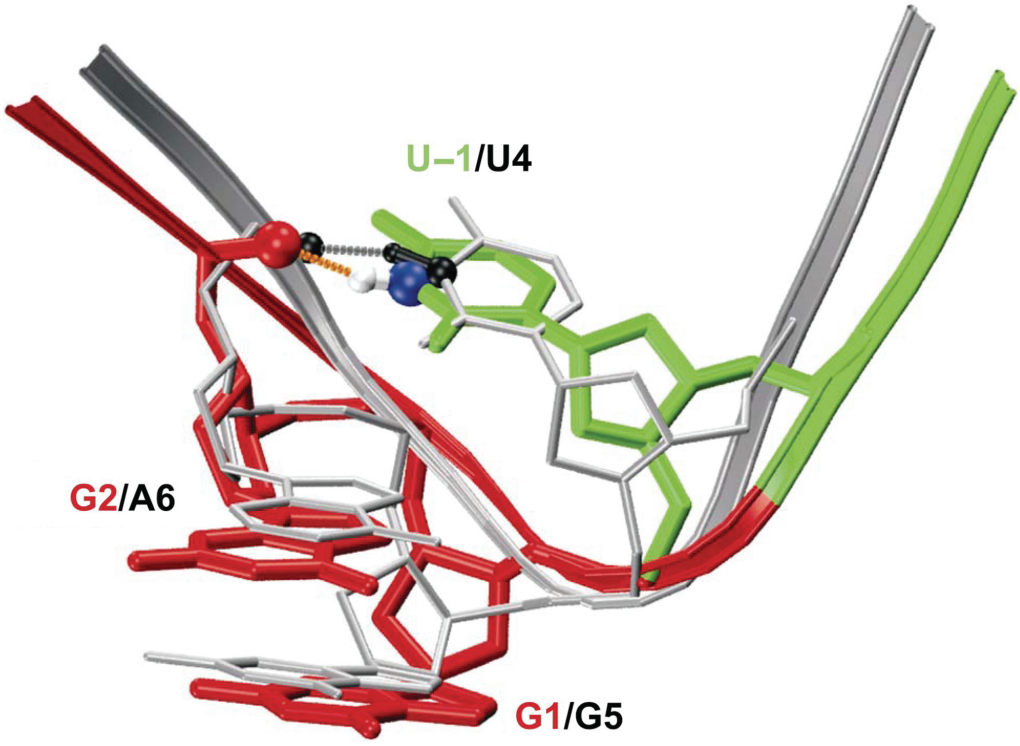


Figure 1.5 U-turn overlap of the HDV and the hammerhead ribozymes. Reproduced with permission from Sefcikova, J.,et al., Nucleic Acids Res, 2007, 35(6): p 1933-46. © 2007 The Author(s).

for single molecule fluorescence studies of its structural dynamics is experimentally explored.

1.3 Catalytic DNAs and their applications

About a decade ago, first examples of DNA enzymes emerged as scientists began understanding the importance of cleavage activity of ribozymes. These DNA enzymes also known as DNAzymes are *in vitro* selected [31-33]. In order to develop a DNA molecule than can be used as a tool to probe, manipulate or assist RNA, DNAzymes were selected to cleave almost any RNA substrate. One such example is the 8-17 DNAzyme which can recognize and cleave specific RNA targets [34, 35]. It is one of the simplest RNA cleaving DNAzymes that binds substrate and performs catalysis only in the presence of divalent metal ions. It has been shown that Zn^{2+} can activate the 8-17 DNAzyme more than other cations [35]. Since the substrate sequence can be changed without losing activity as long as the substrate binding arms of 8-17 are changed to complement the substrate, this enzyme can be used with almost any RNA sequence without further consequence. Furthermore, the 8-17 DNAzyme was selected to cleave a single RNA bond in an otherwise DNA substrate, giving flexibility to the selection of substrates that can be used.

Examples which use the 8-17 DNAzyme are the polycatalytic nano-assemblies called ‘Spiders’, which were developed by scientists to one day program these molecules to perform ordered tasks [36], and to control the behavior of a molecular robot in the same way as a ‘conventional’ robot. These Spiders are controlled by the landscapes they

traverse as well as their environment and their build. In the future, Spiders are expected to be trained to deliver drugs to a specific location of the cells, create biosensors to assist detecting hazardous materials, and perhaps even allow for DNA-based computing.

Spiders consist of a streptavidin "body" and multiple, identical 8-17 DNAzyme "legs" which are attached to the streptavidin giving it the shape of an actual spider (although not necessarily with 8 legs). Previous studies were performed using Surface Plasmon Resonance (SPR) to test for possible movement of Spiders on a densely DNA substrate covered matrix [36]. The substrate is complementary in sequence to the DNAzyme legs, and as previously mentioned, is a DNA substrate with a single RNA bond. In turn, the Spiders are attached to this matrix until a metal ion is added to the environment, which promotes catalysis. Catalysis in turn allows a substrate with a leg attached to be cleaved so that the leg is now free to hybridize with another substrate. The multivalency of Spiders ensures tight binding to the surface and therefore allows a large number of substrate sites to be visited and cleaved before the Spider is released [36]. From SPR data, it was apparent that Spiders were capable of continuous movement in a three-dimensional substrate matrix. Depending on the number of legs, the rate of movement as well as the processivity of Spiders can be varied, demonstrating their dexterity to be used for various applications.

Although Spiders have been shown to cleave substrates and move along a substrate covered matrix, Spider movement on a field of substrate has never been directly observed. Therefore, in Chapter Four, fluorescently labeled Spiders are tracked in real-time using single particle tracking microscopy in collaboration with scientists who developed these nano-machines.

1.4 Fluorescence techniques in biophysics

During the last few decades, there has been a remarkable growth in the use of fluorescence spectroscopy in biophysical studies. Fluorescence is now being employed to understand the properties and dynamics of proteins and nucleic acids, and is applied to clinical chemistry and cellular imaging for identifying the localization and movement of intracellular components. Continued advances in fluorescent techniques and instrumentation have fueled the applications of fluorescence spectroscopy to biological systems. In particular, fluorescence resonance energy transfer has emerged as a phenomenal tool to probe distances, conformational changes and dynamics of ribozymes. In addition, the use of nucleotide analogs as fluorescent probes, as well as of fluorescence microscopy to track nucleic acids or proteins in real-time, has rapidly expanded.

Fluorescence Resonance Energy Transfer

Fluorescence Resonance Energy Transfer (FRET), also called Förster transfer, is the nonradiative transfer of excited state energy between an initially excited donor to an acceptor, where ever the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor [37, 38]. Energy transfer occurs without the appearance of a photon but is the result of long range dipole-dipole interactions between the donor and the acceptor. Energy transfer is strongly dependent on the distance between the donor and acceptor and occurs over distances comparable to the dimensions of macromolecules, typically in the range of 20-90 Å. Hence it is commonly used to measure distances or to

observe the dynamics in the distance between two sites of a protein or a nucleic acid. The rate of energy transfer (k_T) is then given by:

$$k_T = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6$$

where τ_D is the fluorescence decay time of the donor in the absence of an acceptor, r is the distance between the donor and the acceptor, and R_0 is the Förster distance at which energy transfer is 50% efficient [37, 38]. R_0 is given by:

$$R_0^6 = \frac{9000(\ln 10)\kappa^2 Q_D}{128\pi^5 N n^4}\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda$$

where Q_D is the quantum yield of the donor in the absence of acceptor, n is the refractive index of the medium, N is Avogadro's number, κ^2 is the factor describing the relative orientation in space of the donor and acceptor transition dipoles, $F_D(\lambda)$ is the corrected fluorescence intensity of the donor in the wavelength range to λ to $\lambda+\Delta\lambda$, with the total intensity normalized to unity, and $\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor at λ [37].

Single molecule FRET

Observing ribozyme activity in bulk FRET measurements reveals essential conformational rearrangements and folding pathways. Kinetic information about these pathways and statistics can also be extracted from ensemble averaged data. However, to

truly understand the unique characteristics of molecules, it is essential to examine them at the single molecule level. Methods in single molecule FRET (smFRET) have been developed in the recent past that allow one to probe the individuality of each molecule using strategically positioned fluorophores [39-43]. Frequently, the highly photostable fluorophore pair Cy3 and Cy5 has been used for this purpose. To this end, smFRET has revealed the heterogeneity and dynamics of ribozymes that give new insight into their characteristics.

First smFRET studies were done using near-field scanning optical microscopy (NSOM), where Tetramethylrhodamine (TMR) and Texas Red fluorophores were used as donor and acceptor respectively [44]. In NSOM, a sub-wavelength aperture is placed is close proximity to the sample, where it allows the mapping of three-dimensional transition dipole orientations of fluorophores. The system studied was 10- or 20-base pair long double-stranded DNA molecules. Donor and acceptor emissions were recorded with an integration time of 10 ms and both images were obtained simultaneously by separating the emission with a dichroic mirror. Another early example of smFRET use was to observe conformational changes of Staphylococcal nuclease (SNase) [45]. Although many ensemble studies regarding its catalytic mechanism and folding pathways were done with this calcium dependent enzyme, it was not clear how ligand induced conformational changes occur. A laser scanning confocal microscope with a 0.4 μm laser spot was used to image surface immobilized SNase molecules labeled with the FRET pair TMR and Cy5. Their emission time trajectories were measured and FRET efficiencies were calculated as described below [45].

There are several methods for observing smFRET in proteins and nucleic acids. One popular method is total internal reflection fluorescence microscopy (TIRFM), which utilizes total internal reflection at the interface of an optically dense medium such as quartz and an optically less dense medium such as water [46]. With an excitation light (laser beam) directed at an angle of incidence (θ) larger than the critical angle (θ_c), the beam reflects back into the quartz surface and generates an evanescent wave in the aqueous solution. The critical angle is determined by Snell's law as:

$$\sin\theta_c = \frac{n_s}{n_w}$$

where n_s and n_w are indices of refraction of quartz and water. This evanescent field has the maximum intensity *I(z)* at the surface and will decay exponentially with distance away from the surface. This decay is given by:

$$I(z) = I_0 \exp(\frac{-z}{d}) \text{ and } d = \frac{\lambda_0}{4\pi}(n_s^2 \sin^2\theta - n_w^2)^{-1/2}, (\theta > \theta_c)$$

where z is the distance from the quartz in to the solution, d is the depth of the evanescent field, and λ_0 is the excitation wavelength in vacuum. Since generally θ is only a few degrees above θ_c, the evanescent wave is about 100 nm in depth [46, 47]. Therefore, only the molecules that are immobilized within the TIRF distance from the quartz surface are excited and will fluoresce under excitation. This eliminates molecules that are floating free in solution and hence will reduce background signal. Depending on the slide preparation for single molecule studies to be used for TIRFM, the sample concentration has to be optimized to give the lowest possible background signal.

Data collection

The fluorescence emissions from the donor and acceptor are collected through the microscope objective and are sent to a dichroic mirror. Here they are separated into two beams depending on their wavelength, and reflective lenses are used to position the donor and acceptor channels onto separate halves of the CCD (Figure 1.6). This allows visualization of molecules and recording of their respective intensities in real-time. The relative smFRET efficiency E is then calculated as:

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

where I_A and I_D are the intensities arising from the acceptor and donor, respectively.

Slide preparation and surface immobilization of molecules for TIRFM

In order to observe single molecules with the microscope, a flow channel is needed. Quartz slides are used for this purpose because of their low autofluorescence. First, two 1-mm holes are drilled on the slide to allow sample access [48]. The re-usable slide is cleaned thoroughly with the detergent alconox and water, and then boiled for 30 min in a solution containing ammonium hydroxide and hydrogen peroxide. Afterwards, the slide is cleaned with water and flamed until dry. Double-sided tape is applied near the two holes with a gap of about 2-3 mm. A coverslip is now placed on top of the tape, and it is glued to the slide with epoxy. This constitutes a flow channel with a depth of about 100-200 nm, which is essentially determined by the thickness of the sticky tape.

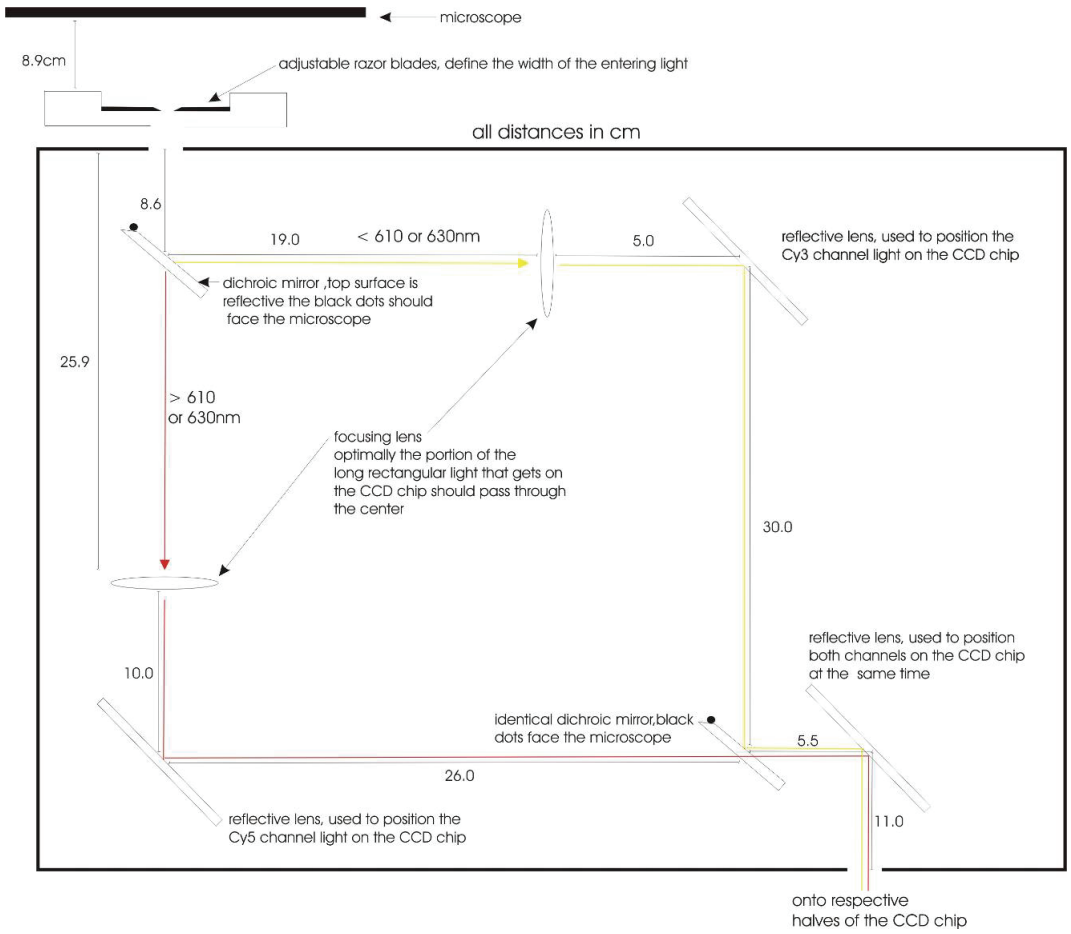


Figure 1.6 Optical setup for single molecule FRET imaging.

There are multiple ways of immobilizing nucleic acids to slides to observe them through TIRFM. A frequently used method is through the biotin-streptavidin interaction. Biotin and streptavidin have the strongest known non-covalent interaction in biology with a dissociation constant as low as 4×10^{-14} M [49]. For use in TIRFM, the flow channel in the quartz slide is treated with biotinylated bovine serum albumin (BSA) that non-specifically adsorbs to quartz tightly. Streptavidin is then added to the slide where it binds to the biotin. As the final step, the molecule of interest labeled with biotin is added and in turn binds to streptavidin [39, 40, 43]. This ensures proper immobilization of the molecule to the quartz slide, which is now ready for TIRFM. This method is used in smFRET experiments described in Chapters Two and Three.

Single-particle tracking

Until recently, the optical resolution limit for two adjacent points observed in a microscope was generally thought to be $\lambda/2\times NA$, where λ is the wavelength of the light source and NA is the numerical aperture. With emission spectra in the visible wavelength range and the maximum NA for a microscope objective of ~1.65, resolving two adjacent spots in a microscopic image would require them to be at a distance of >250 nm [50]. The reason is that the point spread functions (PSF) of two fluorescent spots (from, for example, two single fluorophores) will have a full width at half-maximum of >250 nm and will overlap too substantially to resolve if closer than this distance. However, in the last few years, methods for breaking the diffraction barrier in fluorescence microscopy have emerged that allow one to study biological molecules with improved localization

accuracy, down to the single nanometer range [50-54]. Such methods have been implemented for achieving precise localization of Spiders in real-time to observe them move on a field of substrate.

The earliest examples of quantitative particle tracking were published in the early 1980s by Larry Barak and Watt Webb, where diffusion of a receptor bound low density lipoprotein complex (LDL-RC) was monitored on human fibroblasts [30, 55, 56]. Fluorescence stained probes (biologically active LDL derivative called diI(3)LDL) were used to address the time scale of LDL-RC movement on cell surface to examine its possible interactions with components of the cell cytoskeleton and fluorescent photobleaching recovery experiments were carried out to measure diffusion. Fluorescence was observed with an oil objective and 515 nm excitation of diI(3)LDL. A series of real-time fluorescence images were recorded at 1/30 sec intervals, and positions and displacements of each particle were computed. Furthermore, the average squared displacement was calculated and the diffusion constant was determined [55]. Studies analyzing the clustering and mobility of LDL receptors followed a few years later using similar techniques [57]. In 1994, techniques were developed to detect, recognize and track individual LDL-R molecules and clusters using molecular recognition and simultaneous automated tracking at high precision (down to 30 nm) [58]. More specifically, data acquisition was performed for up to 1,000 images over 45 min using a CCD camera. Each particle's position was determined by the weighted center of mass of its image over the number of pixels that the image occupied. From these positions, the movement of each particle was displayed [58]. In addition, particle tracking techniques

were developed from three-dimensional diffusion studies to the use for optical tweezers [56, 59-62].

Precise nanometer localization of individual molecules using determination of the centroid of fluorescence images of single fluorophores was developed in 2002 [56]. The localization analysis of single particles is dependent on factors such as resolution of the microscope, number of photon collected of the image, background noise, and CCD readout noise. The standard error of the mean (σ_{μ_i}) of the PSF is a measure of localization accuracy and this error can be reduced by increasing the number of photons collected (N):

$$\sigma_{\mu_i} = \sqrt{\left(\frac{s_i^2}{N} + \frac{a^2/12}{N} + \frac{8\pi s_i^4 b^2}{a^2 N^2} \right)}$$

Here, s_i is the standard deviation of the Gaussian distribution that equals 1/2.2 of the PSF width, a is the pixel size, and b is the background. The first term (s_i^2/N) arises from photon noise, the second term represents the effect of the finite pixel size of the camera, and the third term arises from the background signal of the sample [56, 63, 64].

Using this method, Fluorescence Imaging with One Nanometer Accuracy (FIONA) has been developed recently by simply taking the PSF of a fluorophore and locating the center of the fluorescent spot by a two-dimensional Gaussian fit [64]. The fundamental goal was to determine the mean value of a PSF distribution $\mu = (x,y)$ and its uncertainty, σ_{μ_i}. By labeling Myosin V with a single fluorophore at different positions on the calmodulin light chain, its individual step-sizes were measured with 1.5 nm localization using this method for the first time. Their results lead to conclusion that Myosin V moves in a hand-in-hand walking motion rather than the 'inchworm' model

previously suggested [64]. FIONA has also been used for Myosin VI and Kinesin imaging, giving more insight on how they move processively along the cytoskeleton [65, 66]. Thus FIONA was established as a novel imaging technique that can be used in various systems to observe particles with high precision. In Chapter Four, FIONA is used to track Molecular Spiders to obtain their movement in a two dimensional substrate surface.

1.5 Objectives of the dissertation

In light of advances in the biophysical techniques described above, it has become possible to study fundamental aspects of nucleic acids that make them versatile as well their plausible roles in medical and research applications. In this dissertation these possibilities are explored and as a result, functional aspects of the studied systems have become better understood.

The experiments in Chapter Two were designed to compare the performance of the theophylline aptazyme to that of a naturally occurring aptazyme, the *glmS* riboswitch. This aptazyme is potentially important for diagnosing theophylline in human serum, and therefore it is crucial to recognize its limitations. By probing its global and local conformational changes and dynamics as well as its functionality with various ligand concentrations, the characteristics of this aptazyme were observed. By positioning fluorophores at the 3' and the 5' ends of the substrate, ligand induced global conformational changes and dynamics were monitored to identify the differences in aptazyme behavior in the absence and presence of the analyte. Furthermore, by

incorporating 2-aminopurine in either the aptamer or the core domain, local conformational changes upon ligand binding were observed. By combining these two approaches with ligand induced fluorescence cleavage assays, the mechanism of the aptazyme became understood. These new insights were then compared to what is known of the *glmS* riboswitch for contrasting the two aptazymes as well as to generate suggestions for the improvement of aptazymes engineered in the future.

Although the HDV ribozyme has been studied extensively over the years, its U-turn was not identified until recently, when molecular dynamics simulations and activity tests were combined. In order to further characterize the importance of its U-turn, we have begun to experimentally probe the ribozyme with smFRET, and further worked towards probing it with NMR studies. For smFRET studies of the U-turn of the genomic HDV ribozyme, we chose an identical sequence to the construct previously characterized using cleavage assays and Tb^{3+} footprinting [29]. However, in order to use this construct for single molecule studies, it was necessary to make several modifications. These included addition of a tether to immobilize the molecule for smFRET studies, and addition of the Cy3 donor and Cy5 acceptor fluorophores for FRET. Thus, the synthesis, purification, ligation and assembly of several oligonucleotides were required. Chapter Three describes this assembly process in addition to showing preliminary smFRET studies.

Chapter Four sought to understand how Molecular Spiders can randomly walk on a two-dimensional field of substrates using the single particle tracking techniques described above. Experimental and theoretical studies pertaining to Spiders and similar walkers are currently being pursued by a number of laboratories in the Center for

Molecular Cybernetics, which is a National Science Foundation *Chemical Bonding Center*. This 'Spider World' involves many different disciplines and extensive collaborations between diverse experts and has allowed detailed probing of individual molecular robots.

The objective of this dissertation was to study the catalytic mechanisms of nucleic acid enzymes as well as to probe devices developed based on such nucleic acid enzymes. Investigating the structure, function, dynamics, and kinetics of nucleic acid enzymes opens new doors to the development of novel devices such as Molecular Spiders that have many potential applications in biological and medicinal fields. Studies performed in this dissertation have contributed to the further development of the understanding of and ability to manipulate nucleic acids enzymes, by applying advanced biophysical techniques to these fascinating biological systems.

1.6 References

1. *The RNA World*. Third Edition ed, ed. R.F. Gesteland, Cech, T.R., Atkins, J.F.: Cold Spring Harbor Laboratory Press.

2. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.

3. Tucker, B.J. and R.R. Breaker, *Riboswitches as versatile gene control elements.* Curr Opin Struct Biol, 2005. **15**(3): p. 342-8.

4. Long, D.M. and O.C. Uhlenbeck, *Self-cleaving catalytic RNA*. Faseb J, 1993. **7**(1): p. 25-30.

5. Modahl, L.E. and M.M. Lai, *Transcription of hepatitis delta antigen mRNA continues throughout hepatitis delta virus (HDV) replication: a new model of HDV RNA transcription and replication.* J Virol, 1998. **72**(7): p. 5449-56.

6. Hernandez, C. and R. Flores, *Plus and minus RNAs of peach latent mosaic viroid self-cleave in vitro via hammerhead structures.* Proc Natl Acad Sci U S A, 1992. **89**(9): p. 3711-5.

7. Hutchins, C.J., P.D. Rathjen, A.C. Forster, and R.H. Symons, *Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid.* Nucleic Acids Res, 1986. **14**(9): p. 3627-40.

8. Rojas, A.A., A. Vazquez-Tello, G. Ferbeyre, F. Venanzetti, L. Bachmann, B. Paquin, V. Sbordoni, and R. Cedergren, *Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets.* Nucleic Acids Res, 2000. **28**(20): p. 4037-43.

9. Martick, M., L.H. Horan, H.F. Noller, and W.G. Scott, *A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA.* Nature, 2008. **454**(7206): p. 899-902.

10. Ferbeyre, G., V. Bourdeau, M. Pageau, P. Miramontes, and R. Cedergren, *Distribution of hammerhead and hammerhead-like RNA motifs through the GenBank.* Genome Res, 2000. **10**(7): p. 1011-9.

11. Curtis, E.A. and D.P. Bartel, *The hammerhead cleavage reaction in monovalent cations.* Rna, 2001. **7**(4): p. 546-52.

12. Canny, M.D., F.M. Jucker, E. Kellogg, A. Khvorova, S.D. Jayasena, and A. Pardi, *Fast cleavage kinetics of a natural hammerhead ribozyme.* J Am Chem Soc, 2004. **126**(35): p. 10848-9.

13. Bondensgaard, K., E.T. Mollova, and A. Pardi, *The global conformation of the hammerhead ribozyme determined using residual dipolar couplings.* Biochemistry, 2002. **41**(39): p. 11532-42.

14. Bassi, G.S., A.I. Murchie, F. Walter, R.M. Clegg, and D.M. Lilley, *Ion-induced folding of the hammerhead ribozyme: a fluorescence resonance energy transfer study.* Embo J, 1997. **16**(24): p. 7481-9.

15. Rueda, D., K. Wick, S.E. McDowell, and N.G. Walter, *Diffusely bound Mg2+ ions slightly reorient stems I and II of the hammerhead ribozyme to increase the probability of formation of the catalytic core.* Biochemistry, 2003. **42**(33): p. 9924-36.

16. Martick, M. and W.G. Scott, *Tertiary contacts distant from the active site prime a ribozyme for catalysis.* Cell, 2006. **126**(2): p. 309-20.

17. Lewin, A.S. and W.W. Hauswirth, *Ribozyme gene therapy: applications for molecular medicine.* Trends Mol Med, 2001. **7**(5): p. 221-8.

18. Sarver, N., E.M. Cantin, P.S. Chang, J.A. Zaia, P.A. Ladne, D.A. Stephens, and J.J. Rossi, *Ribozymes as potential anti-HIV-1 therapeutic agents.* Science, 1990. **247**(4947): p. 1222-5.

19. Breaker, R.R., *Engineered allosteric ribozymes as biosensor components.* Curr Opin Biotechnol, 2002. **13**(1): p. 31-9.

20. Sekella, P.T., D. Rueda, and N.G. Walter, *A biosensor for theophylline based on fluorescence detection of ligand-induced hammerhead ribozyme cleavage.* Rna, 2002. **8**(10): p. 1242-52.

21. Tinsley, R.A., D.A. Harris, and N.G. Walter, *Magnesium dependence of the amplified conformational switch in the trans-acting hepatitis delta virus ribozyme.* Biochemistry, 2004. **43**(28): p. 8935-45.

22. Lai, M.M., *The molecular biology of hepatitis delta virus.* Annu Rev Biochem, 1995. **64**: p. 259-86.

23. Casey, J.L., G.A. Niro, R.E. Engle, A. Vega, H. Gomez, M. McCarthy, D.M. Watts, K.C. Hyams, and J.L. Gerin, *Hepatitis B virus (HBV)/hepatitis D virus (HDV) coinfection in outbreaks of acute hepatitis in the Peruvian Amazon basin: the roles of HDV genotype III and HBV genotype F.* J Infect Dis, 1996. **174**(5): p. 920-6.

24. Tinsley, R.A. and N.G. Walter, *Long-range impact of peripheral joining elements on structure and function of the hepatitis delta virus ribozyme.* Biol Chem, 2007. **388**(7): p. 705-15.

25. Sefcikova, J., M.V. Krasovska, N. Spackova, J. Sponer, and N.G. Walter, *Impact of an extruded nucleotide on cleavage activity and dynamic catalytic core conformation of the hepatitis delta virus ribozyme.* Biopolymers, 2007. **85**(5-6): p. 392-406.

26. Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna, *Crystal structure of a hepatitis delta virus ribozyme.* Nature, 1998. **395**(6702): p. 567-74.

27. Strobel, S.A., *Ribonucleic general acid.* Nat Chem Biol, 2005. **1**(1): p. 5-6.

28. Das, S.R. and J.A. Piccirilli, *General acid catalysis by the hepatitis delta virus ribozyme.* Nat Chem Biol, 2005. **1**(1): p. 45-52.

29. Sefcikova, J., M.V. Krasovska, J. Sponer, and N.G. Walter, *The genomic HDV ribozyme utilizes a previously unnoticed U-turn motif to accomplish fast site-specific catalysis.* Nucleic Acids Res, 2007. **35**(6): p. 1933-46.

30. Barak, L.S. and W.W. Webb, *Fluorescent low density lipoprotein for observation of dynamics of individual receptor complexes on cultured human fibroblasts.* J Cell Biol, 1981. **90**(3): p. 595-604.

31. Mitchell, A., C.R. Dass, L.Q. Sun, and L.M. Khachigian, *Inhibition of human breast carcinoma proliferation, migration, chemoinvasion and solid tumour growth by DNAzymes targeting the zinc finger transcription factor EGR-1.* Nucleic Acids Res, 2004. **32**(10): p. 3065-9.

32. Wu, Y., L. Yu, R. McMahon, J.J. Rossi, S.J. Forman, and D.S. Snyder, *Inhibition of bcr-abl oncogene expression by novel deoxyribozymes (DNAzymes).* Hum Gene Ther, 1999. **10**(17): p. 2847-57.

33. Schubert, S., D.C. Gul, H.P. Grunert, H. Zeichhardt, V.A. Erdmann, and J. Kurreck, *RNA cleaving '10-23' DNAzymes with enhanced stability and activity.* Nucleic Acids Res, 2003. **31**(20): p. 5982-92.

34. Santoro, S.W. and G.F. Joyce, *A general purpose RNA-cleaving DNA enzyme.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4262-6.

35. Li, J., W. Zheng, A.H. Kwon, and Y. Lu, *In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme.* Nucleic Acids Res, 2000. **28**(2): p. 481-8.

36. Pei, R., S.K. Taylor, D. Stefanovic, S. Rudchenko, T.E. Mitchell, and M.N. Stojanovic, *Behavior of polycatalytic assemblies in a substrate-displaying matrix.* J Am Chem Soc, 2006. **128**(39): p. 12693-9.

37. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*. 1999: Plenum Publishing Corporation.

38. Forster, T., *Zwischenmolekulare Energiewanderung und Fluoreszenz.* Ann. Physik 1948. **437**(55).

39. Pereira, M.J., E.N. Nikolova, S.L. Hiley, D. Jaikaran, R.A. Collins, and N.G. Walter, *Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis.* J Mol Biol, 2008. **382**(2): p. 496-509.

40. Zhuang, X., H. Kim, M.J. Pereira, H.P. Babcock, N.G. Walter, and S. Chu, *Correlating structural dynamics and function in single ribozyme molecules.* Science, 2002. **296**(5572): p. 1473-6.

41. Zhuang, X., L.E. Bartley, H.P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu, *A single-molecule study of RNA catalysis and folding.* Science, 2000. **288**(5473): p. 2048-51.

42. Sako, Y., S. Minoghchi, and T. Yanagida, *Single-molecule imaging of EGFR signalling on the surface of living cells.* Nat Cell Biol, 2000. **2**(3): p. 168-72.

43. Ditzler, M.A., D. Rueda, J. Mo, K. Hakansson, and N.G. Walter, *A rugged free energy landscape separates multiple functional RNA folds throughout denaturation.* Nucleic Acids Res, 2008. **36**(22): p. 7088-99.

44. Ha, T., T. Enderle, D.F. Ogletree, D.S. Chemla, P.R. Selvin, and S. Weiss, *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, 1996. **93**(13): p. 6264-8.

45. Ha, T., A.Y. Ting, J. Liang, W.B. Caldwell, A.A. Deniz, D.S. Chemla, P.G. Schultz, and S. Weiss, *Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism.* Proc Natl Acad Sci U S A, 1999. **96**(3): p. 893-8.

46. Burghardt, T.P. and D. Axelrod, *Total internal reflection fluorescence study of energy transfer in surface-adsorbed and dissolved bovine serum albumin.* Biochemistry, 1983. **22**(4): p. 979-85.

47. Rueda, D. and N.G. Walter, *Single molecule fluorescence control for nanotechnology.* J Nanosci Nanotechnol, 2005. **5**(12): p. 1990-2000.

48. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.

49. Holmberg, A., A. Blomstergren, O. Nord, M. Lukacs, J. Lundeberg, and M. Uhlen, *The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures.* Electrophoresis, 2005. **26**(3): p. 501-10.

50. Toprak, E. and P.R. Selvin, *New fluorescent tools for watching nanometer-scale conformational changes of single molecules.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 349-69.

51. Dyba, M. and S.W. Hell, *Focal spots of size lambda/23 open up far-field fluorescence microscopy at 33 nm axial resolution.* Phys Rev Lett, 2002. **88**(16): p. 163901.

52. Willig, K.I., S.O. Rizzoli, V. Westphal, R. Jahn, and S.W. Hell, *STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis.* Nature, 2006. **440**(7086): p. 935-9.

53. Willig, K.I., R.R. Kellner, R. Medda, B. Hein, S. Jakobs, and S.W. Hell, *Nanoscale resolution in GFP-based microscopy.* Nat Methods, 2006. **3**(9): p. 721-3.

54. Walter, N.G., C.Y. Huang, A.J. Manzo, and M.A. Sobhy, *Do-it-yourself guide: how to use the modern single-molecule toolkit.* Nat Methods, 2008. **5**(6): p. 475-89.

55. Barak, L.S. and W.W. Webb, *Diffusion of low density lipoprotein-receptor complex on human fibroblasts.* J Cell Biol, 1982. **95**(3): p. 846-52.

56. Thompson, R.E., D.R. Larson, and W.W. Webb, *Precise nanometer localization analysis for individual fluorescent probes.* Biophys J, 2002. **82**(5): p. 2775-83.

57. Gross, D. and W.W. Webb, *Molecular counting of low-density lipoprotein particles as individuals and small clusters on cell surfaces.* Biophys J, 1986. **49**(4): p. 901-11.

58. Ghosh, R.N. and W.W. Webb, *Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules.* Biophys J, 1994. **66**(5): p. 1301-18.

59. Kao, H.P. and A.S. Verkman, *Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position.* Biophys J, 1994. **67**(3): p. 1291-300.

60. Edidin, M., S.C. Kuo, and M.P. Sheetz, *Lateral movements of membrane glycoproteins restricted by dynamic cytoplasmic barriers.* Science, 1991. **254**(5036): p. 1379-82.

61. Peters, I.M., Y. van Kooyk, S.J. van Vliet, B.G. de Grooth, C.G. Figdor, and J. Greve, *3D single-particle tracking and optical trap measurements on adhesion proteins.* Cytometry, 1999. **36**(3): p. 189-94.

62. Sako, Y., A. Nagafuchi, S. Tsukita, M. Takeichi, and A. Kusumi, *Cytoplasmic regulation of the movement of E-cadherin on the free cell surface as studied by optical tweezers and single particle tracking: corralling and tethering by the membrane skeleton.* J Cell Biol, 1998. **140**(5): p. 1227-40.

63. Yildiz, A. and P.R. Selvin, *Fluorescence imaging with one nanometer accuracy: application to molecular motors.* Acc Chem Res, 2005. **38**(7): p. 574-82.

64. Yildiz, A., J.N. Forkey, S.A. McKinney, T. Ha, Y.E. Goldman, and P.R. Selvin, *Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization.* Science, 2003. **300**(5628): p. 2061-5.

65. Yildiz, A., H. Park, D. Safer, Z. Yang, L.Q. Chen, P.R. Selvin, and H.L. Sweeney, *Myosin VI steps via a hand-over-hand mechanism with its lever arm undergoing fluctuations when attached to actin.* J Biol Chem, 2004. **279**(36): p. 37223-6.

66. Yildiz, A., M. Tomishige, R.D. Vale, and P.R. Selvin, *Kinesin walks hand-over-hand.* Science, 2004. **303**(5658): p. 676-8.

CHAPTER TWO

LEAKAGE AND SLOW ALLOSTERY LIMIT PERFORMANCE OF SINGLE DRUG-SENSING APTAZYME MOLECULES BASED ON THE HAMMERHEAD RIBOZYME[1]

2.1 Introduction

Riboswitches represent a recently discovered gene regulatory paradigm wherein binding of a small-molecule metabolite to an RNA motif embedded in the 5′- or 3′-untranslated region of a pre-mRNA controls gene expression via transcription, translation or alternative splicing [1-7]. Expression of an estimated 2-3% of all bacterial genes and an unknown number of eukaryotic genes is controlled in this economic protein-free fashion [1, 3, 6]. Ligand binding to a riboswitch is thought to typically induce an RNA conformational change by either kinetic or thermodynamic control that affects an adjacent expression platform [8]. A notable exception is the *glmS* catalytic riboswitch, or aptazyme, located upstream of the *glmS* gene of Gram-positive bacteria [1, 3, 9]. It employs its glucosamine-6-phosphate (GlcN6P) ligand as a chemical cofactor [10] in a self-cleavage reaction that causes mRNA degradation [11], without the involvement of any detectable conformational change [12-15].

[1]Adapted with permission from de Silva, C. and Walter, N.G., *RNA* 15, 76-84. Copyright 2009, Cold Spring Harbor Press Laboratory.

In pioneering work initiated years before naturally occurring riboswitches were discovered, artificial aptazymes were engineered for use as biosensor components and artificial genetic switches by fusing an *in vitro* selected aptamer to a naturally occurring ribozyme, affording a ligand dependent RNA conformational change as mechanism that activates the ribozyme motif [2, 16-21]. In a prototypical such biosensor aptazyme, a theophylline-specific aptamer is incorporated via a communication module into Stem II of the hammerhead ribozyme [16, 22, 23] (Fig. 2.1A), the smallest known natural catalytic RNA [24-27]. Theophylline binding activates the ribozyme, leading to cleavage of a substrate strand in *cis* or *trans*, depending on strand connectivity, which in turn can be coupled to a radioactive [16] or fluorescent signal [23, 28]. Theophylline, or dimethylxanthine (Fig. 2.1B), is a bronchodilating drug used in therapy for respiratory diseases including asthma, bronchitis, and emphysema. It is characterized by a narrow therapeutic range of between ~55 and 110 μM in blood plasma, with numerous side effects above this range, necessitating that an administered dose be closely monitored [29, 30]. The theophylline aptamer was previously *in vitro* selected to have micromolar ligand affinity and to discriminate 10,000-fold against the ubiquitous, structurally similar caffeine [23, 31, 32] (Fig. 2.1B), making it an attractive aptazyme component. Yet in tests of its dynamic range, an important quality of molecular tools in analytical applications [33], the theophylline aptazyme falls about three orders of magnitude short [22, 23] of the 100,000-fold ligand-induced activation observed for the naturally occurring *glmS* aptazyme [9, 10].

To further optimize engineered aptazymes for use as biosensor components [16] and RNA-based switches for controlling gene expression [19] it is critical to understand

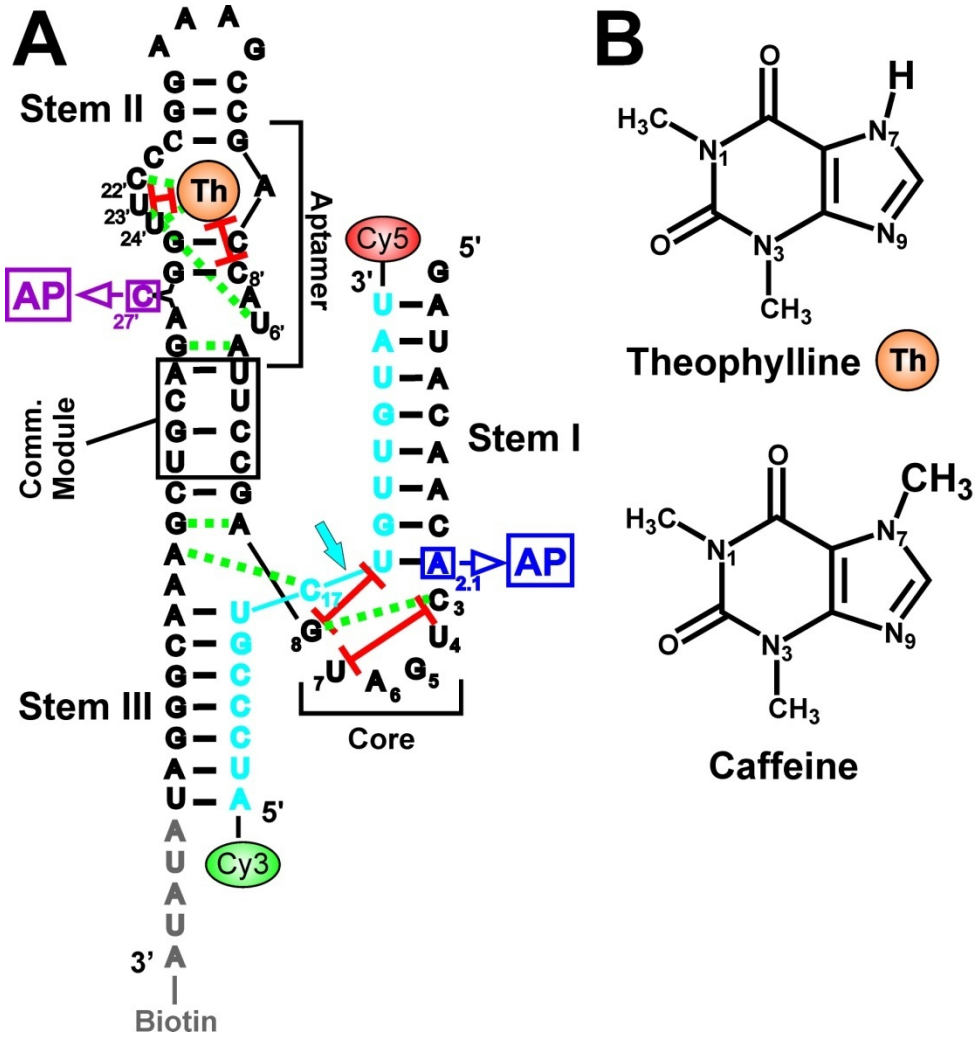


Figure 2.1 Theophylline aptazyme used in this study. (A) The assembled aptazyme (black) and substrate strands (cyan), forming Stems I, II, and III and the aptamer, communication, and catalytic core domains as indicated with canonical numbering schemes. The substrate is labeled with a FRET donor (Cy3) and acceptor (Cy5), and its cleavage site is indicated by an arrow. The aptazyme strand is 3′-biotinylated for surface immobilization during smFRET experiments. Theophylline (Th) ligand binding to the aptamer domain is probed by substituting either C27′ or A2.1 with AP as indicated. Important hydrogen bond and stacking interactions as observed by X-ray crystallography [49] are indicated as green dashed and red solid lines, respectively. (B) Chemical structures of theophylline and caffeine.

the weaknesses of existing aptazyme designs. Here we probe the allosteric mechanism of the prototypical theophylline aptazyme by single molecule and ensemble fluorescence spectroscopy. We find that the engineered aptazyme is characterized by leaky and slow conformational changes, rationalizing the superior performance of the cofactor-activated *glmS* aptazyme. These findings lead to suggest specific strategies for improvement of artificial engineered aptazymes.

2.2 Results

The prototypical theophylline-dependent aptazyme shows limited dynamic range and significant background cleavage in the absence of ligand

A previously developed ensemble fluorescence resonance energy transfer (FRET) assay for the theophylline aptazyme where a 5'-donor (Cy3) and 3'-acceptor (Cy5) labeled substrate (which positions the fluorophores into Stems III and I, respectively, of the hammerhead ribozyme, Fig. 2.2A) is cleaved in *trans*, leading to a conveniently monitored FRET decrease upon theophylline addition [23, 28]. This assay is utilized here to measure a dynamic range for theophylline detection as a key indicator of biosensor performance [33] that is quite modest, consistent with earlier studies [22, 23]: We observe an ~20-fold increase in cleavage activity between 0 and 10 mM theophylline at 10 mM $MgCl_2$, which increases to ~70-fold at 100 mM $MgCl_2$ (Fig 2.2B). This dynamic range does not measure up to that of the naturally occurring *glmS* aptazyme, which shows an ~100,000-fold activity increase between 0 and 10 mM GlcN6P [9, 10]. Notably, a

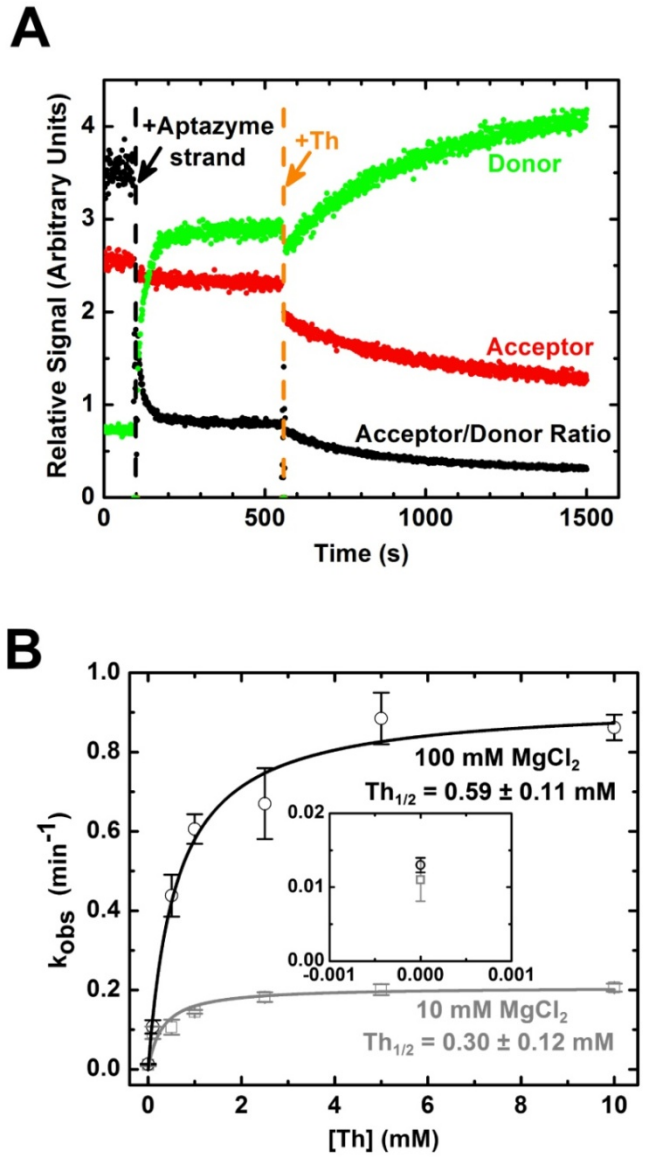


Figure 2.2 Steady-state ensemble fluorescence signals. (A) FRET labeled cleavable substrate upon sequential addition of aptazyme strand and 10 mM theophylline (Th) under standard conditions of 50 mM Tris-HCl, pH 7.5, 25 mM DTT, 100 mM $MgCl_2$, at 25˚C. (B) Theophylline concentration dependence of the cleavage rate constants derived from single-exponential fits to acceptor/donor ratio time cses like that in panel c, yielding the indicated theophylline half-titration points, $Th_{1/2}$, at 10 mM and 100 mM $MgCl_2$ (Materials and Methods). The inset zooms in on the rate constants in the absence of theophylline.

significant background ("leakage") cleavage rate constant of 0.013 min^{-1} is observed in the absence of theophylline (100 mM $MgCl_2$, which is defined as standard conditions henceforth; inset of Fig. 2.2B) and confirmed by radioactive cleavage assays ([23] and data not shown), which drastically limits the dynamic range (calculated as a fold increase relative to the background reaction).

Single molecule FRET reveals distinct global conformational isomers

It can be hypothesized that the cause for the modest performance of the theophylline aptazyme compared to the naturally occurring *glmS* ribozyme may be linked to the fact that a ligand-induced conformational change is required for the theophylline aptazyme, while none is found in the natural aptazyme [12-15]. To test this hypothesis and probe for global conformational changes upon ligand binding in the absence of dynamics caused by catalysis, we studied the aptazyme by single molecule FRET (smFRET) in the presence of a non-cleavable 2′-O-methyl substrate modification at the cleavage site (C17) (Fig. 2.1A). Fluorescence of Cy3 and Cy5 on the 5′- and 3′-ends of the substrate strand, respectively, was monitored in real-time using prism-based total internal reflection fluorescence microscopy (TIRFM) on surface immobilized aptazymes, essentially as described [34-36]. Attachment of the smFRET donor-acceptor pair to Stems I and III (Fig. 2.1A) was the same as for ensemble cleavage assays (Fig. 2.2A) and chosen based on previous ensemble FRET studies of the hammerhead ribozyme suggesting that Stems I and III are proximal in an inactive conformation, but become distal upon Mg^{2+} titration and catalytic activation [37-39]. Indeed, smFRET time

trajectories reveal anti-correlated changes in the donor and acceptor signals in the presence of 10 mM theophylline (which is saturating as judged from the activity titration in Fig. 2.2B), with states of high (H, ~0.8) and low (L, ~0.2) smFRET (Fig. 2.3A). Observation of an H state is consistent with a global conformation in which Stems I and III come into close proximity (below the Förster distance of Cy3 and Cy5 of ~54 Å [40]), while the L state is consistent with a large distance between the two stems (well beyond the Förster distance). The fact that these H and L states are readily detected and persist for seconds to minutes at a time suggests that they are structurally quite stable and well defined. smFRET thus uncovers the existence of an equilibrium distribution of previously unobserved, distinct conformational isomers in a hammerhead ribozyme with either proximal or distal Stems I and III.

Single aptazymes access pairs of high and low smFRET states of distinct kinetic behavior

Aptazymes switch between the H and L states with kinetics that are well resolved at the single molecule level (Fig. 2.3A). A prominent long-lived L state is observed in many trajectories, while molecules that are in the H state take frequent, but brief excursions to the L state (Fig. 2.3B). An aggregate smFRET histogram of 45 molecules underscores the resulting preference (68%) for time spent in the L state while in the presence of 10 mM theophylline (Fig. 2.3C). Notably, a similar bimodal smFRET distribution is observed in the absence of theophylline (also 45 molecules), although with a less abundant L state (39%, Figure 2.3C). Histograms of different subsets of molecules

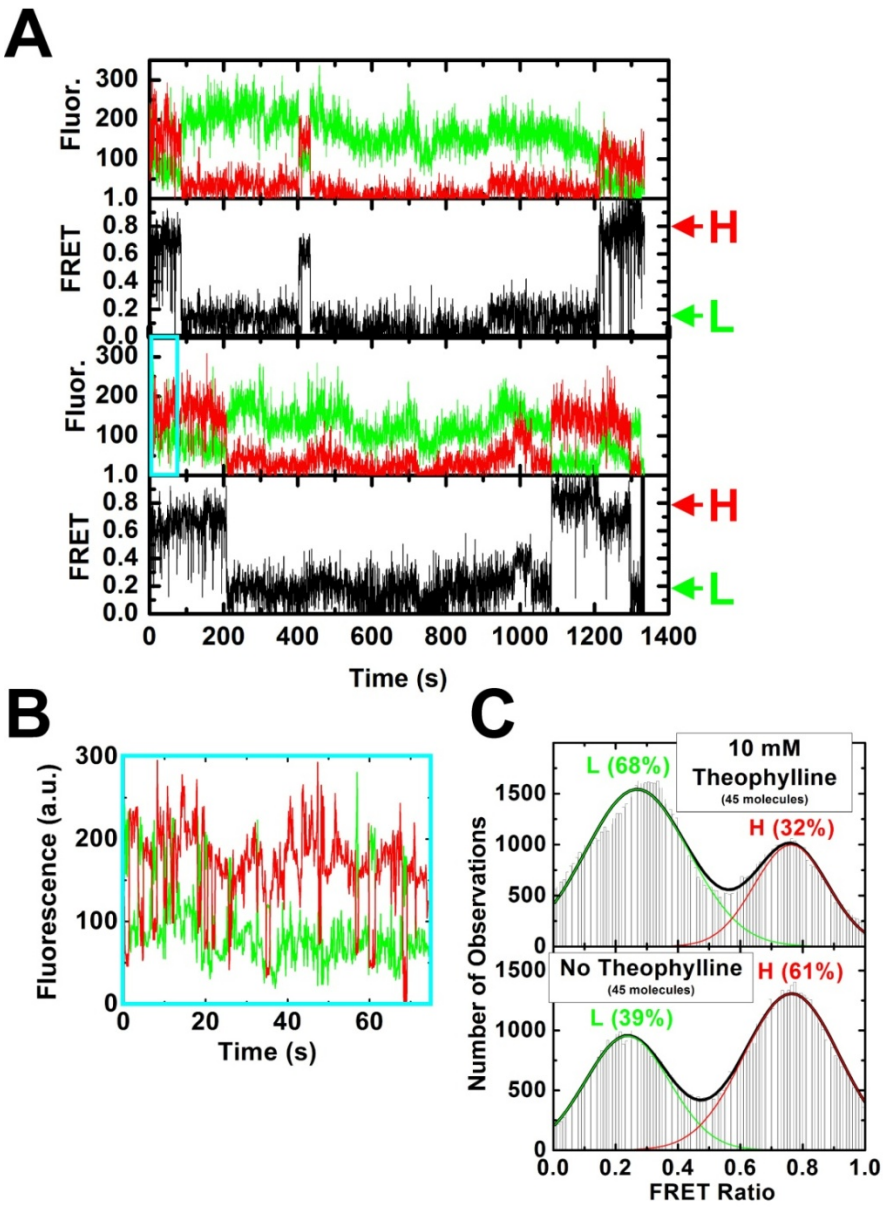


Figure 2.3 smFRET analysis of the theophylline aptazyme. (A) Representative smFRET time traces of two aptazyme molecules (top and bottom pairs of plots) under standard conditions in the presence of 10 mM theophylline. The raw Cy3 donor and Cy5 acceptor fluorescence signals are green and red, respectively, and the corresponding smFRET ratio ($I_A/(I_A+I_D)$) is shown in black with the observed H and L states indicated. (B) Expanded portion from the cyan box in panel a, showing rapid conformational changes within the molecule. (C) Aggregate smFRET histograms in the presence and absence of 10 mM theophylline.

produce similar results (data not shown). The equilibrium shift towards the L state upon addition of theophylline leads us to hypothesize that this state is associated with the observed ligand-induced catalytic activation.

Rate constants for interconversion of the H to the L state and vice versa in the presence of 10 mM theophylline are deduced from cumulative plots of all single molecule dwell times in the H and L states, respectively (Fig. 2.4A). Double-exponentials are required to fit both sets of dwell times (Fig. 2.4A also shows the best single-exponential fits for comparison), in accord with the heterogeneous behavior of single molecules over extended time periods in Figure 2.3A. The deviation from simple single-exponential kinetics is further highlighted by plotting the probability density functions (Fig. 2.4B). The rate constants from the cumulative dwell time plots are corrected for photobleaching and observation time limitations as previously described [41]. The resulting pair of H-to-L transition rate constants (0.21 s^{-1} and 0.05 s^{-1}; ~4-fold difference) are less distinct from one another than those characterizing the L-to-H transition (0.93 s^{-1} and 0.0007 s^{-1}; ~1300-fold difference), consistent with the qualitative observation that a very short-lived and a very long-lived L state co-exist in single molecule trajectories (Fig. 2.3A). Notably, we observe no correlation (memory) between either consecutive L state dwell times or between the L and adjacent H state dwell times (Fig. 2.5), further suggesting that each molecule can freely access each of the underlying f states (two H, two L states) without stochastic bias from its preceding behavior. Similar observations of f rate constants and associated states extend to conditions of low and zero theophylline (Fig. 2.6).

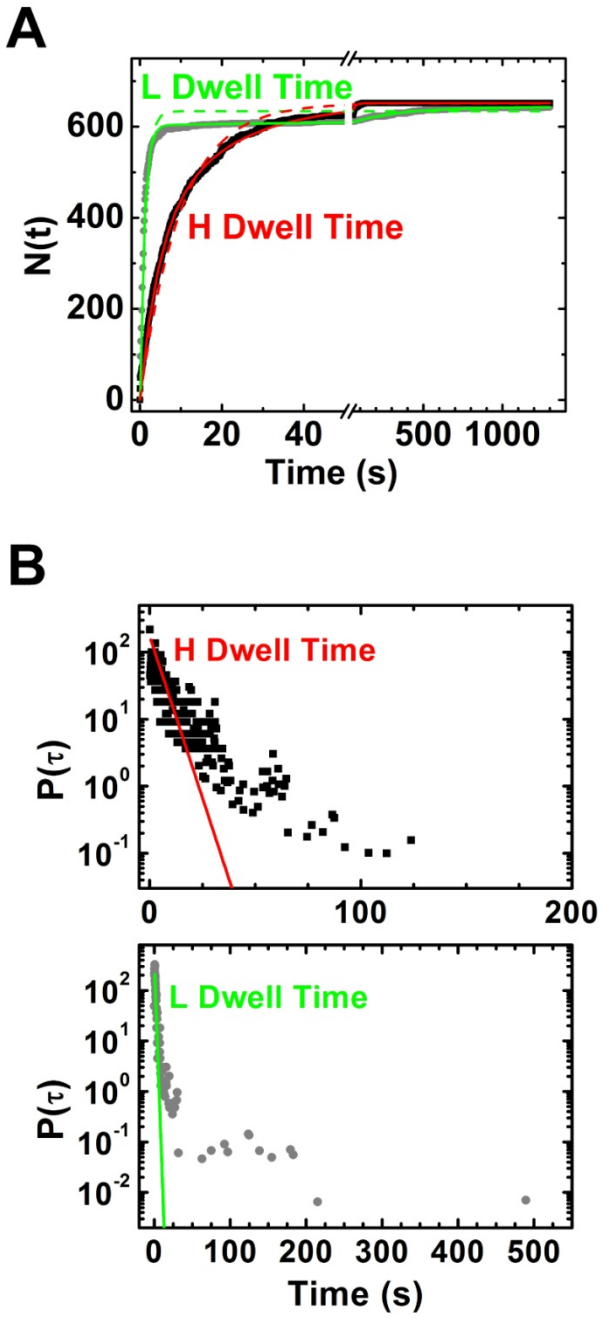


Figure 2.4 Cumulative dwell time histograms. (A) L and H smFRET states for calculating the L-to-H and H-to-L transition rate constants, respectively, and their corresponding fractions at 10 mM theophylline (Materials and Methods). (B) Probability density plots for the H and L state dwell times. The steep decay portions were fit with single exponentials (colored lines) to highlight deviations thereof at 10mM theophylline.

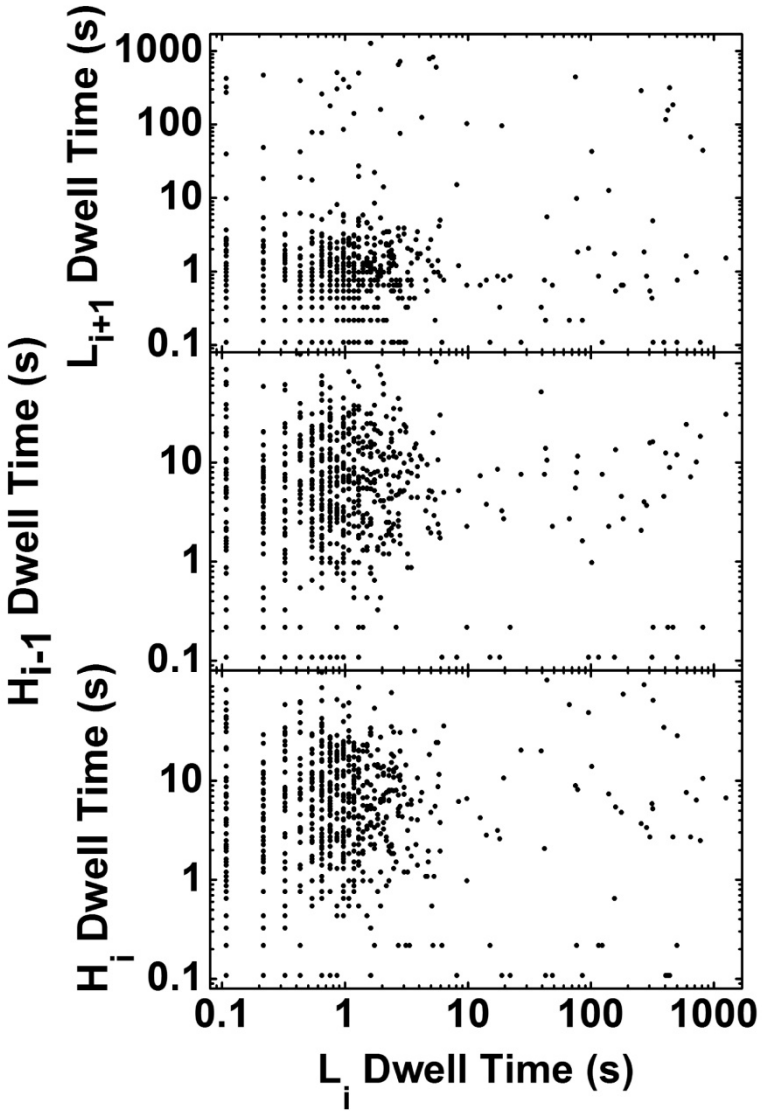


Figure 2.5 Dwell time analysis of L and H states. Plots of the immediately following L state dwell time (L_{i+1}) as well as immediately following and preceding H dwell times (H_{i+1} and H_i, respectively) as a function of a given L dwell time L_i at 10 mM theophylline.

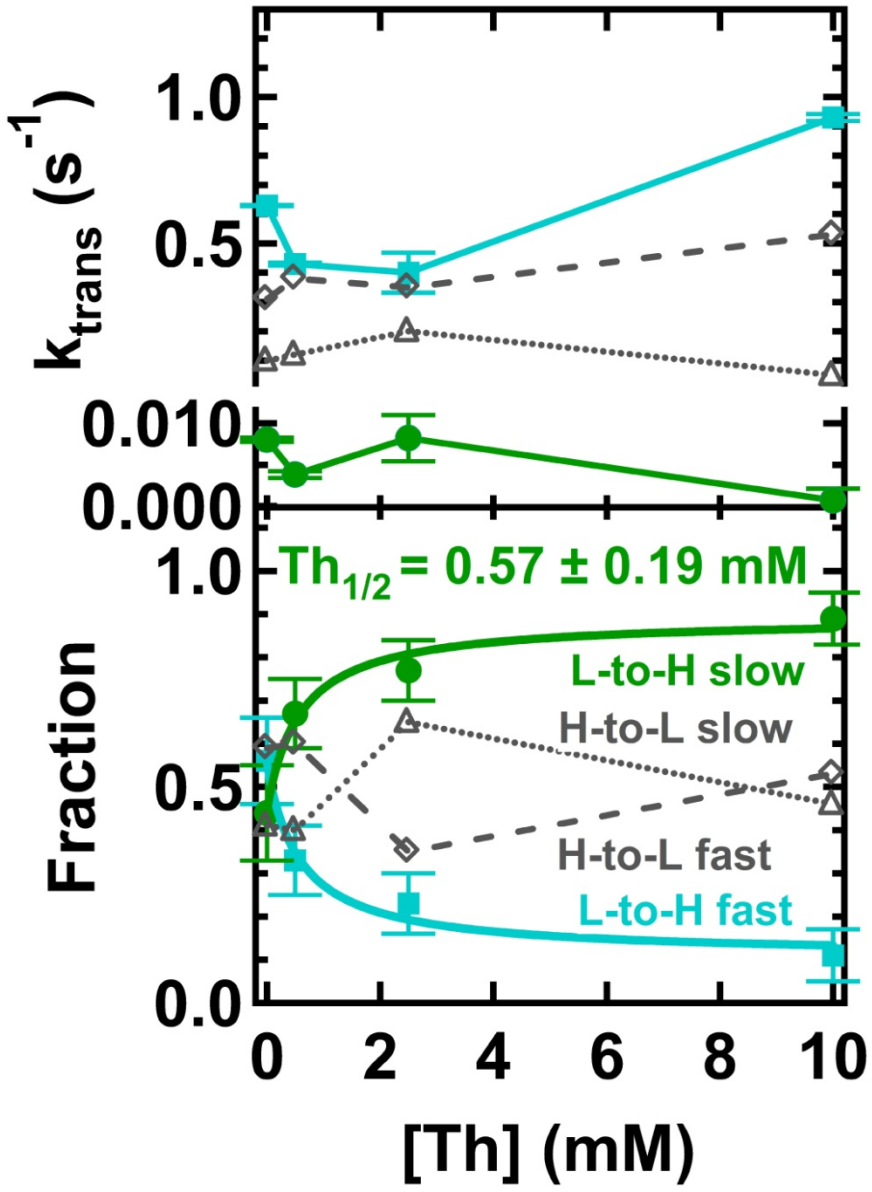


Figure 2.6 Theophylline dependence of the aptazyme. Corrected transition rate constants (top) and their corresponding fractions (bottom) as deduced from cumulative dwell time histograms similar to the one shown in panel D (Materials and Methods). The fractions of the two L-to-H rate constants were fit globally to obtain the indicated theophylline half-titration point $Th_{1/2}$.

Catalytic activity is associated with the long-lived low smFRET state

The corrected H-to-L and L-to-H transition rate constants are shown in Figure 2.6 (top panel) over varying theophylline concentrations. The bottom panel of Fig. 2.6 shows the relative fractional amplitudes of the transitions (adding up to unity), as calculated from the exponential fits to the cumulative dwell time histograms (Materials and Methods). All rate constants and the relative fractional amplitudes of the H-to-L rate constants vary little and/or show no consistent trend. By contrast, the fractional amplitude of the slower L-to-H transition significantly increases with the theophylline concentration (bottom panel, green curve) while the faster L-to-H fractional amplitude decreases (bottom panel, cyan curve), observations that are consistent with an increasing dominance of the long-lived L state as evidenced in the aggregate smFRET histogram of Fig. 2.3C. Global fitting of the two L-to-H fractional amplitudes yields a theophylline half-titration point of $Th_{1/2}$ = 0.57 ± 0.19 mM, within error identical to the $Th_{1/2}$ value (0.59 ± 0.11 mM) of catalytic activation under the same conditions (100 mM $MgCl_2$, Fig. 2.2B). This identity strongly supports a link between catalytic activity and specifically the long-lived L state. We propose that this long-lived L state encompasses the conformationally docked active state of the hammerhead-derived theophylline aptazyme, which forms the expected [26, 27] Y-shape with distal Stems I and III (Fig. 2.1A), while the H states represent undocked, catalytically inactive conformations.

Site-specific 2-aminopurine fluorescence probing provides evidence for slow allosteric signal transduction

To also probe for local conformational changes in the aptamer and catalytic core domains, we substituted the fluorescent adenine analog 2-aminopurine (AP) for the non-conserved C27′ and A2.1 residues, respectively, in separate aptazyme mutants (Fig. 2.1A). Using either stopped-flow or manual mixing combined with steady-state fluorescence detection, changes in AP fluorescence as a probe for local base stacking interactions [42] were monitored upon addition of varying theophylline concentrations. As expected, the AP fluorescence increases in the aptamer (C27′→AP) mutant upon addition of 10 mM theophylline (Fig. 2.7A, upper panel) since nucleotide 27′ (Fig. 2.1A) becomes solvent-exposed in the ligand-bound aptamer [43]. By contrast, the AP fluorescence decreases in the core (A2.1→AP) mutant upon 10 mM theophylline addition (Fig. 2.7A, lower panel), presumably due to increased base stacking of the core around the U1.1-AP2.1 base pair [44] (Fig. 2.1A).

The kinetics of the structural rearrangement detected by AP in the aptamer and core mutants are strikingly different (Fig. 3). The rate constant of the local conformational change in the aptamer domain, as measured in the aptamer mutant, becomes as fast as 63 s^{-1} at saturating theophylline concentrations, only ~7-fold slower than in the isolated aptamer [43]. The observed saturation behavior with $Th_{1/2} < 0.1$ mM suggests that under these conditions the conformational change, not theophylline binding, is rate-limiting (Fig. 2.7B). Notably, the conformational changes in the catalytic core, detected in the core mutant, occur at a much slower rate constant of at most 0.08 s^{-1} under

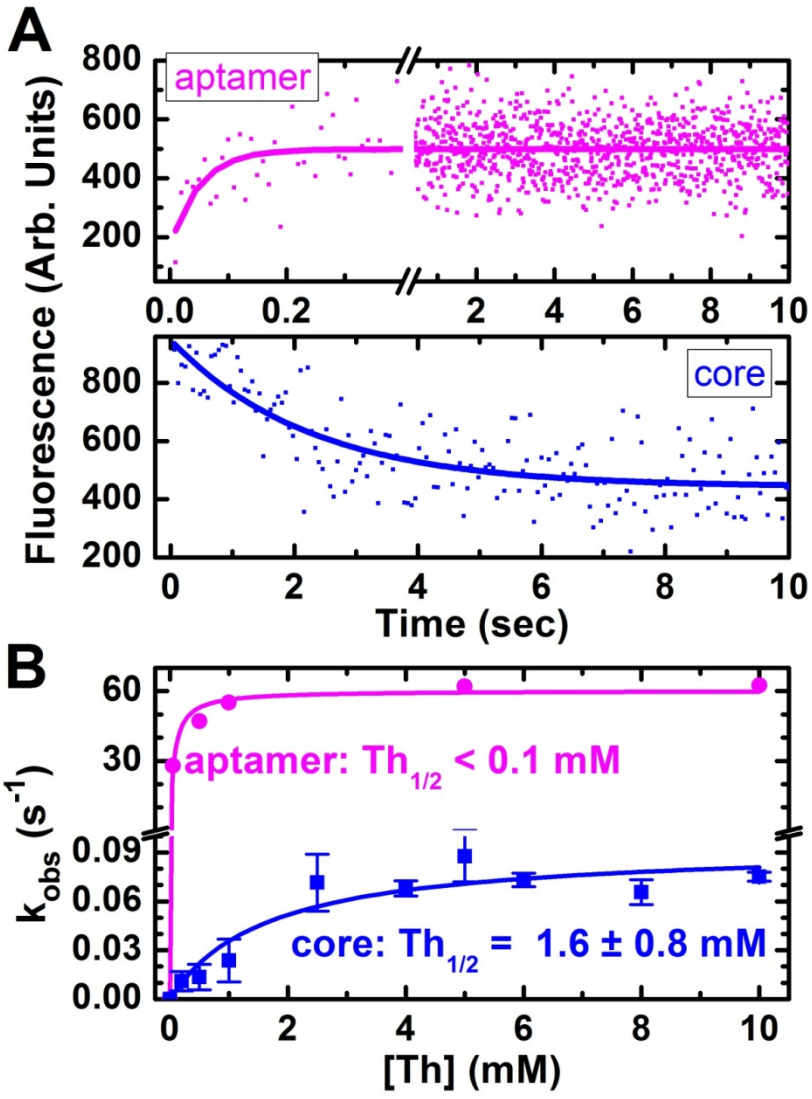


Figure 2.7 Fluorescence changes upon addition of 10 mM theophylline to AP labeled aptazymes. (A) Increasing AP fluorescence signal of the aptamer (C27′→AP) mutant (top) and decreasing AP fluorescence signal of the catalytic core (A2.1→AP) mutant (bottom). (B) Theophylline concentration dependence of the theophylline binding rate constants derived from single-exponential fits to AP fluorescence time cses like those in panels A and B, yielding the indicated theophylline half-titration points, $Th_{1/2}$, for the aptamer and core mutants (Materials and Methods).

saturating (rate-limiting) conditions ($Th_{1/2}$ = 1.6 ± 0.8 mM, Fig. 2.7B)[2]. That is, the conformational change induced in the catalytic core near the cleavage site of the aptazyme by distal theophylline binding (Fig. 2.1A) is ~800-fold slower than the structural rearrangement in the theophylline binding aptamer itself. These observations suggest that signal transduction by conformational rearrangements from the aptamer domain through the communication module to the catalytic core is relatively slow, limiting the maximal cleavage output of the aptazyme. This upper limit further reduces the attainable dynamic range (ratio of maximal activity over background cleavage) of the aptazyme as a biosensor for theophylline.

2.3 Discussion

Here, single molecule FRET has been used to observe in real-time the global dynamics between Stems I and III of a drug-sensing hammerhead-derived aptazyme, indicating that the active conformation docks into a Y-shape. Ligand binding to the aptamer domain incorporated into Stem II favors a long-lived, catalytically active docked state, revealing the ligand-induced conformational activation mechanism of the aptazyme. This same docked state is also visited, albeit less frequently, in the absence of ligand, accounting for both the observed background activity (leakage) and part of the limited dynamic range of sensing. Together with the observation from specific 2-aminopurine labeling that allosteric signal transduction from the aptamer to the catalytic

[2] Correctly weighted error analysis shows a $Th_{1/2}$ value of 1.47±0.47 mM (Appendix One). This value does not change the conclusions of this chapter.

core is slow, these results suggest that the designs of this and other biosensing aptazymes can be further improved to match the performance of nature's only known aptazyme, the *glmS* ribozyme.

Despite extensive studies since the mid 1980's, details on the structural dynamics of the hammerhead ribozyme have remained elusive until now. Data presented here reveal that at equilibrium an aptazyme designed based on the hammerhead ribozyme fluctuates between several inactive conformations and the active, expectedly Y-shaped [26, 27] conformer. As evident from Figures 2.3-2.6, smFRET probing specifically suggests that: (i) A pair of H (high smFRET) states as well as a pair of L (low smFRET) states exist that are distinguished by their dwell times (Figs. 2.3A, 2.3B, 2.4A, 2.4B), but not their smFRET values as the histograms of Fig. 2.3C show close-to-ideal bimodal distributions with no evidence for additional states of distinct smFRET value; (ii) it is the particularly long-lived L state with distal Stem I and III whose prevalence correlates with catalytic activity (compare Figs. 2.2B and 2.6); and (iii) each single molecule can freely interchange between all f states, that is, the heterogeneity described here exists between pairs of structurally, yet not kinetically similar H and L states that dynamically exchange for any given molecule without memory of any preceding behavior (Fig. 2.5). This behavior is distinct from that described for the similarly small hairpin ribozyme, where distinct sub-populations of molecules exist that do not interconvert on any experimentally accessible timescale [34, 41, 45-47].

AP fluorescence probing highlights the fact that the aptamer domain is relatively unperturbed by incorporation into the aptazyme, as solvent exposure of nucleotide position 27' upon theophylline binding still occurs and is only 7-fold slower (Fig. 2.7B)

than without appended ribozyme [43]. Yet, coupling of the ligand-induced, localized conformational changes of the aptamer domain through the communication domain into the catalytic core of the hammerhead ribozyme motif (Fig. 2.1A) is ~800-fold slower (Fig. 2.7B), suggesting that a significant energetic barrier of ~4 kcal/mol, equivalent to the breaking of several hydrogen bonds, has to be overcome in this long-range structural communication. Such a significant barrier is consistent with the "slippage" mechanism proposed to be responsible for properly aligning base pair in the communication domain upon ligand binding [16, 18, 22]. These findings suggest that the design or in vitro selection of aptazymes with such a barrier has been inherently limiting the achievable dynamic range and response time of a biosensor.

Figure 2.8 schematically shows the simplest kinetic reaction pathway consistent with all the data obtained. Each molecule switches between two high and two low smFRET states that are distinguished by their dwell time kinetics. The low smFRET state characterized by a very slow conversion back to a high smFRET state can also spontaneously, but very slowly, reach a catalytically active conformation that generates cleavage product in the absence of theophylline ligand ("leakage pathway"). Rapid local binding of theophylline to the aptamer domain stabilizes a conformation that slowly rearranges the catalytic core into the active state to generate product ("biosensing pathway"). Here, the ratio of the overall rate constants through the biosensing and leakage pathways have been determined as <100-fold (Fig. 2.2B), which significantly limits the inherent dynamic range of the engineered biosensing aptazyme.

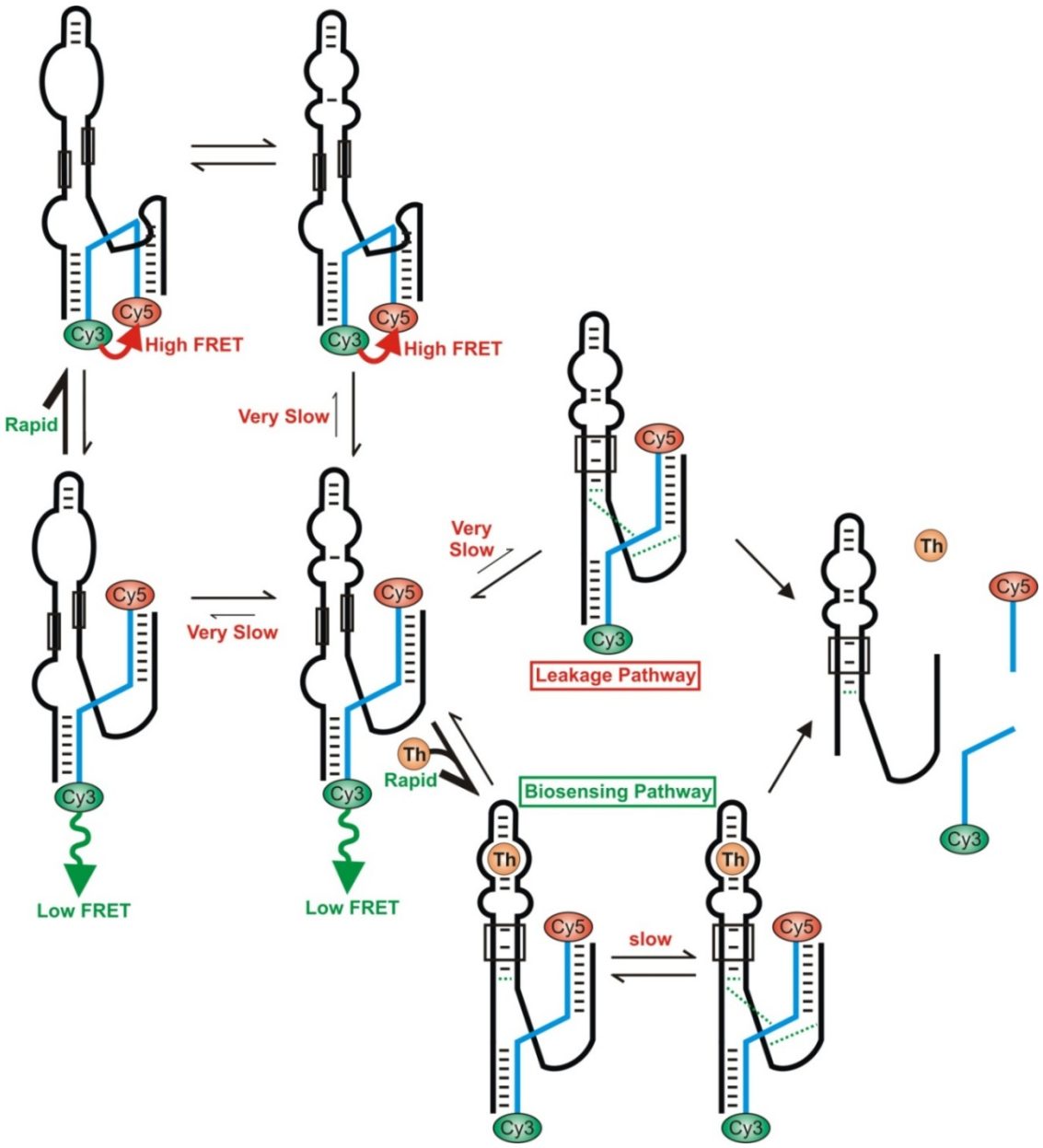


Figure 2.8 Proposed simplest kinetic pathway of the theophylline aptazyme consistent with cleavage activity, smFRET, and AP fluorescence data.

Engineered aptazymes based on the hammerhead ribozyme were first developed before the discovery of the natural *glmS* aptazyme [16, 22]. Nature chose the GlcN6P ligand to act as a chemical participant of the *glmS* aptazyme reaction [10], contributing general acid (or possibly base) capacity so that ligand binding and catalytic activation can quickly and efficiently occur without the need for a conformational change [12-15]. By contrast, the hammerhead ribozyme is thought to utilize two of its own guanines (and possibly bound water molecules) for acid-base catalysis [48-50], making design of a ligand-induced, activating conformational change necessary to afford ligand dependence. We show here that single molecules of the prototypical theophylline aptazyme exhibit f global conformational states that are readily distinguished by the extent of smFRET between Stems I and III as well as their interconversion kinetics. We identify two drawbacks that limit the dynamic range of the aptazyme as a biosensor component: (i) Leakage, i.e., significant residual activity in the ligand-free state; and (ii) slow allostery, i.e., slow transduction of the ligand binding signal from the aptamer to the catalytic core domain.

While similar studies of other engineered aptazymes will need to follow to show the generality of findings on the prototypical theophylline aptazyme, these results suggest three clear strategies to improve aptazyme performance: (i) Design aptazymes with a similar mode of action as the *glmS* ribozyme, i.e., where the ligand is used as a chemical cofactor. This strategy will of course be limited to certain types of ligands with chemically active functional groups. (ii) Prevent the leakage reaction by more stringent selection against aptazymes that cleave in the absence of ligand without a loss in ligand-induced activity. (iii) Facilitate (accelerate) signal transduction from the aptamer to the

catalytic core domain by lowering the associated energetic barrier. This latter strategy may be implemented, for example, by designing aptamer domains that more fully envelop the ligand for a maximum gain in binding free energy, as observed in naturally occurring riboswitches [5, 7, 51, 52]. This enhanced binding free energy could then be harvested to more fully, and possibly more rapidly, rearrange the catalytic core into an active conformation. The future (combinatorial) exploration of such design strategies promises to overcome the limitations of currently available aptazymes for use as biosensor components and artificial genetic switches.

2.4 Materials and Methods

RNA Preparation

The aptazyme strand (Fig. 2.1A) was in vitro transcribed using T7 RNA polymerase from a single-stranded DNA template with double-stranded promoter purchased from Invitrogen and purified by denaturing polyacrylamide gel electrophoresis (D-PAGE) [24]. For smFRET the transcribed aptazyme strand was 3′-biotinylated by oxidization with periodate and reaction with biotin hydrazide as described [53]. The chemically synthesized substrate and two AP modified aptazyme strands were purchased from the Keck Foundation Biotechnology Research Laboratory at the Yale University School of Medicine, deprotected as recommended by the manufacturer, and purified by D-PAGE and C_8 reverse phase HPLC as described [54]. In all single-molecule FRET and

AP-fluorescence based structure probing experiments a non-cleavable substrate analog with a 2′-O-methyl modification at the cleavage site (C17) was employed. For all FRET experiments the substrate was labeled with Cy3 and Cy5 on the 5′- and 3′-ends during and after synthesis (the latter by coupling the dye succinimidyl ester to a C7 3′-Amino-Modifier), respectively, and repurified by C_8 reverse phase HPLC as described [54].

Steady-State and Stopped-Flow Ensemble Fluorescence Assays

Steady-state ensemble cleavage assays monitored by FRET were performed on an Aminco-Bowman 2 spectrofluorometer at 25ºC in standard buffer (50 mM Tris-HCl, pH 7.5, 25 mM DTT) with either 10 mM or 100 mM $MgCl_2$ as indicated. 200 nM aptazyme was added to 50 nM cleavable substrate (both final concentrations) in the cuvette, and reactions were initiated by addition of theophylline to varying final concentrations as previously described [28]. Theophylline binding kinetics were monitored by AP fluorescence of an annealed complex of 100 nM AP labeled aptazyme mutant and 400 nM non-cleavable substrate analog (without fluorophores) at 25˚C in standard buffer with 100 mM $MgCl_2$ in an SF-2001 stopped-flow spectrofluorometer (KinTek Corp., Austin, TX) fitted with a 75 W Xe-arc lamp. AP was excited at 315 nm (4 nm slit width) and its fluorescence was monitored using a 340 nm long-pass filter. Time traces of acceptor/donor signal ratio and of AP fluorescence were analyzed by least-squares fitting single exponentials to the data using Microcal Origin 7.0 software. The obtained rate constants were plotted against the theophylline concentration (*[8]*) and fit with a hyperbolic, non-cooperative binding equation of the form:

$$k_{obs} = k_{\max} \frac{[Th]}{[Th] + Th_{1/2}}$$

where k_{max} is the cleavage rate at saturating theophylline concentration and $Th_{1/2}$ is the theophylline half-titration point.

smFRET Assays

The 3′-biotinylated aptazyme strand (500 nM final concentration) was annealed with the non-cleavable Cy3/Cy5 doubly labeled substrate strand (1 μM) in standard buffer at 70˚C for 2 min and cooled on ice for 15 min. The assembled RNA complex was diluted to 50 pM in standard buffer with 100 mM $MgCl_2$, supplemented with an oxygen scavenger system to prolong the lifetime of the fluorophores as described [41], and bound at 25˚C to a streptavidin coated quartz slide surface within a microfluidic channel as described [55] until a density of ~0.1 molecules per μm^2 was reached. A home-built prism-based total internal reflection fluorescence (TIRF) microscope with intensified CCD camera (I-PentaMAX, Roper Scientific Inc.) was used to observe the donor (I_D) and acceptor (I_A) fluorescence signals and the resulting smFRET ratio of $I_A/(I_A+I_D)$ at 100-ms time resolution, essentially as described [35, 36, 56]. The dwell times of all L and H states of each trajectory were calculated and cumulative dwell time histograms plotted and fit (using Microcal Origin 7.0) with double exponentials to obtain transition rate constants k_{trans} after correction for photobleaching and observation time limitations as described [34, 41]. Using the amplitudes of the double-exponential fits, fractions of each observed L-to-H and H-to-L rate constant were calculated after also correcting for

photobleaching and observation time limitations [41]. The theophylline concentration (*[8]*) dependence of the fast and slow L-to-H fractions *f* (which add up to unity) were globally fit with a hyperbolic, non-cooperative binding equation of the form:

$$f = f_0 + (-1)^a f_{\max} \frac{[Th]}{[Th] + Th_{1/2,\, global}}$$

where $a = 0$ leads to an increasing and $a = 1$ to a decreasing curve, f_0 and f_{max} characterize the start and end points of the two titrations, respectively, and $Th_{1/2,\, global}$ is the common theophylline half-titration point.

Aggregate histograms of the counts of video frames with a specific smFRET value from each one of 45 representative molecules were plotted and fit with double gaussians to calculate the percentage time spent in the L and H states. We generated a continuous probability density plot of the dwell time data by weighing each event by the average time between its nearest neighbor events. This procedure has negligible impact on the short dwell times that are characterized by multiple events in each time bin, yet represents an improved statistical estimate of the probability density of long dwell times where the event density is low [57]. Dwell times of adjacent L and H states were also plotted against one another to detect correlations between them. Figure 2.6 was plotted using Igor Pro software, whereas all other plots were generated in Microcal Origin 7.0 software.

2.5 Acknowledgments

This work was supported by grants from the NIH (GM62357) and NASA (NNA04CD01G) to NGW. The authors thank Dr. David Rueda for setting up the single molecule microscope and for help with initial experiments, Drs. Carol Fierke and John Hsieh for the use of and training on the stopped-flow spectrofluorometer and for stimulating discussions of the kinetic pathway and all Walter group members for helpful suggestions.

2.6 References

1. Mandal, M. and R.R. Breaker, *Gene regulation by riboswitches.* Nat Rev Mol Cell Biol, 2004. **5**(6): p. 451-63.

2. Breaker, R.R., *Natural and engineered nucleic acids as tools to explore biology.* Nature, 2004. **432**(7019): p. 838-45.

3. Winkler, W.C. and R.R. Breaker, *Regulation of bacterial gene expression by riboswitches.* Annu Rev Microbiol, 2005. **59**: p. 487-517.

4. Coppins, R.L., K.B. Hall, and E.A. Groisman, *The intricate world of riboswitches.* Curr. Opin. Microbiol., 2007. **10**(2): p. 176-81.

5. Serganov, A. and D.J. Patel, *Ribozymes, riboswitches and beyond: regulation of gene expression without proteins.* Nat. Rev. Genet., 2007. **8**(10): p. 776-90.

6. Wakeman, C.A., W.C. Winkler, and C.E. Dann, 3rd, *Structural features of metabolite-sensing riboswitches.* Trends Biochem. Sci., 2007. **32**(9): p. 415-24.

7. Montange, R.K. and R.T. Batey, *Riboswitches: emerging themes in RNA structure and function.* Annu. Rev. Biophys., 2008. **37**: p. 117-33.

8. Al-Hashimi, H.M. and N.G. Walter, *RNA dynamics: it is about time.* Curr. Opin. Struct. Biol., 2008. **18**(3): p. 321-329.

9. Winkler, W.C., A. Nahvi, A. Roth, J.A. Collins, and R.R. Breaker, *Control of gene expression by a natural metabolite-responsive ribozyme.* Nature, 2004. **428**(6980): p. 281-6.

10. McCarthy, T.J., M.A. Plog, S.A. Floy, J.A. Jansen, J.K. Soukup, and G.A. Soukup, *Ligand requirements for glmS ribozyme self-cleavage.* Chem. Biol., 2005. **12**(11): p. 1221-6.

11. Collins, J.A., I. Irnov, S. Baker, and W.C. Winkler, *Mechanism of mRNA destabilization by the glmS ribozyme.* Genes Dev., 2007. **21**(24): p. 3356-68.

12. Hampel, K.J. and M.M. Tinsley, *Evidence for preorganization of the glmS ribozyme ligand binding pocket.* Biochemistry, 2006. **45**(25): p. 7861-71.

13. Klein, D.J. and A.R. Ferre-D'Amare, *Structural basis of glmS ribozyme activation by glucosamine-6-phosphate.* Science, 2006. **313**(5794): p. 1752-6.

14. Cochrane, J.C., S.V. Lipchock, and S.A. Strobel, *Structural investigation of the GlmS ribozyme bound to Its catalytic cofactor.* Chem. Biol., 2007. **14**(1): p. 97-105.

15. Tinsley, R.A., J.R. Furchak, and N.G. Walter, *Trans-acting glmS catalytic riboswitch: locked and loaded.* RNA, 2007. **13**(4): p. 468-77.

16. Breaker, R.R., *Engineered allosteric ribozymes as biosensor components.* Curr. Opin. Biotechnol., 2002. **13**(1): p. 31-9.

17. Silverman, S.K., *Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA.* RNA, 2003. **9**(4): p. 377-83.

18. Hall, B., J.R. Hesselberth, and A.D. Ellington, *Computational selection of nucleic acid biosensors via a slip structure model.* Biosens. Bioelectron., 2007. **22**(9-10): p. 1939-47.

19. Wieland, M. and J.S. Hartig, *Improved aptazyme design and in vivo screening enable riboswitching in bacteria.* Angew. Chem. Int. Ed. Engl., 2008. **47**(14): p. 2604-7.

20. Blount, K., I. Puskarz, R. Penchovsky, and R. Breaker, *Development and application of a high-throughput assay for glmS riboswitch activators.* RNA Biol, 2006. **3**(2): p. 77-81.

21. Mayer, G. and M. Famulok, *High-throughput-compatible assay for glmS riboswitch metabolite dependence.* Chembiochem, 2006. **7**(4): p. 602-4.

22. Soukup, G.A. and R.R. Breaker, *Engineering precision RNA molecular switches.* Proc. Natl. Acad. Sci. USA, 1999. **96**(7): p. 3584-9.

23. Sekella, P.T., D. Rueda, and N.G. Walter, *A biosensor for theophylline based on fluorescence detection of ligand-induced hammerhead ribozyme cleavage.* RNA, 2002. **8**(10): p. 1242-52.

24. Milligan, J.F. and O.C. Uhlenbeck, *Synthesis of small RNAs using T7 RNA polymerase.* Methods Enzymol., 1989. **180**: p. 51-62.

25. Salehi-Ashtiani, K. and J.W. Szostak, *In vitro evolution suggests multiple origins for the hammerhead ribozyme.* Nature, 2001. **414**(6859): p. 82-4.

26. Blount, K.F. and O.C. Uhlenbeck, *The structure-function dilemma of the hammerhead ribozyme.* Annu Rev Biophys Biomol Struct, 2005. **34**: p. 415-40.

27. Nelson, J.A. and O.C. Uhlenbeck, *Hammerhead redux: does the new structure fit the old biochemical data?* RNA, 2008. **14**(4): p. 605-15.

28. Rueda, D. and N.G. Walter, *Fluorescent energy transfer readout of an aptazyme-based biosensor.* Methods Mol Biol, 2006. **335**: p. 289-310.

29. Hendeles, L. and M. Weinberger, *Theophylline. A "state of the art" review.* Pharmacotherapy, 1983. **3**(1): p. 2-44.

30. Miyazawa, Y., L.P. Starkey, A. Forrest, J.J. Schentag, H. Kamimura, H. Swarz, and Y. Ito, *Effects of the concomitant administration of tamsulosin (0.8 mg/day) on the pharmacokinetic and safety profile of theophylline (5 mg/kg): a placebo-controlled evaluation.* J. Int. Med. Res., 2002. **30**(1): p. 34-43.

31. Jenison, R.D., S.C. Gill, A. Pardi, and B. Polisky, *High-resolution molecular discrimination by RNA.* Science, 1994. **263**(5152): p. 1425-9.

32. Zimmermann, G.R., C.L. Wick, T.P. Shields, R.D. Jenison, and A. Pardi, *Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer.* Rna, 2000. **6**(5): p. 659-67.

33. Mairal, T., V.C. Ozalp, P. Lozano Sanchez, M. Mir, I. Katakis, and C.K. O'Sullivan, *Aptamers: molecular tools for analytical applications.* Anal. Bioanal. Chem., 2008. **390**(4): p. 989-1007.

34. Zhuang, X., H. Kim, M.J. Pereira, H.P. Babcock, N.G. Walter, and S. Chu, *Correlating structural dynamics and function in single ribozyme molecules.* Science, 2002. **296**(5572): p. 1473-6.

35. Pereira, M.J., E.N. Nikolova, S.L. Hiley, D. Jaikaran, R.A. Collins, and N.G. Walter, *Single VS Ribozyme Molecules Reveal Dynamic and Hierarchical Folding Toward Catalysis.* J Mol Biol, 2008. **382**(2): p. 496-509.

36. Walter, N.G., C.Y. Huang, A.J. Manzo, and M.A. Sobhy, *Do-it-yourself guide: how to use the modern single-molecule toolkit.* Nat. Methods, 2008. **5**(6): p. 475-89.

37. Bassi, G.S., N.E. Mollegaard, A.I. Murchie, and D.M. Lilley, *RNA folding and misfolding of the hammerhead ribozyme.* Biochemistry, 1999. **38**(11): p. 3345-54.

38. Rueda, D., K. Wick, S.E. McDowell, and N.G. Walter, *Diffusely bound Mg2+ ions slightly reorient stems I and II of the hammerhead ribozyme to increase the probability of formation of the catalytic core.* Biochemistry, 2003. **42**(33): p. 9924-36.

39. Penedo, J.C., T.J. Wilson, S.D. Jayasena, A. Khvorova, and D.M. Lilley, *Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements.* RNA, 2004. **10**(5): p. 880-8.

40. Sabanayagam, C.R., J.S. Eid, and A. Meller, *Using fluorescence resonance energy transfer to measure distances along individual DNA molecules: corrections due to nonideal transfer.* J. Chem. Phys., 2005. **122**(6): p. 061103.

41. Rueda, D., G. Bokinsky, M.M. Rhodes, M.J. Rust, X. Zhuang, and N.G. Walter, *Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance.* Proc. Natl. Acad. Sci. USA, 2004. **101**(27): p. 10066-71.

42. Gondert, M.E., R.A. Tinsley, D. Rueda, and N.G. Walter, *The catalytic core structure of the trans-acting HDV ribozyme is subtly influenced by sequence variation outside the core.* Biochemistry, 2006. **45**: p. 7563-7573.

43. Jucker, F.M., R.M. Phillips, S.A. McCallum, and A. Pardi, *Role of a heterogeneous free state in the formation of a specific RNA-theophylline complex.* Biochemistry, 2003. **42**(9): p. 2560-7.

44. Menger, M., T. Tuschl, F. Eckstein, and D. Porschke, *Mg(2+)-dependent conformational changes in the hammerhead ribozyme.* Biochemistry, 1996. **35**(47): p. 14710-6.

45. Bokinsky, G., D. Rueda, V.K. Misra, M.M. Rhodes, A. Gordus, H.P. Babcock, N.G. Walter, and X. Zhuang, *Single-molecule transition-state analysis of RNA folding.* Proc. Natl. Acad. Sci. USA, 2003. **100**(16): p. 9302-7.

46. Okumus, B., T.J. Wilson, D.M. Lilley, and T. Ha, *Vesicle Encapsulation Studies Reveal that Single Molecule Ribozyme Heterogeneities Are Intrinsic.* Biophys. J., 2004. **87**(4): p. 2798-806.

47. Ditzler, M.A., D. Rueda, J. Mo, K. Håkansson, and N.G. Walter, *A deeply furrowed, non-ergodic folding landscape is intrinsic to functional RNA.* Nucleic Acids Res., 2008: p. in press.

48. Han, J. and J.M. Burke, *Model for general acid-base catalysis by the hammerhead ribozyme: pH-activity relationships of G8 and G12 variants at the putative active site.* Biochemistry, 2005. **44**(21): p. 7864-70.

49. Martick, M. and W.G. Scott, *Tertiary contacts distant from the active site prime a ribozyme for catalysis.* Cell, 2006. **126**(2): p. 309-20.

50. Walter, N.G., *Ribozyme catalysis revisited: is water involved?* Mol. Cell, 2007. **28**(6): p. 923-929.

51. Batey, R.T., S.D. Gilbert, and R.K. Montange, *Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine.* Nature, 2004. **432**(7015): p. 411-5.

52. Serganov, A., Y.R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A.T. Phan, C. Hobartner, R. Micura, R.R. Breaker, and D.J. Patel, *Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs.* Chem Biol, 2004. **11**(12): p. 1729-41.

53. Newby Lambert, M., E. Vocker, S. Blumberg, S. Redemann, A. Gajraj, J.C. Meiners, and N.G. Walter, *Mg2+-induced compaction of single RNA molecules monitored by tethered particle microscopy.* Biophys. J., 2006. **90**(10): p. 3672-85.

54. Walter, N.G., *Probing RNA structural dynamics and function by fluorescence resonance energy transfer (FRET).* Curr. Protoc. Nucleic Acid Chem., 2002. **11.10**: p. 11.10.1-11.10.23.

55. Ha, T., *Single-molecule fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 78-86.

56. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat. Methods, 2008. **5**(6): p. 507-16.

57. Kuno, M., D.P. Fromm, H.F. Hamann, A. Gallagher, and D.J. Nesbitt, *"On"/"off" fluorescence intermittency of single semiconductor quantum dots.* J. Chem. Phys., 2001. **115**(2): p. 1028-1040.

CHAPTER THREE

DNA AND RNA LIGASE MEDIATED METHODS FOR THE ASSEMBLY OF THE HEPATITIS DELTA VIRUS (HDV) RIBOZYME FOR SINGLE MOLECULE FRET STUDIES[3]

3.1 Introduction

Chemical synthesis of RNA molecules beyond ~80 nucleotides has been shown to be prohibitively inefficient due to exceedingly low yields. Consequently, joining several shorter pieces of RNA by various ligation methods to generate large RNAs is an attractive alternative route to studying lengthy RNA molecules [1, 2]. Furthermore, the addition of site-specific modifications to nucleic acids has been a powerful tool for biophysical and biochemical studies, and ligation methods for synthetic RNA strands facilitate incorporation of a variety of unnatural nucleotides into RNA. One established ligation approach uses DNA ligase to mediate successful ligation of two pieces of RNA (Hereafter referred to as Dligation) [3]. Alternatively, RNA ligase has been used as an effective enzyme for RNA ligation (Hereafter referred to as Rligation). Both of these techniques require unique parameters to be met for the reaction to occur. Moreover, they

[3]Transcription assays and Dligation performed by Chamaree de Silva and Robin Johnson, Rligation performed by Chamaree de Silva and Dr. Shiamalee Perumal, and smFRET experiments carried out by Chamaree de Silva.

have distinctive advantages and drawbacks that require investigation and optimization unique to each RNA system of consideration.

DNA ligase mediated RNA ligation

DNA ligase can be used to catalyze formation of a natural 3´-5´ phosphodiester linkage between the 5´ phosphate of an RNA (called the donor RNA) and the 3´ hydroxyl group of another RNA (called the acceptor) [3, 4]. Therefore, the donor RNA is required to carry a monophosphate at the 5´ end instead of the tri-phosphate obtained by standard transcription from a DNA template. A suitable DNA ligase can be isolated from bacteriophage T4 (widely called T4 DNA ligase), and is ATP dependent. T4 DNA ligase, which normally joins nicks in a double-stranded DNA, requires the assistance of a specially designed DNA strand (called the DNA splint) to hold the two RNA ends in apposition in order to form the chemical bond between them during ligation [1, 3-5]. To accomplish this task, the DNA splint is required to hybridize to the acceptor and donor strands, and hence is required to be complementary to segments of both RNAs. T4 DNA ligase recognizes this nicked, double stranded hybrid duplex, and performs the ligation generally within a few hours at room temperature when provided in near-stoichiometric amounts. Furthermore, T4 DNA ligase is not heavily dependent on the exact sequence at the ligation point besides the required hybridization between the RNAs and the DNA splint on both sides of the junction. Nonetheless, these reactions require large amounts of T4 DNA ligase and frequently produce low yields of ligation material [4, 5].

RNA ligase mediated RNA ligation

An alternate technique to Dligation is to use an RNA ligase as the enzyme to catalyze a ligation reaction between two RNAs [6-8]. Donor and acceptor strands with identical requirements as for Dligation are necessary for the reaction to occur. However, RNA ligase can perform without the assistance of a DNA splint in a non-templated fashion, as it requires single stranded RNA at the ligation position [8]. Although single stranded regions are a prerequisite for RNA ligase activity, a DNA splint can be used to hold the two RNAs together, provided that a few nucleotides at and near the ligation position are not hybridized. These substrate requirements are due to the fact that the natural substrate for T4 RNA ligase, the main enzyme used for the reaction, are nicked RNA loops. Additionally, however, T4 RNA ligase indiscriminately joins any RNAs with the required 5´ and 3´ termini, which is useful for techniques such as labeling with radioactive pCp, but also forms unintended circular RNAs and oligomers during ligation of longer RNAs [4, 6, 7, 9]. Studies of Rligation have shown that the reaction can occur as fast as within 5 minutes, which is an advantage compared to that of Dligation [4]. Rligation can also be used to join a very small RNA strand (few nucleotides in length) to a larger RNA. This method is useful in situations where the shorter RNA would carry a chemical modification necessary for the study of the ligated RNA, such as the attachment of a fluorophore or radioactive label. By contrast, Dligation is not useful in these circumstances where hybridization of a DNA splint to a very short RNA strand is not thermodynamically favorable.

For smFRET studies of the genomic HDV ribozyme, an identical sequence to the construct that has been previously characterized using cleavage assays and Tb^{3+} footprinting was employed [10]. However, in order to adapt this construct for single molecule studies it was necessary to make several modifications. These include addition of a tether (capture strand) to immobilize the molecule on a quartz slide for smFRET studies, and addition of the donor Cy3 and acceptor Cy5 fluorophores for FRET. Thus, the synthesis, purification, ligation and assembly of several oligonucleotides were required. The HDV ribozyme with an extended 5′ end is 111 nucleotides in length, and Dligation and Rligation methods were used and compared in order to assemble the complete molecule.

The HDV ribozyme and its U-turn

The hepatitis delta virus (HDV) ribozyme is the only known ribozyme found in a human pathogen, the hepatitis delta virus, which acts as a satellite virus for the hepatitis B virus. The HDV virion contains a single stranded circular RNA genome that replicates through a double-rolling circle mechanism. This process generates multimeric genomic and antigenomic HDV RNAs of ~1,700 nucleotides length. The embedded minimal genomic and antigenomic ribozymes are 85 nucleotides in length and have very similar sequences. Site-specific self-cleavage entails nucleophilic attack by the adjacent, base-activated 2′-oxygen on the cleavage site phosphorus to substitute the acid-activated 5′-oxygen, generating 5′- and 3′- cleavage products with 2′,3′- cyclic phosphate and 5’-hydroxyl termini, respectively [10]. It has been shown that the nucleotide C75 in the

genomic ribozyme is crucial for catalytic activity [11, 12]. However, whether it acts as a general acid or as the general base remains elusive. From enzymologic studies it was suggested that C75 acts as a general acid to protonate the leaving group, which requires C75 to be N3-protonated just before cleavage. Molecular Dynamics (MD) simulations found that an unprotonated C75 shows weak hydrogen bonding between its C75(N3) and U1(O2´) as required by C75 to act as a general base [10]. Taken together even with the wealth of information shown by crystal structures and biochemistry studies, the exact mechanism of catalysis of the HDV ribozyme remains highly controversial and elusive. The recent discovery of the RNA U-turn motif at the heart of the HDV ribozyme highlights even more the intricacies of its structure and function. A previously noticed functional role of the 5´-substrate sequence is explained by the required formation of the U-turn motif to bend the cleavage site into the correct conformation [10] (Figure 3.1).

Although bulk measurements have been performed on the HDV ribozyme, observations at the single molecule level are expected to reveal for the first time details of the ribozyme's global dynamics. Single molecule FRET is an excellent tool to observe such structural dynamics and to gather kinetic and folding information of ribozymes, as shown previously [2, 13-15]. By carefully placing fluorophores at particular positions of the ribozyme, global conformational changes can be directly observed. In addition, future NMR studies and additional MD simulations of the HDV ribozyme will generate complementary information on structural and functional dynamics in a multi-timescale fashion to gain a deeper understanding of the functional properties of the U-turn motif, with relevance for a multitude of biologically relevant RNAs.

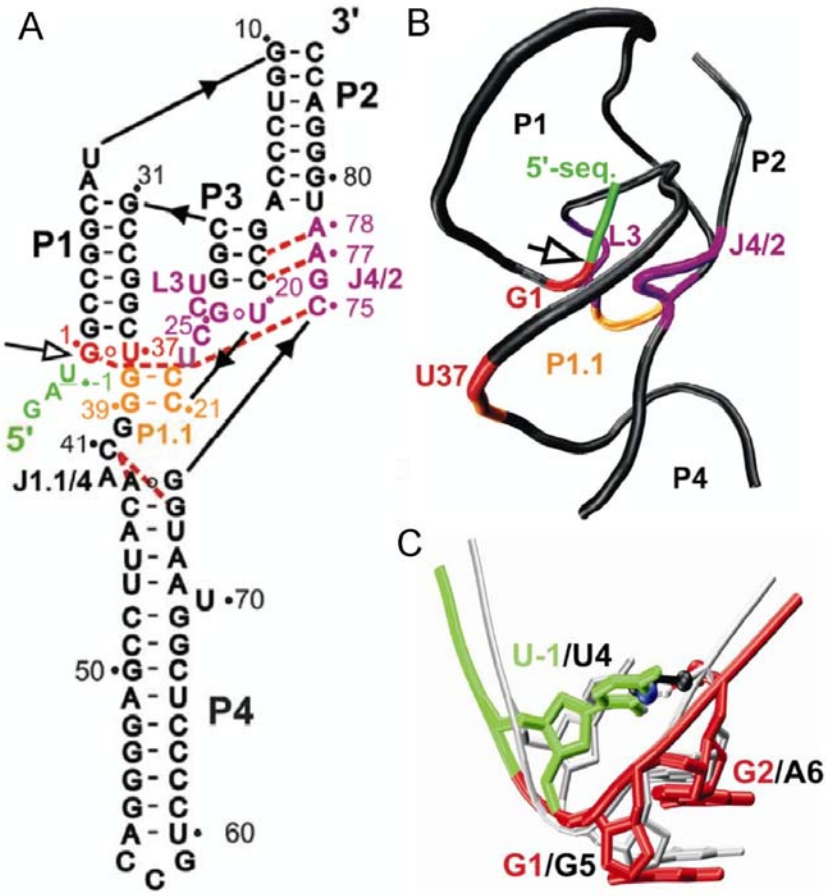


Figure 3.1 Sequence and structure of the cis-acting genomic HDV ribozyme and its U-turn. (A) Secondary structure with the 5´-sequence (green) immediately upstream of the cleavage site (open arrow). Nucleotides in color correspond to important structural elements in the catalytic core. Red dashed lines are functionally relevant tertiary interactions. (B) Backbone ribbon representation of the precursor crystal structure, color-coded as in A. (C) Overlay of the U-turn from the initial HDV crystal structure-based model from MD simulations with the crystal structure of the hammerhead ribozyme (silver). Reproduced with permission from Sefcikova, J., et al., Nucleic Acids Res, 2007, 35(6): p 1933-46. © The Author(s).

The HDV ribozyme was synthesized in three different segments to introduce the necessary modifications before ligating them together into one RNA molecule (Figure 3.2). The extended 5′ sequence is designed to be hybridized with the capture strand (termed HCAPT) immediately prior to single molecule experiments. This capture strand can be immobilized on a quartz surface with the assistance of a 3′ biotin, and is Cy3 labeled at the 5′ end to act as the donor for FRET studies (Figure 3.3). The 5′ segment of the ribozyme (C-26 to G10) is comprised of 36 nucleotides (termed HD36) and is chemically synthesized with a 2′ O-methyl modification at the cleavage site (U-1) in order to obtain a non-cleavable ribozyme. The 73 nucleotide segment between G11 and A83 (termed H73) was originally transcribed from a DNA template and ligated to HD36 by Dligation. The remaining two-nucleotide segment of the very 3′ end of the molecule (called H2) was then synthesized with an amino linker on the 3′ end for post-synthetic Cy5 labeling. This short segment was the ligated to the rest of the molecule using Rligation. However, this yielded very little product, prompting chemical synthesis of the 75 nucleotide segment between G11 and C85 (termed H75) to be then ligated to HD36 via Dligation. As a result, the 111 nucleotide long HDV ribozyme was assembled and single molecule FRET studies could be performed. The assembly required numerous optimizations of transcription and ligation protocols. The two ligation processes are explained in detail below.

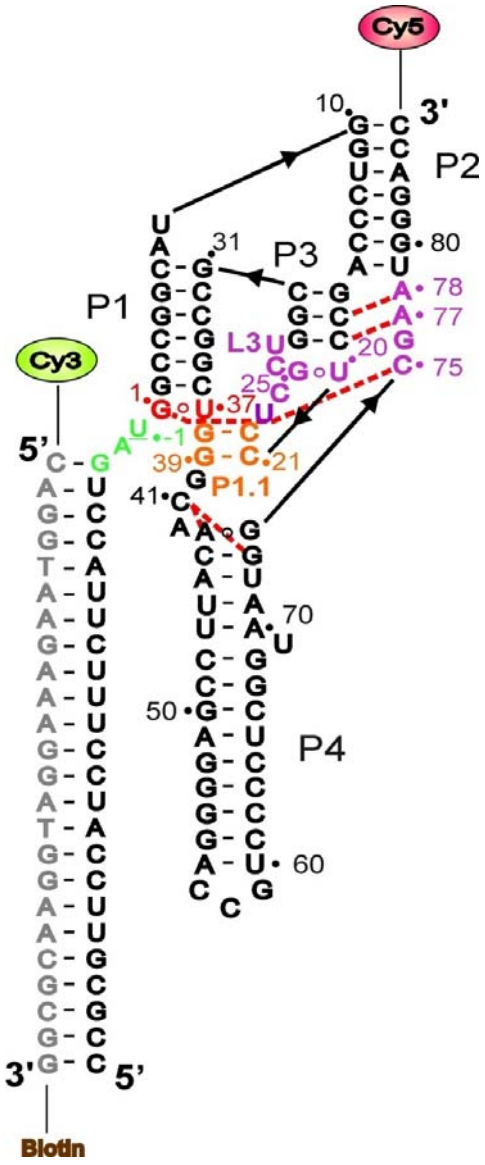


Figure 3.2 The genomic HDV ribozyme used for single molecule FRET experiments. The capture strand (grey) has a biotin on the 3′ end to immobilize the molecule to the surface. It is labeled with Cy3 on the 5′ end to act as the donor of the FRET pair. The ribozyme strand is labeled with the acceptor Cy5 at the 5′ end. Cleavage is blocked with a 2′ O-methyl modification at the 5′ sequence.

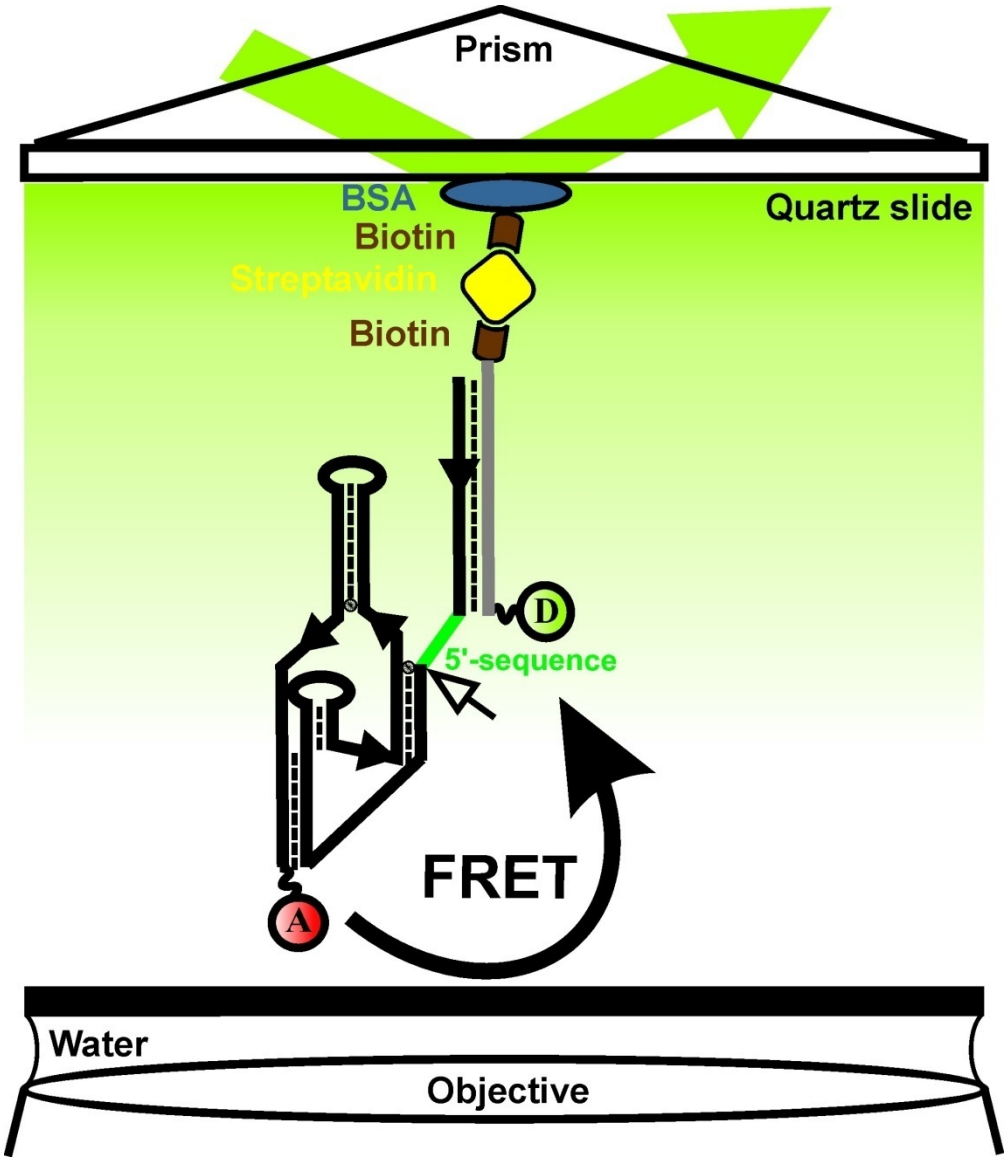


Figure 3.3 The HDV ribozyme construct for TIRFM based single-molecule FRET experiments.

3.2 Results

Optimization of a transcription reaction

The segment H73 could have been chemically synthesized to contain a 5´ monophosphate, but due to its length, in vitro transcription was easily attainable and was thought to be more economical. In order to transcribe H73, two DNA templates were required. One DNA template (termed H93) was partially complementary to the desired H73 and has 20 extra nucleotides that were complementary to an additional DNA promoter sequence (termed PROMOTOR). Once H93 and PROMOTOR strands were hybridized, RNA polymerase can attach one nucleotide at a time to produce H73 after binding to the TATA box on H93. Generally, a transcription reaction is completed in a few hours at 37˚C. Although many protocols have been developed for RNA transcription, each RNA system in itself is different and needs optimization to produce an optimal RNA yield.

With the use of H93 and PROMOTOR, variations of an already established transcription protocol were performed [16] (Figure 3.4). Several reactions were run under standard buffer conditions (Materials and Methods), yet concentrations of many reagents were modified to find optimal conditions. These include increased cation and pyrophosphatase concentrations, variations of template and promoter concentrations as well as increased RNA polymerase concentrations (Materials and Methods). In addition, GMP was introduced to the reaction in order to include a monophosphate at the 5´ end of

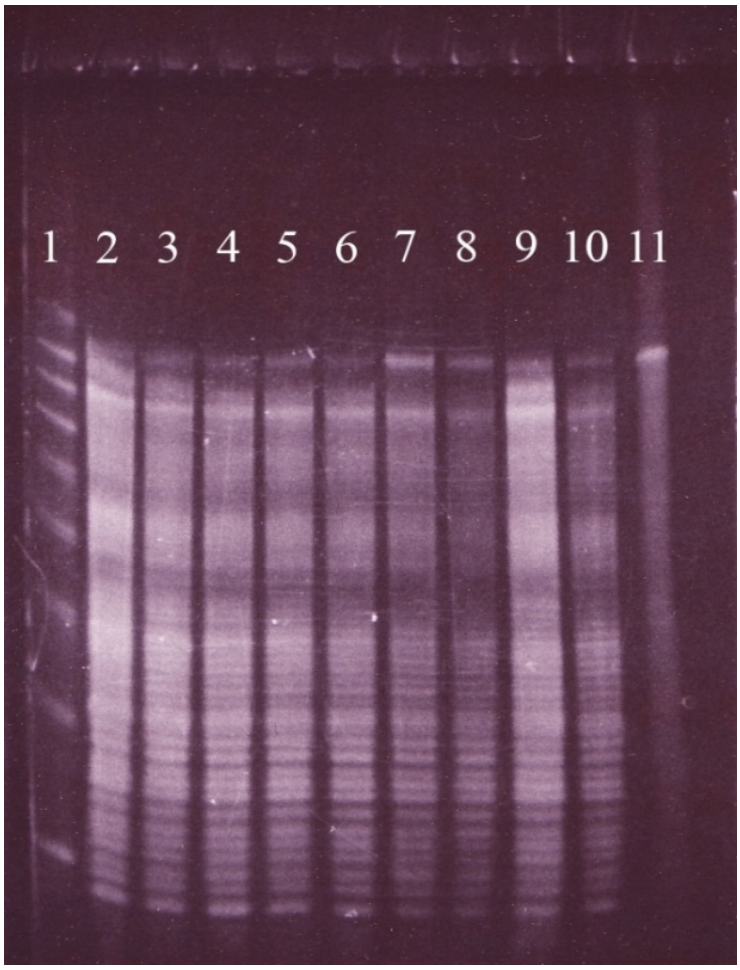


Figure 3.4 Analytical scale H73 transcription reactions. See Materials and Methods for details.

H73 which is required for Dligation. The amount of GMP present in the reaction as well as the GTP concentration strongly affects the H73 yield [3]. Furthermore, it is essential to gain as much 5′ monophosphate tagged H73 RNA as possible to obtain higher yields during subsequent Dligation reactions. Consequently, different GMP:GTP ratios were also tested to achieve the highest H73 yield. Therefore, several parallel analytical-scale transcription reactions were carried out, and the yields were visually compared by separating the RNA strands by length on a polyacrylamide gel. It was clear that the reaction without GMP used as a control yielded the highest amount of desired length RNA, although this transcription product was unusable for ligation assays due to its tri-phosphate at the 5′ end. By comparison, the reaction with increased amounts of pyrophosphatase produced the next highest yield. Pyrophosphatases are acid anhydride hydrolases that act upon the diphosphate discarded after each nucleotide attachment by the polymerase. Free diphosphate can interfere with transcription (by binding to free Magnesium in the reaction) and hence pyrophosphatases are used in the reaction [17]. Although it has been used as an optional reagent in previous transcription protocols, it was found to be a crucial element in H73 transcription.

A preparative scale transcription reaction was then performed with the optimal conditions found from the analytical transcription reactions to obtain large quantities of H73. However, the RNA yield was low compared to similar reactions previously performed in other systems. Therefore, a DNA template fully complementary to that of H93 (termed H93_COMPL) was used in place of PROMOTOR. It was anticipated to produce higher yields due to full hybridization between the two DNA templates which would in turn allow the polymerase to bind and function optimally. Under identical conditions to the

previously derived optimal environment with increased pyrophosphatase amounts, this reaction produced 35-fold higher yields. Therefore, the use of the fully complementary DNA duplex was used to produce H73 RNA.

Optimization of Dligation

Once enough H73 was produced, it was then subjected to Dligation with HD36 at the G11 position to produce an RNA strand that encompasses 109 nucleotides from C-26 to A83 (termed H109). As mentioned earlier, a DNA splint was required to hold H73 and HD36 together and to bring the two ends together for ligation. This DNA splint (termed HSPLINT) was 33 nucleotides in length and was partially complementary to both RNA strands. A Dligation protocol at the G11 position was used as previously described [3]. In this previously demonstrated protocol, the acceptor RNA used was significantly shorter than HD36. Therefore, additional optimizations were required to obtain the highest possible yield of H73.

Several parallel analytical scale Dligation reactions were done under standard buffer conditions (Materials and Methods). Variations of T4 DNA ligase concentrations, incubation temperatures and incubation times were tested so that the most favorable Dligation could be achieved (Figure 3.5). A commercially available T4 DNA ligase was also compared with that of home-made enzyme isolated from *E. coli*. In addition, the concentration ratios of the three strands (donor, acceptor and the splint) in the reaction mixture were also varied. ATP was used to initiate Dligation and its concentration was not varied among tested reactions. After Dligation was completed, the reactions were

visually compared on a polyacrylamide gel to identify the conditions that yielded the highest amount of H109 material. Many of the reactions generated visually similar yields of H109. However, the reaction with H73: HSPLINT: HD36 proportion of 2:1.5:1 was chosen as the highest H109 producing conditions. Then, a preparative scale reaction under these conditions was carried out to obtain high amounts of H109.

Optimization of Rligation

Once H109 was produced by ligation of HD36 and H73 RNAs by Dligation, it was then ligated to H2 by Rligation. This method was chosen over Dligation due to the length of H2, which is not accessible to DNA splint hybridization. H2 was synthesized with a monophosphate at the 5' end in order to facilitate ligation with H109. Prior to this reaction, H2 was labeled with Cy5, and was purified using a home-made anion exchange column (Materials and Methods). Rligation was performed with slight changes to the protocol suggested by the T4 RNA ligase manufacturer (Materials and Methods). After several attempts, H111 was produced, identified and was isolated by polyacrylamide gel electrophoresis. However, the yield of H111 (labeled with Cy5) was too low even for single molecule FRET experiments. It was evident that combining both ligation methods to produce H111 was not economical.

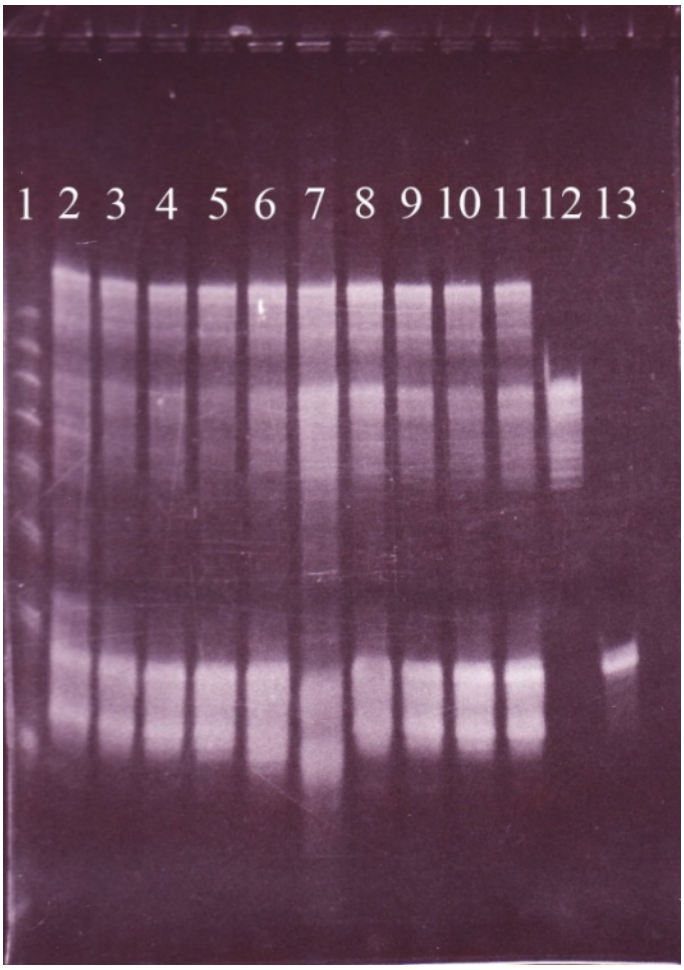


Figure 3.5 Analytical scale Dligation reactions for HD36 and H73. See Materials and Methods for details.

Revisiting Dligation and further optimizations

H75 (described earlier) was chemically synthesized with a monophosphate at the 5´ end, along with an amino linker attached to the 3´ end for post synthetic Cy5 labeling. The Cy5 labeled H75 was then ligated to HD36 using the optimal conditions found in the Dligation between HD36 and H73. However, the yield was still too low for experiments. It was then attempted to ligate unlabeled H75 to HD36, and several analytical scale reactions were done to optimize conditions (Figure 3.6). Dligation was successful in this situation (Materials and Methods), and the resulting H111 was subsequently labeled with Cy5 to obtain the final HDV ribozyme construct used.

Several further optimizations were carried out to obtain the maximum possible yield of H111. It was challenging to purify such small quantities of RNA from polyacrylamide gels. The volume of 1 mM EDTA used to extract the RNA from gel pieces was reduced to 400 µl (generally 4 mL) in order to precipitate RNA in a smaller volume of ethanol (Materials and Methods). Furthermore, Cy5 labeling was attempted both before and after gel purification of H111 because cleaner H111 without free Cy5 is desirable for smFRET experiments to prevent non-specific binding of Cy5 to the surface. Using the Cy5-labeled H111 obtained in this fashion, smFRET experiments were performed in saturating magnesium concentrations to observe the ribozyme under the conditions where it is most active.

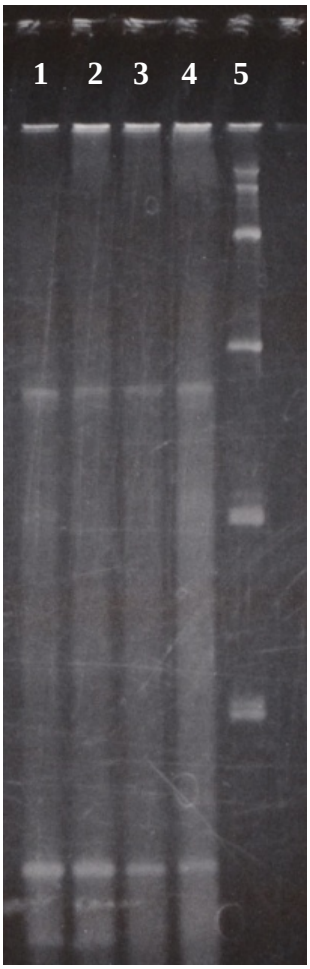


Figure 3.6 Analytical scale Dligation reactions for HD36 and H75. See Materials and Methods for details.

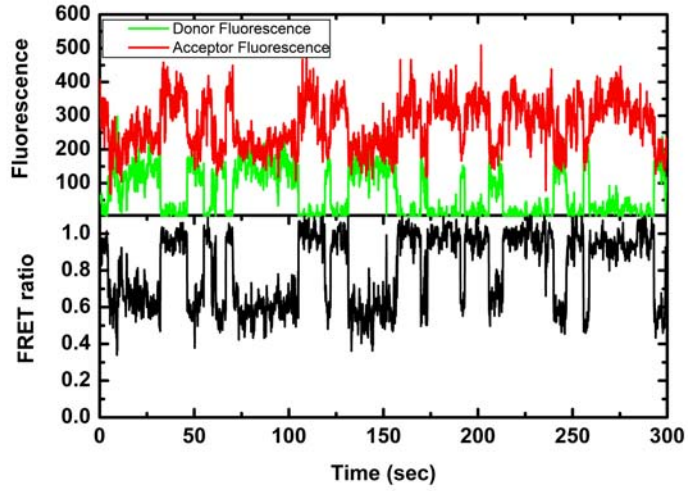
HDV ribozyme shows global dynamics

Single molecule FRET studies performed on the HDV ribozyme show clear global dynamics between the two fluorophores (Figure 3.7). FRET trajectories observed show at least three different FRET states. These states include a high (~ 0.8), a mid (~0.6) and a low (~0.2) FRET state. These preliminary data thus suggest that global dynamics exist around the U-turn of the HDV ribozyme. To further analyze these states, a histogram was plot using FRET ratios of 17 different molecules that shows a bias towards high FRET states (Figure 3.8).

3.3 Discussion

Assembling the 111 nucleotide long HDV ribozyme construct was challenging and attempts to obtain enough material for smFRET experiments by ligation of H109 and H2 were not successful. This Rligation did yield, however, some small amounts of material, hence raising the possibility that after further optimizations more material can be produced. In the recent past, a new RNA ligase called RNA ligase 2 has been used for RNA ligation, and has shown to produce less circular products, further enhancing the yield [18]. It is hence likely that using RNA ligase 2 might enhance the ligation reaction between H109 and H2. In addition, other locations on the ribozyme can be tested as potential ligation positions, such as the mid-P4 stem, for example, at the G53 nucleotide. This approach also gives rise to annealing the two strands instead of ligating them

A



B

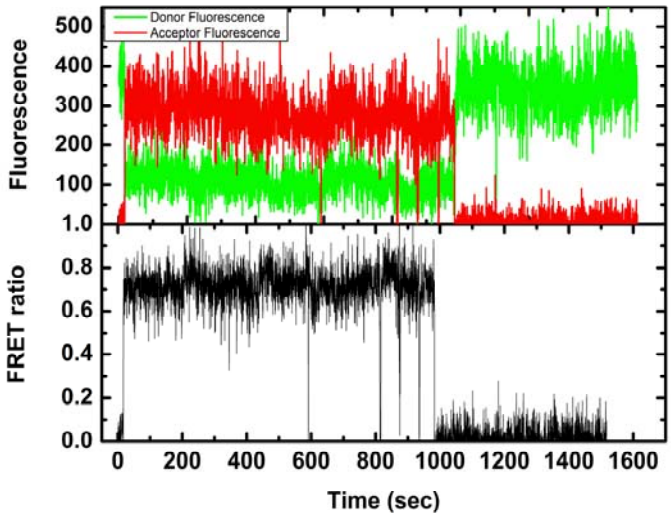


Figure 3.7 Single molecule FRET traces of two molecules. (A) Dynamics between a high (~ 1.0) FRET state and a mid (~ 0.6) FRET state. (B) Dynamics between a high (~ 0.8) FRET state and a low (~ 0.2) FRET state.

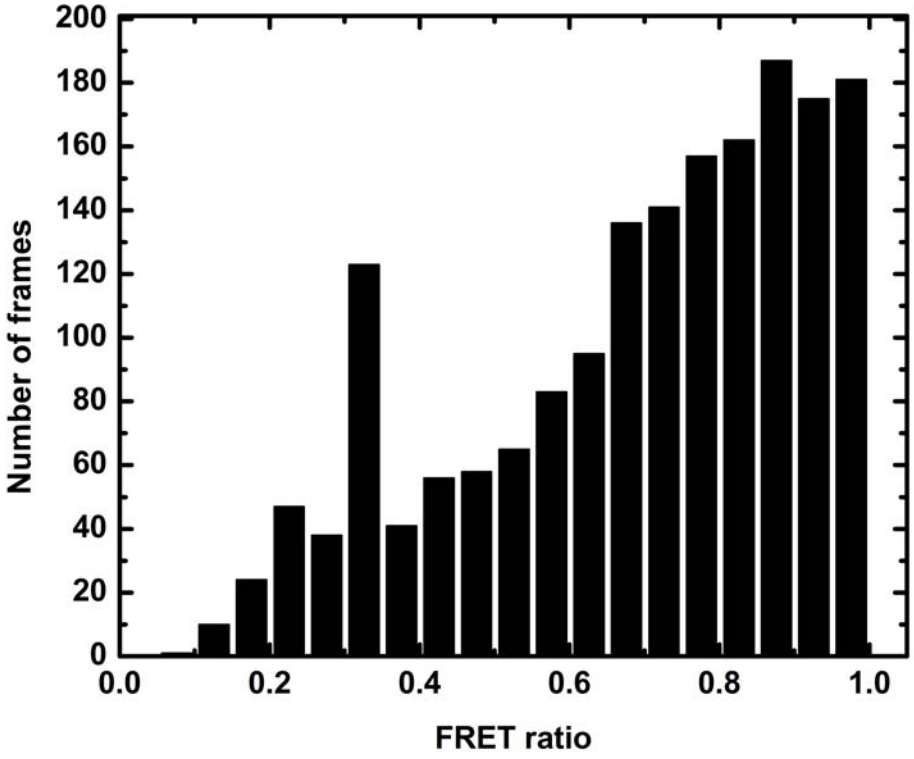


Figure 3.8 Cumulative FRET ratio histogram from 17 molecules.

together in order to monitor FRET between the two fluorophores since Watson-Crick base pairing would alone stabilize P4 to some extent.

From analytical scale transcriptions, it was clear that use of pyrophosphatase increased the H73 yield. As stated before, it is an optional reagent in many transcription protocols. These results suggest that it is indeed important to test the effect of pyrophosphatase in such reactions since it can be a significant factor. Furthermore, it is possible that with additional optimizations, the H73 yield can be further increased. One such method may be to use a DNA template with 2´-O-methyl modifications in the terminal two nucleotides as shown previously [19].

It is apparent that performing two ligation reactions to obtain one RNA strand proves to be challenging even when requiring only low amounts of RNA for single molecule studies. For studies that require large amounts of material, for example NMR studies, it is sensible to reduce the number of ligation points if feasible. For the assembly of H111, it is clear that Dligation between HD36 and H75 produces larger quantities of H111 than combining Dligation and Rligation between HD36, H73 and H2. Furthermore, it is possible to chemically synthesize the entire H111 strand with an amino linker at the 3’ end for post synthesis Cy5 labeling, but the yield of such synthetic 111 nucleotide long RNA is expected to be extremely low due to its length.

With the assembly of Cy5 labeled H111, preliminary smFRET studies were successfully performed on the genomic HDV ribozyme. Although a high FRET level was expected due to 3´ and 5´ ends coming close together in saturating magnesium concentrations, at least three different FRET states were observed. It is interesting to note that any given trajectory showed only two of the three main FRET states, suggesting a

possible memory effect as revealed previously in the hairpin ribozyme [13, 14]. A FRET ratio histogram of 17 molecules illustrates the expected shift towards a high FRET state, but the acquisition of more trajectories will be necessary to confirm these results and explain details of individual molecule behaviors, as well as provide a more comprehensive distribution of cumulative FRET ratios and kinetic information. Only then will it be possible to fully reveal the global dynamics around the U-turn as well as any FRET states and kinetic information that contribute to catalysis.

3.4 Materials and Methods

DNA and RNA sequences and preparation

All DNA strands were gel purified before use. HD36 RNA was deprotected as suggested by the manufacturer, gel purified and HPLC purified. H2 RNA was deprotected using manufacturer suggested 2´-deprotection buffer (100 mM acetic acid, adjusted pH 3.8 with TEMED) and lyophilized. H75 was deprotected as described for H2 before use. The following are all sequences purchased. All DNA oligonucleotides were purchased from Invitrogen, HD36 was purchased from Keck Foundation Resource Laboratory and H2 and H75 were purchased from Dharmacon. Their sequences are shown below.

HCAPT (DNA): *5'-Cy3-CAG GTA AGA AAG GAT GGA ACG CGG-Biotin-3'*.

HD36 (RNA): *5'-CCG CGU UCC AUC CUU UCU UAC CUG AUG GCC GGC AUG-3'*.

H93 (DNA): *5'-TCC CAT TCG CCA TTA CCG AGG GGA CGG TCC CCT CGG AAT GTT GCC CAG CCG GCG CCA GCG AGG AGG CTG GGA CTA TAG TGA GTC GTA*

TTA CTG-3'.

PROMOTOR (DNA): *5´-CAG TTA TAC GAC TCA CTA TAG-3´*.

H93_COMPL (DNA): *5´- CAG TAA TAC GAC TCA CTA TAG TCC CAG CCT CCT CGC TGG CGC CGG CTG GGC AAC ATT CCG AGG GGA CCG TCC CCT CGG TAA TGG CGA ATG GGA – 3´*.

HSPLINT (DNA): *5'-GCC AGC GAG GAG GCT GGG ACC ATG CCG GCC ATC-3'*.

H2 (RNA): *5'-pCC-3'-Amino-Modifier N6* (p = 5´- phosphate; amino-modifier to be labeled with Cy5).

H75 (RNA): *5´ – monophosphate - GUC CCA GCC UCC UCG CUG GCG CCG GCU GGG CAA CAU UCC GAG GGG ACC GUC CCC UCG GUA AUG GCG AAU GGG ACC – aminolinker - 3´*.

Transcription reaction of H73

All reactions were done under standard conditions, with the exceptions noted below. The standard buffer conditions were 30mM Tris-HCl pH 8.1, 25 mM $MgCl_2$, 2 mM Spermidine, 10 mM DTT and 0.01% Triton X-100. This solution along with both DNA templates (200 nM) each is heated to 70˚C for 2 minutes, and then allowed to cool down for 10 minutes. Subsequently, 5mM ATP, CTP, UTP and 1 mM GTP were added to the reaction. In addition, 0.001 u/µl pyrophosphatase , 0.1 mg/mL T7 RNA polymerase and ddH_2O was added. Lastly, 20 mM GMP (added last to avoid precipitation) was added and incubated at 37˚C for 3 hours. These analytical scale reactions were then run on a 10% polyacrylamide gel (Figure 3.4).

Exceptions from standard conditions are:

0.004 u/µl pyrophosphatase (lane 2)

10 mM GMP, 1 mM GTP (lane 3)

0.5 mg/mL T7 RNA polymerase (lane 4)

20 mM GMP, 1 mM GTP (lane 5)

100 nM DNA templates (lane 6)

400 nM DNA templates (lane 7)

50 mM $MgCl_2$ (lane 8)

No GMP, 5 mM GTP (lane 9)

25 mM GMP, 1.25 mM GTP (lane 10)

In addition, lane 1 has a 10 base paired DNA ladder, and lane 12 is H93.

Dligation of H73 and HD36

All reactions were done under standard buffer conditions of 66 mM Tris-HCl pH 7.6, 6.6 mM $MgCl_2$ and 10 mM DTT. A solution containing HD36, HSPLINT and H73 in 4 µM, 3 µM and 2µM respectively (2:1.5:1 ratio) and the standard buffer was heated at 90˚C for 2 minutes, and was then allowed to cool at room temperature for 10 minutes. Afterwards, 0.5 mM ATP, ddH_2O and home made 4 µM T4 DNA ligase was added, and incubated at 30˚C for 6 hours. Exceptions to this protocol were done in analytical scale ligation reactions as noted below. The reactions were then run on a 10% polyacrylamide gel to visually compare the yields of H109 (Figure 3.5).

Exceptions from the standard conditions are:

Use of commercially available (Promega) T4 DNA ligase (lane 2)

3 hour incubation time (lane 4)

Overnight incubation (lane 5)

8 µM ligase concentration (lane 6)

2:1.5:1 proportion of H73: HSPLINT: HD36 (lane 7)

3 hour incubation at 37˚C (lane 8)

6 hour incubation at 37˚C (lane 9)

6 hour incubation at 25˚C (lane 10)

6 hour incubation at 16˚C (lane 11)

In addition, lane 1 has a 10 base paired DNA ladder, lane 12 and lane 13 have H73 and HD36 respectively.

Rligation of H109 and H2

Purification of H2 (both labeled and unlabeled with Cy5) was performed with a home made anion exchange column. Approximately 500 µl of Sephedex A-25 powder in an enclosed column with 10 mL of ddH_2O was rotated and equilibrated overnight at room temperature in order to create the column. Water was then eluted and the column was packed to 1 mL. Subsequently, it was equilibrated with 10 mL of 50 mM TEAA pH 7.0. The RNA sample was then diluted to 10 mL, and passed through the column. It was then washed with 5 mL of 0.5 M TEAA pH 7.0, and eluted RNA with 3 mL of 2M TEAA pH 7.0. An additional 3 mL of ddH_2O was added and was dried overnight.

Rligation was performed with 5 µl of 10x buffer (500 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP), 1 µl of T4 RNA ligase (Takara), 10 µl of 100% DMSO, 3 µl of 0.1% BSA, ddH_2O, H109 and H2 (25x access to H109). The total volume of this reaction was 50 µl. It was then incubated for 16-18 hours at 4˚C.

Dligation between HD36 and H75

Four Dligation reactions in analytical scale were performed to find the optimal conditions. All reactions were done under standard buffer conditions (see *Dligation of H73 and HD36*). The ratio of the three strands was varied with commercially available DNA ligase except for lane 4 (Figure 3.6). Ratio of HD36: HSPLINT: H75 used were 1:1.5:2 (lane 1), 1:1:1 (lane 2), 2:1.5:1 (lane 3) and 2:1.5:1 (lane 4, with home-made DNA ligase). Lane 1 was chosen as producing the highest yield. Lane 5 shows an RNA ladder where the shown bands are 50, 80, 150 and 300 nucleotides from the bottom, respectively.

smFRET assays

The 3´-biotinylated HCAPT strand (50 pM final concentration) was annealed with the noncleavable Cy5 labeled ribozyme strand (1 nM final concentration) in standard buffer (50 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$) at 90°C for 2 minutes and cooled at room temperature for 15 min. The assembled RNA complex was supplemented with an

oxygen scavenger system to prolong the lifetime of the fluorophores, and bound at 25°C to a streptavidin coated quartz slide surface within a microfluidic channel as described until a density of ~0.1 molecules/mm^2 was reached. A home-built prism-based total internal reflection fluorescence (TIRF) microscope with intensified CCD camera (I-PentaMAX, Roper Scientific Inc.) was used to observe the donor (ID) and acceptor (IA) fluorescence signals and the resulting smFRET ratio of IA/(IA + ID) at 100-msec time resolution. The dwell times of all L and H states of each trajectory were calculated and FRET ratio histograms were plotted with Microcal Origin 7.0 software.

3.5 Acknowledgements

The authors wish to thank Drs. Ailong Ke and Jennifer Doudna for information on the G11 ligation position, Dr. John Hsieh for gifting home-made T4 DNA ligase, Dr. Hannah Townsend for Figure 3.8 and members of the Walter group for helpful suggestions.

3.6 References

1. Moore, M.J. and P.A. Sharp, *Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites.* Science, 1992. **256**(5059): p. 992-7.

2. Pereira, M.J., E.N. Nikolova, S.L. Hiley, D. Jaikaran, R.A. Collins, and N.G. Walter, *Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis.* J Mol Biol, 2008. **382**(2): p. 496-509.

3. Ke, A. and J.A. Doudna, *Crystallization of RNA and RNA-protein complexes.* Methods, 2004. **34**(3): p. 408-14.

4. Stark, M.R., J.A. Pleiss, M. Deras, S.A. Scaringe, and S.D. Rader, *An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs.* RNA, 2006. **12**(11): p. 2014-9.

5. Moore, M.J. and C.C. Query, *Joining of RNAs by splinted ligation.* Methods Enzymol, 2000. **317**: p. 109-23.

6. Romaniuk, P.J. and O.C. Uhlenbeck, *Joining of RNA molecules with RNA ligase.* Methods Enzymol, 1983. **100**: p. 52-9.

7. Gumport, R.I. and O.C. Uhlenbeck, *T4 RNA ligase as a nucleic acid synthesis and modification reagent.* Gene Amplif Anal, 1981. **2**: p. 313-45.

8. Tessier, D.C., R. Brousseau, and T. Vernet, *Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase.* Anal Biochem, 1986. **158**(1): p. 171-8.

9. England, T.E., A.G. Bruce, and O.C. Uhlenbeck, *Specific labeling of 3' termini of RNA with T4 RNA ligase.* Methods Enzymol, 1980. **65**(1): p. 65-74.

10. Sefcikova, J., M.V. Krasovska, J. Sponer, and N.G. Walter, *The genomic HDV ribozyme utilizes a previously unnoticed U-turn motif to accomplish fast site-specific catalysis.* Nucleic Acids Res, 2007. **35**(6): p. 1933-46.

11. Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna, *Crystal structure of a hepatitis delta virus ribozyme.* Nature, 1998. **395**(6702): p. 567-74.

12. Das, S.R. and J.A. Piccirilli, *General acid catalysis by the hepatitis delta virus ribozyme.* Nat Chem Biol, 2005. **1**(1): p. 45-52.

13. Ditzler, M.A., D. Rueda, J. Mo, K. Hakansson, and N.G. Walter, *A rugged free energy landscape separates multiple functional RNA folds throughout denaturation.* Nucleic Acids Res, 2008. **36**(22): p. 7088-99.

14. Zhuang, X., H. Kim, M.J. Pereira, H.P. Babcock, N.G. Walter, and S. Chu, *Correlating structural dynamics and function in single ribozyme molecules.* Science, 2002. **296**(5572): p. 1473-6.

15. de Silva, C. and N.G. Walter, *Leakage and slow allostery limit performance of single drug-sensing aptazyme molecules based on the hammerhead ribozyme.* RNA, 2009. **15**(1): p. 76-84.

16. Hoerter, J.A., M.N. Lambert, M.J. Pereira, and N.G. Walter, *Dynamics inherent in helix 27 from Escherichia coli 16S ribosomal RNA.* Biochemistry, 2004. **43**(46): p. 14624-36.

17. Cunningham, P.R. and J. Ofengand, *Use of inorganic pyrophosphatase to improve the yield of in vitro transcription reactions catalyzed by T7 RNA polymerase.* Biotechniques, 1990. **9**(6): p. 713-4.

18. Nichols, N.M., S. Tabor, and L.A. McReynolds, *RNA ligases.* Curr Protoc Mol Biol, 2008. **Chapter 3**: p. Unit3 15.

19. Martick, M. and W.G. Scott, *Tertiary contacts distant from the active site prime a ribozyme for catalysis.* Cell, 2006. **126**(2): p. 309-20.

CHAPTER FOUR

IMAGING NANOMETER-SCALE AUTONOMOUS MOVEMENT OF MOLECULAR SPIDERS BY SINGLE PARTICLE FLUORESCENCE TRACKING[4]

4.1 Introduction

In recent years, nanometer scale particles have emerged as microscopic machines that resemble and mimic macroscopic motors, ratchets, rotors and even vehicles [1-9]. These nano-machines are promising devices that are being developed for application in engineering, medicinal and biological fields of studies. Furthermore, molecular recognition properties of nucleic acids (such as the biosensor described in Chapter Two, with binding properties that differentiate between small molecules) can be utilized for the construction of nano-machines as intelligent drug delivery devices, to perform information processing tasks, to be used as biosensors to detecting biohazards and to perform various mechanical tasks such as pulling and stretching [6, 10-13]. Among the most important current challenges in this field is designing a nano-walker that can traverse specific landscapes and take directions to start, stop and direct itself on

[4]All experiments and analysis were performed by Chamaree de Silva, Matlab software used for Gaussian fitting of PSFs (Appendix Two) was provided by Dr. Anthony J. Manzo, Matlab software for displacement and velocity calculations (Appendix Two) was provided by Alexander Johnson-Buck, and Spiders were prepared by Dr. Steven Taylor.

designated path. Such movers can potentially become energy efficient autonomous machines for clever delivery of drugs or other molecules.

In the past few years, scientists have developed various such molecular nano-walkers that incorporate nucleic acids to afford high binding specificity. One such example is the molecular Spider, also called Nanoassembly Incorporating Catalytic Kinesis (NICK) [14]. Molecular Spiders resemble the anatomy of a spider with two to six 8-17 DNAzymes ("legs") attached to one or two streptavidin molecules ("bodies") by the strong biotin-streptavidin interaction (Figure 4.1A). Here, a Spider with three identical catalytically active legs and one inactive leg is used. It has been shown that Spider legs can hybridize with complementary hybrid DNA/RNA molecules (hereafter called substrates) displayed on a solid surface in the presence of monovalent ions (Figure 4.1B). The substrate is DNA with the exception of one ribonucleotide where the backbone is cleaved when the substrate is hybridized to a Spider leg in the presence of a divalent cation, which acts as a cofactor of the reaction. Substrate turnover occurs most efficiently in the presence of Zn^{2+}, and involves binding of substrates to legs, cleavage of substrates, and dissociation of Spider legs from the reaction products. Given a high surface density of substrates and due to the multivalency of the spider it is likely that a free leg rebinds a fresh substrate before all legs detach, keeping the spider associated with the substrate coated surface. As a result of the lower affinity of the spider legs with the cleaved products remaining on the surface compared to the substrate, the spider will continuously move towards regions of fresh substrate. This walk has been studied using Surface Plasmon Resonance (SPR) and characterized using multiple different Spiders and

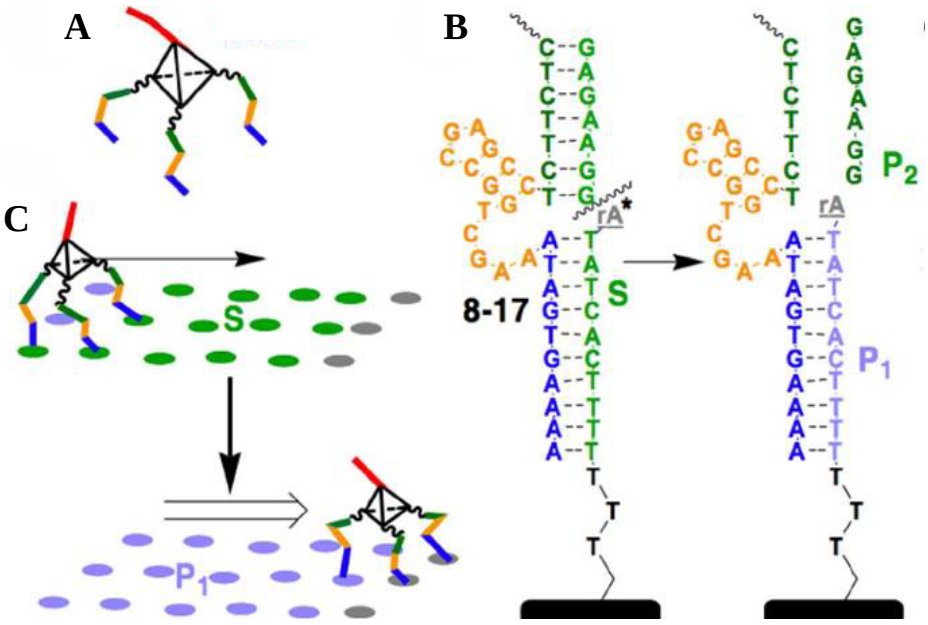


Figure 4.1 Spider with three active legs and one inactive leg. (A) Inactive leg is shown in red (B) Sequences of Spider leg (8-17 DNAzyme) and substrate. The arrow indicates addition of Zn^{2+} (C) Movement of Spider on a two dimensional substrate surface as hypothesized from SPR experiments. Figure by Dr. Kyle Lund.

environments [14]. In addition, theoretical modeling studies have been performed to predict Spider behavior [15, 16].

SPR measurements demonstrate the interaction between substrates and Spider legs, but do not reveal specific information about the movement of individual Spiders on a substrate matrix. Modeling predicts that Spiders would move on a substrate field in a self-repellent random walk (Figure 4.1C). Modern single molecule imaging offers a remarkable tool to directly visualize the movement of individual Spiders, as previously explained in Chapter One. Such imaging techniques have been previously developed to track particles in real-time with nanometer precision. A frequently used method of particle tracking is Fluorescent Imaging Of Nanometer Accuracy (FIONA), wherein the centroids of imaged Point Spread Functions (PSFs) are determined by fitting with two-dimensional Gaussian functions [17]. To apply this method to Spiders, their streptavidin body was labeled with multiple Cy3B fluorophores. Substrates were immobilized on a quartz slide, and the fluorescence emitted by the Spider fluorophores was monitored by Total Internal Reflection Fluorescence Microscopy (TIRFM).

For the Spiders to freely move on their substrate matrix, it is important that the distance between substrates on the quartz slide is less than the span of Spider legs (~ 25 nm). After using several different methods to achieve an optimal substrate density, a final 4-5 nm distance between substrates was reached (Appendix Two). Spiders were then hybridized to immobilized substrates in a flow channel (Materials and Methods), and incubated with Zn^{2+} before monitoring the fluorescence with TIRFM video imaging. After each experiment, the PSFs representing a single fluorophore labeled Spider over consecutive images were analyzed using FIONA (Materials and Methods), allowing us to

monitor Spider movement in real-time. By visualizing individual Spider trajectories, it was clear that some Spiders move across the substrate field quite freely, with total distance travelled exceeding 1000 nm within two hours.

4.2 Results and Discussion

In preliminary fluorescence studies of Spider movement, it became apparent that stage-drift due to thermal or vibrational fluctuations could be as large as the expected movement of Spiders. This was established by the observation that all Spiders in a given experiment displayed similar movement patterns (Appendix Two). To alleviate this problem, Quantum Dots (Q-dots) of 705 nm emission were immobilized on the quartz slide to be used as reference points (fiduciary markers) for drift correction. Since the emission from Q-dots was recorded on a different section of the CCD, they were easily differentiable from the Cy3B labeled Spiders (Figure 4.2A). Q-dots also show higher fluorescence intensity than Cy3B labeled Spiders under the same illumination conditions (532 nm laser excitation). PSFs of Q-dots were then subjected to Gaussian fitting and their movement was observed throughout the experiment. This movement occurs mostly due to stage-drift since Q-dots are immobilized to the quartz slide. As expected, all Q-dots displayed similar, but not identical drift patterns (Figure 4.2B). Slight differences between the patterns can be attributed to the limited resolution of the fitting procedure, which is calculated to be ~20 nm, and slight differences in drift across a field of view. To further improve the accuracy, (x,y) coordinates of Q-dot trajectories from the same field of view were averaged. The remaining error or difference between any two Q-dot

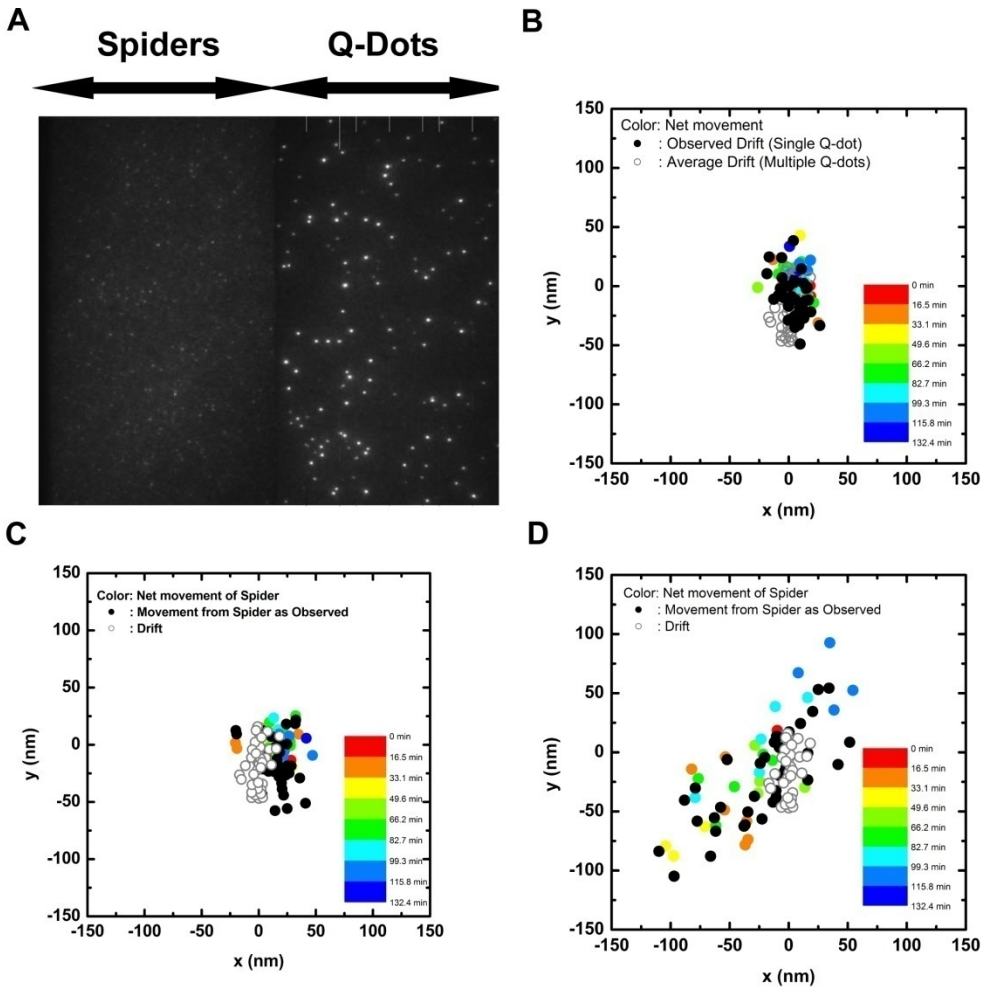


Figure 4.2 Movement of Spiders. (A) Dual view of Spiders (left half) and Q-dots (right half) on the slide. (B) Observed stage-drift during a two hour experiment (C) Net movement of a 'stationary' Spider (D) Net movement of a 'walking' Spider.

trajectories was observed to be <50 nm. From Q-dot tracking, it was also clear that the stage-drift can significantly vary between experiments. Therefore, coordinates from multiple (~5-7) Q-dots that were arbitrarily picked to represent the entire field of view were averaged to obtain a drift trajectory for each experiment, which in turn was subtracted from Spider trajectories. This correction for stage-drift allowed calculation of two-dimensional net movement of Spiders (Figure 4.2 C-D).

In order to observe multiple Spiders, multiple identical experiments were performed and PSFs of fluorescence spots were analyzed. A typical experiment was two hours long with 41 imaged frames (Materials and Methods). From these spots, ones that did not completely photobleach for a minimum of 30 out of 41 frames were selected. Each of the PSFs of these remaining spots was considered to be a Spider, and was corrected for stage-drift for their net movement against the fiduciary Q-dot makers as explained above. These net trajectories fell into three categories: (1) "Movement" only within an error similar to the residual error observed for Q-dots (<50 nm each direction – Figure 4.3A). (2) Movement >50 nm in any direction, but no clear directional path, i.e. noisy (Figure 4.3B). (3) Movement >100 nm at least in one direction with clear directionality (Figure 4.3C-F).

Spiders that fell into the first category were interpreted as not exhibiting any movement and hence were denoted as 'stationary'. We cannot exclude that these Spiders moved within <50 nm, but such movement is within our resolution limit. Out of 76 total fluorescent spots analyzed, 34 (43%) showed such behavior. Spiders that fell into the second category showed movement beyond 50 nm, but their trajectories showed no directionality. This observation was consistent with Spiders grazing substrates in a certain

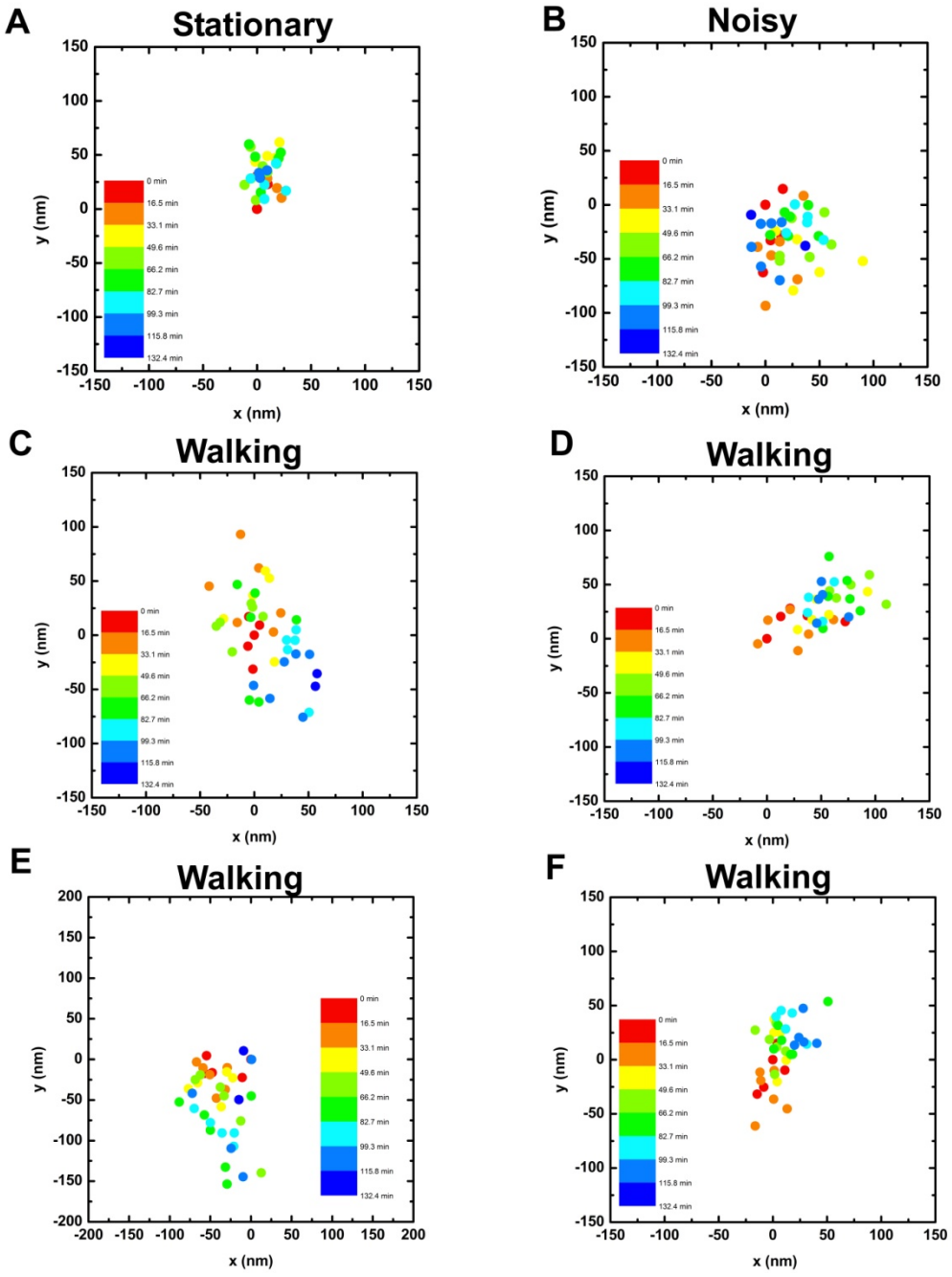


Figure 4.3 Net movements of Spiders. (A) 'stationary' Spider (B) 'noisy' Spider (C-F) 'walking' Spiders.

area and being trapped with no fresh substrates to hybridize around them. 23 (30%) of the analyzed fluorescence spots showed such behavior. Although studies performed on substrate densities on the slide showed a homogeneous spread of substrates on the slide surface, it is possible that substrate patches exist on the slide where Spiders can get trapped. At the same time, Spiders can trap themselves by grazing on a perimeter and moving into the center of the grazed area. Such Spiders were categorized as 'noisy'.

Twenty (27%) of the analyzed fluorescent spots showed significant movement compared to that of our resolution limit and showed directionality over time. These were considered to be "walking" Spiders. These Spider trajectories provide direct evidence of two-dimensional movement, which could not be directly established by the previous SPR measurements. It is important to note that each such walking Spider in a given experiment shows its own unique trajectory of movement, further validating the interpretation as autonomous Spiders and indicating a certain level of individuality. Therefore, we conclude that we here demonstrate by single particle tracking that fluorescently tagged Spiders move on a two-dimensional field of substrate and that we can begin to compare their behaviors. By allowing Spiders to move on an unrestricted pathway (i.e., the area covered by substrates is far greater than the Spider's net movement), we demonstrate here that they display the expected random walk. Spiders from all three categories (stationary, noisy and walking) were observed in each experiment, further validating out conclusion. Figure 4.4 shows the layout of all fluorescence spots in a particular experiment, and stationary Spiders (red circles), noisy Spiders (orange circles) and walking Spiders (green circles) are distributed throughout the field of view.

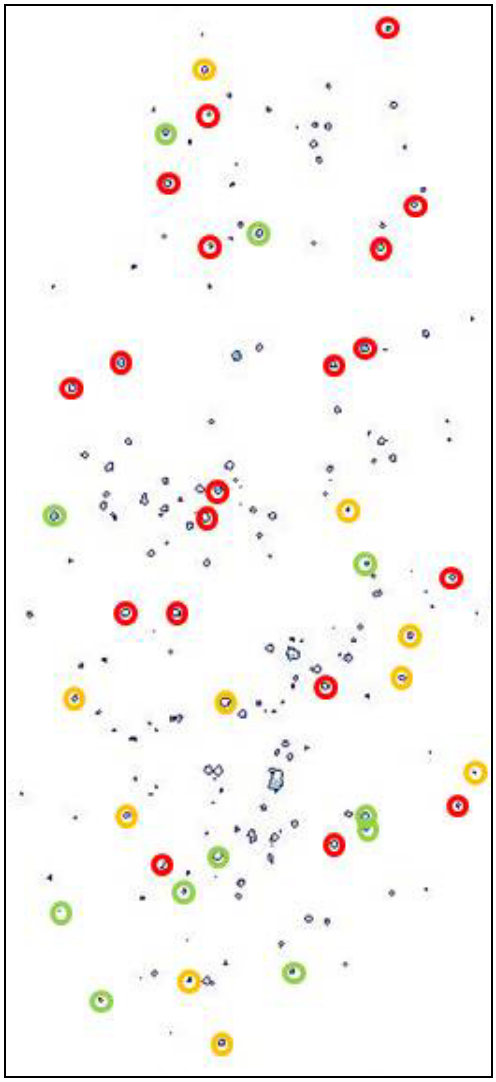


Figure 4.4 Spiders on a field of view. Red, orange and green circles denote ‘stationary’, ‘noisy’ and ‘walking’ Spiders, respectively.

In analyzing the movement of 20 Spiders that show walking behavior, distances traversed and velocities were calculated. More specifically, for each spider the total distance travelled was estimated by calculating the distance between its locations in each pair of frames, and adding them together over the entire duration of the trajectory (Fig 4.5 - top panel). This total distance travelled varied from ~950 nm to 2,500 nm for different Spiders. It is important to note that this distance is a lower bound of the total distance travelled since it is unlikely that Spiders move in a straight line between frames. In fact, Spiders were observed to change direction and often moved towards the starting location, giving rise to a net displacement significantly lower than the total travelled distance, from ~2 nm to 300 nm (Fig 4.5 – middle panel). It is thought to be thermodynamically unfavorable for Spiders to return to already visited (cleaved) substrates since only a part of the Spider leg can hybridize to the product left after cleavage after dissociation of the other product (half of the substrate). This driving force is predicted to propel Spiders forward. However, by observing Spiders that move a considerable distance and show changes in directionality, it is clear that Spiders can move back towards the start location quite often. Since substrates are ~4-5 nm apart and the Spider leg span is ~25 nm, the Spider body is expected to be positioned at a location central to all of its legs hybridized to substrate. Therefore, by observing fluorophores attached to the Spider body (and given the resolution limit), the exact location of any single leg cannot be resolved, and hence it cannot be determined whether a particular substrate is indeed visited twice upon return of the Spider, or rather a neighboring one (Figure 4.3 C-E).

The velocity for each Spider was estimated by dividing the total distance travelled by the duration of the observation (Fig 4.5 – bottom panel). Velocities ranged from ~8 to

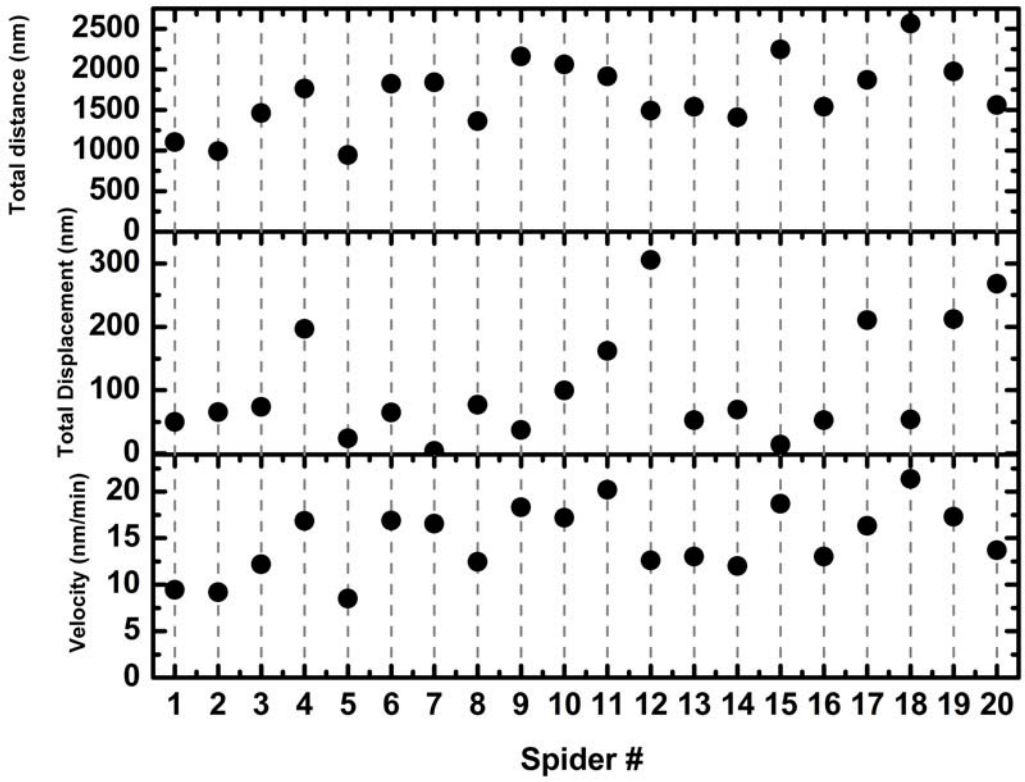


Figure 4.5 Displacement and velocities of 'walking' Spiders. The total distance traversed (top panel), total displacement travelled (middle panel) and the average velocity (bottom panel) of individual Spiders are shown.

22 nm/min with an average value of ~15 nm/min. These values for individual spiders are consistent with their experimentally measured ensemble leg cleavage rate [14] and thus further supports the notion that Spiders can move on a two-dimensional substrate surface. These results begin to reveal details of the *Spider World*, however, it is important to probe Spiders further by varying environmental conditions to study their programmability. For example, experiments with no Zn^{2+} added should be performed to analyze and understand behavior under which Spiders are thought to be stationary. In addition, experiments can be performed on an only product covered surface (Figure 4.1B – P1). In principle, the Spiders should dissociate under such conditions. However, if we were to observe that Spiders can indeed move on products, it would raise the possibility that Spider legs visit the same substrate multiple times. Even though more experiments need to be performed in order to fully understand Spider behavior, information gathered here by monitoring individual Spiders using single molecule fluorescence microscopy corroborates previous models and brings us one step closer to manipulating Spiders for drug delivery.

4.3 Materials and Methods

Spider Preparation

Synthesis and purification of the modified DNA strands used to construct the Spiders were carried out by Integrated DNA Technologies (Coralville, IA) and used as received. Streptavidin was obtained from Pierce, product number 21125, (Rockford, IL).

IEHPLC purification was performed using a Shimadzu LC-6AD pump equipped with a Shimadzu SPD-M10A PDA detector, with separation carried out on an anion exchange TSKgel DEAE-NPR column, 4.6x50 mm (IDxL), (Tosoh Biosciences). Concentrations of oligonucleotides were determined on an Amersham Biosciences Ultrospec 3300 pro UV/visible spectrophotometer. Here, iSp18 is a hexa-ethyleneglycol internal spacer, Bio is biotin and BioTEG is biotin-tetra-ethyleneglycol conjugate. Assembly of Spider was performed in three parts.

Part A; (5' – /5Cy3/GCC GAG AAC CTG ACG CAA GT/iSp18//iSp18//3Bio/ - 3') (termed **C**) (47 nmoles in 10 mL of 20 mM triethanolamine /150 mM NaCl/pH 7.4) was added drop-wise to a stirred solution of streptavidin (STV) (5 mg, 94 nmoles in 1 mL of 10 mM 3PO4/pH6.5). The desired one-to-one conjugate product ("STV-(**C**)1") was purified by ion exchange (IE) HPLC.

Part B; (5' - /5BioTEG//iSp18//iSp18/TCT CTT CTC CGA GCC GGT CGA AAT AGT GAA AA - 3') (termed **L**) (100 µM, in water) was titrated into the collected 1:1 conjugate HPLC fraction from 'Part A' above, until all three remaining biotin binding sites of the 1:1 conjugate "STV-(**C**)1" were occupied by **L** to give the final desired product "STV-(**C**)1(**L**)3". The titration was monitored by IE-HPLC, and was deemed complete when a slight excess of **L** was observed with no intermediate species, i.e. "STV-(**C**)1(**L**)1" or "STV-(**C**)1(**L**)2", present. The assembly was purified by IE-HPLC and the volume of the elutent reduced (by centrifugation) to give a final concentration of 2.3 µM, as determined

by absorbance at 260 nm. Characterization of the assembly was carried out by IE-HPLC and PAGE. The assembly was stable at -20 °C for at least six months.

Part C; the volume of Cy3B labeled Spider, fraction isolated by HPLC, was concentrated to 1 mL (0.834 nmoles) and Cy3 Mono NHS ester (20 nmoles) (PA13101, Lot number 359269, GE Healthcare) dissolved in DMSO added to the solution containing the assembly (giving a total DMSO concentration of 10%). The resulting mixture was incubated at room temperature overnight, protected from light. Excess dye was separated from the Cy3B labeled Spider by gel filtration (PD-10 column, 17-0851-01, lot 367770, GE Healthcare). Ratio of fluorophore to streptavidin-DNA assembly was obtained by determining concentrations at 550 nm (ε_{max} 150,000 M^{-1} cm^{-1}) and 260 nm (ε_{max} 1,220,000 M^{-1} cm^{-1}) respectively.

Slide Preparation for Single Molecule Fluorescence Studies

Two holes were drilled on each quartz slide that was used to allow sample flow. To clean the slides, they were immersed in a 5% (v/v) ammonium hydroxide and 14% (v/v) hydrogen peroxide solution and boiled for 20-30 minutes, followed by sonication in 1M KOH for one hour. The slides were then flamed for about one minute each. To aminosialanize the slides, they were emerged in a solution containing 5% (v/v) solution of 3-aminopropyltriethoxysilane in acetone for one hour. Then the slides were rinsed with acetone and dried for one hour at 80°C. A layer of the bifunctional crosslinking agent para-diisothiocyanatobenzene (PDC) was covalently coupled to the aminosilanized

surface by incubating the slides for two hours in a solution of 0.2 % (v/v) PDC, 10% (v/v) pyridine in DMF. Afterwards, the slides were thoroughly rinsed with methanol and acetone. Then 70 μL of 0.5 mg/mL avidin was added to each slide, and the middle sections of the slides with the holes were covered with a glass cover slip. After two hours incubation in a container with moisture to avoid drying out, the cover slips were removed, and the slides were washed with deionized water, followed by a solution containing 1 M NaCl and 40 mM NaOH. Then the slides were thoroughly washed with deionized water again and dried under nitrogen. Two pieces of double sided tape were placed parallel to the connecting line of the two holes in order to make a flow channel about 2-3 mm wide. A fresh glass coverslip was placed on top of the tape and then was glued around the perimeter to seal the edges. The slides were stored at 4°C in a vacuum desiccator up to 3-4 weeks.

Sample Preparation for Single Molecule Fluorescence Microscopy

The slides were taken out of 4°C (unless they were made fresh the day of data collection), and incubated at room temperature for 10-15 minutes. 1X buffer (150 mM NaCl, 10mM HEPES-KOH pH 7.4) was flowed through the channel to moist the avidin. Then the slide was incubated with 10 μM substrate and 2 pM of 705 nM emission quantum dots for 30 minutes, followed by 1X buffer to wash away the unbound substrates and quantum dots. Afterwards, 3 pM Spider was added with the oxygen scavenging system (OSS) containing ProtoCatechuate Dioxygenase (PCD), protocatechuate (PCA), and Trolox in 1X buffer [18]. After a 10 minute incubation of

Spiders, OSS in 1X buffer was added to the channel to wash the unbound Spiders. After an additional 15 minute incubation, 1mM $ZnSO_4.7H_2O$ in OSS was added to activate the Spiders. This preparation was done in near darkness to protect the quantum dots and fluorophores from photobleaching.

Data Collection for Single Particle Tracking

Slides prepared with samples were imaged by total internal reflection microscopy as described previously [19, 20], and Cy3B and Q-dots were excited by a 532nm laser at 25°C. Then, a section of the channel with suitable Spider and Q-dot density was imaged with a 60X objective for typically two hours. Integration time for one frame was 2.5 seconds (continuous laser light) followed by 177.5 seconds of darkness (laser light cut off) to decrease photobleaching. This gives one frame every 180 seconds (three minutes). Therefore, a typical movie was 41 frames in length.

Data Analysis by Gaussian Fitting to Point Spread Functions

Point spread functions (PSFs) from Cy3B and quantum dots were imaged on spectrally separated halves of an intensified CCD camera, resulting in intensity traces that reflect the photon count of each peak of the PSF in at each frame. Stepwise photobleaching of fluorophores were sometimes observed, although traces with no photobleaching events were optimal to obtain traces for the entire time they were imaged.

Traces that were excluded based on low or absent signal intensity from Cy3B likely resulted from other fluorescent contaminants on the slide.

Each PSF was fit frame-by-frame with a two-dimensional Gaussian function of the form of:

$$f(x,y) = z_0 + A\exp\left\{\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_x}\right)^2\right]\right\}$$

Multiple quantum dots from different locations on the field of view were fitted with PSFs and the position values μ_x and μ_y for each were obtained, normalized for the start position (frame one) to be (0,0). Coordinates from such trajectories were averaged over the number of used quantum dots. These averaged coordinates were used as the stage drift correction factor for the Spider traces. Some experiments had considerable drift through x-y plane, while others were quite stable.

4.4 References

1. Leigh, D.A., J.K. Wong, F. Dehez, and F. Zerbetto, *Unidirectional rotation in a mechanically interlocked molecular rotor.* Nature, 2003. **424**(6945): p. 174-9.

2. van Delden, R.A., M.K. ter Wiel, M.M. Pollard, J. Vicario, N. Koumura, and B.L. Feringa, *Unidirectional molecular motor on a gold surface.* Nature, 2005. **437**(7063): p. 1337-40.

3. Fletcher, S.P., F. Dumur, M.M. Pollard, and B.L. Feringa, *A reversible, unidirectional molecular rotary motor driven by chemical energy.* Science, 2005. **310**(5745): p. 80-2.

4. Kelly, T.R., H. De Silva, and R.A. Silva, *Unidirectional rotary motion in a molecular system.* Nature, 1999. **401**(6749): p. 150-2.

5. Serreli, V., C.F. Lee, E.R. Kay, and D.A. Leigh, *A molecular information ratchet.* Nature, 2007. **445**(7127): p. 523-7.

6. Khatua, S., Guerrero, J.M., Claytor, K., Vives, G., Kolomeisky, A.B., Tour, J.M., Link, S., *Micrometer-scale translation and monitoring of individual nanocars on glass.* ACS Nano, 2009.

7. Morin, J.F., T. Sasaki, Y. Shirai, J.M. Guerrero, and J.M. Tour, *Synthetic routes toward carborane-wheeled nanocars.* J Org Chem, 2007. **72**(25): p. 9481-90.

8. Shirai, Y., J.F. Morin, T. Sasaki, J.M. Guerrero, and J.M. Tour, *Recent progress on nanovehicles.* Chem Soc Rev, 2006. **35**(11): p. 1043-55.

9. Shirai, Y., A.J. Osgood, Y. Zhao, K.F. Kelly, and J.M. Tour, *Directional control in thermally driven single-molecule nanocars.* Nano Lett, 2005. **5**(11): p. 2330-4.

10. Kolpashchikov, D.M. and M.N. Stojanovic, *Boolean control of aptamer binding states.* J Am Chem Soc, 2005. **127**(32): p. 11348-51.

11. Seelig, G., D. Soloveichik, D.Y. Zhang, and E. Winfree, *Enzyme-free nucleic acid logic circuits.* Science, 2006. **314**(5805): p. 1585-8.

12. Stojanovic, M.N., S. Semova, D. Kolpashchikov, J. Macdonald, C. Morgan, and D. Stefanovic, *Deoxyribozyme-based ligase logic gates and their initial circuits.* J Am Chem Soc, 2005. **127**(19): p. 6914-5.

13. Lederman, H., J. Macdonald, D. Stefanovic, and M.N. Stojanovic, *Deoxyribozyme-based three-input logic gates and construction of a molecular full adder.* Biochemistry, 2006. **45**(4): p. 1194-9.

14. Pei, R., S.K. Taylor, D. Stefanovic, S. Rudchenko, T.E. Mitchell, and M.N. Stojanovic, *Behavior of polycatalytic assemblies in a substrate-displaying matrix.* J Am Chem Soc, 2006. **128**(39): p. 12693-9.

15. Antal, T. and P.L. Krapivsky, *Molecular spiders with memory.* Phys Rev E Stat Nonlin Soft Matter Phys, 2007. **76**(2 Pt 1): p. 021121.

16. Antal, T., P.L. Krapivsky, and K. Mallick, *Molecular Spiders in One Dimension.* J Stat Mech, 2007. **2007**: p. P08027.

17. Yildiz, A., J.N. Forkey, S.A. McKinney, T. Ha, Y.E. Goldman, and P.R. Selvin, *Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization.* Science, 2003. **300**(5628): p. 2061-5.

18. Aitken, C.E., R.A. Marshall, and J.D. Puglisi, *An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments.* Biophys J, 2008. **94**(5): p. 1826-35.

19. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.

20. Walter, N.G., C.Y. Huang, A.J. Manzo, and M.A. Sobhy, *Do-it-yourself guide: how to use the modern single-molecule toolkit.* Nat Methods, 2008. **5**(6): p. 475-89.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Introduction

Nucleic acid enzymes and applications using these enzymes have enriched biological and medicinal fields of studies for decades. The importance of nucleic acids to cellular processes is now understood to rival that of proteins. It has become clear that nucleic acid enzymes can play a vital role in cellular processes from translation to gene regulation and that further studies are required to truly understand such systems and to manipulate them in useful ways. In this dissertation, ribozymes as well as engineered nucleic acid enzymes have been explored using fluorescence based techniques. These molecules were probed at the single molecule level in order to understand their heterogeneous characteristics that ensemble studies cannot observe. Here, three systems were investigated with fluorescence microscopy to observe one molecule at a time. The three systems are: (1) An engineered hammerhead aptazyme, (2) the Hepatitis Delta Virus ribozyme, and (3) nano-robots with DNA enzyme legs known as “Molecular Spiders”.

5.2 Hammerhead aptazyme

Summary: The hammerhead aptazyme is leaky and inefficient compared to the naturally occurring glmS catalytic riboswitch

The hammerhead ribozyme incorporating a theophylline aptamer was particularly suited for study at the single molecule level. The work in this dissertation is the first application of single molecule techniques to this aptazyme, and our work has highlighted both the strengths and weaknesses of such an engineered riboswitch in Chapter Two. In addition, no single molecule studies were previously performed on the hammerhead ribozyme itself to understand its dynamics in the absence of the aptamer domain. It was only a few years ago that a naturally occurring catalytic riboswitch was discovered, which presented a unique opportunity to compare the characteristics of the engineered theophylline aptazyme to that of the naturally evolved *glmS* riboswitch.

The hammerhead aptazyme was previously subjected to fluorescence and radioactive studies in ensemble measurements, and did not show heterogeneity under time-resolved FRET studies [1]. In order to more deeply understand the dynamic link between the theophylline-binding aptamer domain and the catalytic core of the hammerhead ribozyme, structure and dynamics of the aptazyme was monitored by smFRET. From smFRET studies, it became clear that two distinct FRET states, Low (L) and High (H) existed, and molecules show global dynamics between the two states. Furthermore, increasing theophylline concentrations shift the equilibrium from the H to the L state, indicating a correlation between catalysis and the L state. By analyzing the

kinetics of inter-conversion of the two FRET states, it also became evident that each two kinetic states exist for both L and H state. One of the L states is very long lived, and the fraction in this state increases with increasing ligand concentration. By globally fitting both L state fractions with respect to the ligand concentration, $Th_{1/2}$ was obtained. Intriguingly, the $Th_{1/2}$ calculated from ensemble cleavage assays is equal within error to that of $Th_{1/2}$ obtained from L states in single molecule studies. This provided further evidence that the L state, and the long lived L state in particular, has a direct correlation to catalysis. Moreover, 2-AP fluorescence experiments were performed to observe local conformational changes in the aptamer and core domains. 2-AP data illustrated that although ligand binding in the aptamer is rapid, the core formation is quite slow. This suggests structural rearrangements that occur after ligand binding are the rate limiting step for catalysis. Finally, it was observed by steady-state FRET that even in the absence of the ligand, the aptazyme is partially active. Thus, between the slow conformational change, and the leaky activity in the absence of ligand, the hammerhead aptazyme is a less optimal biosensor than the naturally evolved *glmS* ribozyme.

Future directions for the aptazyme project

Work presented in Chapter Two clearly suggests that the engineered aptazyme falls short of its expected efficiency, in part due to slow conformational change steps in the kinetic mechanism. Although a simplest possible kinetic pathway was proposed from observed and calculated kinetic rates, further work is needed to truly understand the complete pathway. For example, the cleavage rate when ligand is not present (leakage

pathway) observed with steady-state FRET assays gives only an overall rate composed of many reaction steps of the pathway that leads up to cleavage (Chapter Two - Figure 2.7). Therefore, it is important to find rates for additional steps on the reaction pathway in order to truly understand the system. If more information could be obtained on at least a few unknown kinetic rates, it would be possible to simulate rates for a complete kinetic pathway. It may be possible to observe multiple individual steps on the pathway by using methods such as NMR, which could distinguish between two conformational states by monitoring changes in hydrogen bonding patterns. This would likely be difficult in practice, however, since it is not physically possible to separate certain steps of the pathway, for example, the two high FRET states depicted.

Although the aptazyme lacks desirable speed compared to a naturally occurring riboswitch, leakage plays a more vital problem in a therapeutic environment. Since the therapeutic range for theophylline is about 55 – 110 μM in blood plasma, the cleavage rate between such amounts and leakage is only about 10 fold in difference at 100 mM $MgCl_2$ [2-4]. This is quite dangerous since it is important to accurately measure theophylline concentrations at micro molar concentrations. The hammerhead ribozyme is known to cleave even in the absence of cations, and therefore the aptazyme perhaps inherits leakage activity. It is possible to design the aptazyme where it is completely misfolded in the absence of the ligand to eliminate leakage. This however will possibly contribute to an even slower aptazyme of recognizing theophylline. Other suggestions on improving the efficiency could be to shorten the communication module or to position the aptamer domain in Stems I or III, although only thorough analysis of such designs can give information on their competence as biosensors.

5.3 HDV ribozyme

Summary: The HDV ribozyme can be assembled from fragments by ligation and is dynamic

The theophylline biosensor was developed using basic studies performed on the well-characterized hammerhead ribozyme. A wealth of knowledge regarding the ribozyme was needed to understand the minimal structure required for cleavage, its structure and function as well as catalytic rates. Therefore, it is important to closely explore other ribozymes to uncover their own unique characteristics that will be useful for future applications. The HDV ribozyme is the only ribozyme from a pathogen found in humans (so far) and is also embedded in the human genome, making it of particular interest. Although many studies have been performed on the HDV ribozyme regarding its catalytic properties and the effect of sequence variation on catalysis, it has not been scrutinized at the single molecule level. Information derived from single molecule studies such as global dynamics and kinetics around the catalytic domain of the ribozyme can possibly assist future therapeutics for the Hepatitis delta virus, as well as help understand behaviors of other small ribozymes. Furthermore, single molecule studies will reveal detailed information about the mechanism of catalysis around the U-turn as mentioned in Chapters One and Three.

Assembling a functionally relevant RNA molecule to be used in fluorescence studies can be challenging due to the necessary RNA length and the need for site-specific modifications. In Chapter Three, HDV ribozyme assembly is optimized using various

RNA ligation methods. An attempt to assemble it with three RNA strands (termed HD36, H73 and H2) using both DNA and RNA ligase produced a very low yield of the final material. Therefore, a second protocol was developed to ligate just two RNA strands (HD36 and H75) using DNA ligase. This method was successful, and enough material for single molecule studies was obtained. Preliminary studies show that the ribozyme is dynamic at high magnesium concentrations.

Future directions for the HDV ribozyme project

To better understand the characteristics of the HDV ribozyme, it is important to compare the native sequence to a construct containing several mutations. The wild type genomic form of the ribozyme has a U nucleotide immediately 5' of the cleavage site (U-1), which yields the highest cleavage rate constant, while changing the identity of this nucleotide slows cleavage, with G-1 showing the slowest cleavage rate constant. Therefore, comparing global dynamics of U-1 and G-1 ribozymes will shed new light on the structure of important features of the active site such as the orientation of the U-turn structural motif and the catalytically essential C75 nucleotide. From such single molecule studies, kinetic information can be calculated (such as discussed in Chapter Two) and a more detailed understanding of the cleavage mechanism will be achieved. Furthermore, the ribozyme could be studied under various environmental conditions such as varying cation concentrations, pH values, and temperatures. In addition, a non-catalytic C75U mutation could be used (generating an already established inactive ribozyme) to observe possible changes to the observed global dynamics of the wild type ribozyme. Kinetic

information obtained in such studies would provide further detail on the role C75 plays in the dynamics of HDV ribozyme catalysis. These further studies will likely dissect the contributions of structural motifs, such as the U-turn motif, to active site architecture and activity of the HDV ribozyme.

5.4 Molecular Spiders

Summary: Spiders walk and their paths can be tracked in real-time

With expertise from such ribozyme studies, several DNA enzymes (DNAzymes) have been assembled into “Molecular Spiders”. Mobility of these nano-assemblies on surfaces relies on the cleavage reaction of DNAzymes, as described in Chapter Four. The design of these Molecular Spiders is thus an excellent example of how basic research on nucleic acid enzymes can facilitate development of novel materials. Spiders were designed to be possibly used as a drug delivery method, in which case they would be required to recognize and obey specific instructions just as aptazymes do.

Although previously performed SPR measurements showed that Spiders cleave substrate DNA molecules on a surface, these measurements alone could not positively identify ‘walking-like’ behavior [5]. In order to visualize such behavior, where the Spiders follow a continuous path along a surface of substrate molecules, fluorescent tracking was employed to record x,y-coordinates of individual Spider molecules over time. After measuring location coordinates of individual Spiders using FIONA-based fitting of imaged point-spread functions and subtracting the stage-drift, it became evident

that Spiders indeed showed movement across substrate surfaces that were able to track to ~20 nm precision.

Molecular Spiders demonstrate external control of movement such as start and stop commands (for example with addition of zinc as 'start' and addition of EDTA (a chelator of zinc) as 'stop'). They also show landscape-based control of movement in that Spiders will only travel where substrate is present. To further illustrate this controlled motion, Spider movement on specifically designed substrate tracks with certain shapes (for example, tracks which are similar to the letter U) has been tested and observed[5].

Future directions for the Spider World

The studies in this dissertation have only addressed one type of Spider, and it would be interesting to explore Spiders with different a number of legs or differently shaped bodies, and to subject them to various environmental conditions. Furthermore, to test the feasibility of Spiders as a drug delivery method, it is important to observe Spiders as they carry a load (such as a Q-dot or small molecule), which would be necessary if they are to carry drug cargo to a biological target. In addition to testing different Spiders and their environments, the resolution of Spider imaging can also be improved. In order to achieve this, a higher magnification can be used to reduce the pixel size. In addition, the integration time for each frame can be increased to improve the number of photons received, although this can contribute to early photobleaching of fluorophores. On the other hand, Spiders labeled with a greater number of fluorophores, or labeled with a brighter fluorophore could be used to achieve a higher photon count.

[5] Unpublished data from Dr. Anthony J. Manzo, Alexander Johnson-Buck and Nicole Michelotti.

Further studies of Spider behavior would assist in the future design of autonomous Spider-based nano-machines, which could be capable of tasks such as releasing constant amounts of products such as drug molecules as well as nano-patterning of conductive materials for electrical circuits. Additional studies such as Spider races where either the same kind of Spider or different varieties of Spiders move from a start position to an end position with a trigger to 'start' to the race (such as addition of Zn^{2+}) could be performed. A Spider moving through different hurdles to get to a final location is yet another test that will provide further insight into the characteristics of molecular movers and their ability to traverse more complex environments which will be important for future use in biological and medicinal applications.

5.5 Conclusions and overall outlook

This dissertation has focused on applying steady-state and single molecule fluorescence spectroscopy to basic nucleic acid enzyme research as well as to the understanding of engineered devices that employ nucleic acid enzymes. Studies of an engineered hammerhead aptazyme have provided insight into its shortcomings compared to a naturally occurring riboswitch. Now it is clear that more stringent selection methods should be applied to future aptazymes in order to obtain better-performing biosensors. Additionally, protocols and optimizations developed for assembling the HDV ribozyme for smFRET studies contributed to the field by providing a blueprint for successful assembly of longer RNA molecules, and by testing the placement of fluorophores on the ribozyme for fluorescence studies. With this information, future studies can be performed

on the HDV ribozyme to fully understand the characteristics, behavior and functional role of the U-turn and the catalytically assisting nucleotides at the cleavage site. Finally, imaging Molecular Spiders has shed new light on molecular robots and their behavior on traversing a specific landscape. These studies can be used to improve future Spider designs, enabling them to perform specific tasks (with specific speeds, directions and behaviors), as well as to aid development of other nano-devices.

The three systems studied here were observed at the single molecule level, allowing the characteristics of individual molecules to be identified and analyzed. In each case, heterogeneous behavior was observed, with implications for the function and efficiency of each system. Single molecule fluorescence imaging is a powerful tool that here was used successfully to investigate unique structural and mechanistic details about a variety of nucleic acid molecules and complexes. The work presented in this dissertation suggests that single-molecule fluorescence imaging can greatly contribute to the biophysical understanding of biological systems with applications from therapeutics to engineering complex molecular machines.

5.6 References

1. Sekella, P.T., D. Rueda, and N.G. Walter, *A biosensor for theophylline based on fluorescence detection of ligand-induced hammerhead ribozyme cleavage.* Rna, 2002. **8**(10): p. 1242-52.

2. Hendeles, L. and M. Weinberger, *Theophylline. A "state of the art" review.* Pharmacotherapy, 1983. **3**(1): p. 2-44.

3. de Silva, C. and N.G. Walter, *Leakage and slow allostery limit performance of single drug-sensing aptazyme molecules based on the hammerhead ribozyme.* RNA, 2009. **15**(1): p. 76-84.

4. Miyazawa, Y., L. Paul Starkey, A. Forrest, J.J. Schentag, H. Kamimura, H. Swarz, and Y. Ito, *Effects of the concomitant administration of tamsulosin (0.8 mg) on the pharmacokinetic and safety profile of intravenous digoxin (Lanoxin) in normal healthy subjects: a placebo-controlled evaluation.* J Clin Pharm Ther, 2002. **27**(1): p. 13-9.

5. Pei, R., S.K. Taylor, D. Stefanovic, S. Rudchenko, T.E. Mitchell, and M.N. Stojanovic, *Behavior of polycatalytic assemblies in a substrate-displaying matrix.* J Am Chem Soc, 2006. **128**(39): p. 12693-9.

APPENDIX ONE

SUPPLEMENTARY MATERIAL TO CHAPTER TWO

Error analysis

Fluorescence changes upon addition of 10 mM theophylline to 2-AP labeled hammerhead aptazymes is shown in Figure 2.7. Here, $Th_{1/2}$ is calculated for aptamer and core domains with theophylline titrations. The $Th_{1/2}$ that was calculated for the core domain gives a value of 1.65±0.78 mM (Figure 2.7) using Microcal Origin 7.0 software. Here, the mean k_{obs} for each theophylline concentration are taken in to account while data-fitting. However, due to rather large experimentally derived error bars, some k_{obs} values were weighted more (ex: data at 5 mM theophylline), and some were weighted less (ex: data at 0.5 mM theophylline). Therefore, the $Th_{1/2}$ value and its error calculated only depending on the mean values is not accurate (Figure A1.1 – blue curve). To alleviate this issue, the mean values were fitted weighting their respective errors (Figure A1.1 – red curve) using the weighting option of non-linear curve fitting in the software program. This gives a more accurate $Th_{1/2}$ value of 1.47±0.47 mM. Since the two values are within close proximity, the corrected value does not change the conclusions of Chapter Two.

Amplitude correction calculations for Figure 2.6 (bottom panel)

The relative fractional amplitudes of transitions were calculated from the exponential fits to the cumulative dwell time histograms as shown in Chapter Two. These fractions (*f*, which add up to unity) were calculated using:

$$f = \frac{A_{H\to L}\left(\left(\frac{1}{k_{H\to L}}\right)+\left(\frac{1}{k_{L\to H}}\right)\right)}{\sum_{k=1}^{2} A_{H\to L}\left(\left(\frac{1}{k_{H\to L}}\right)+\left(\frac{1}{k_{L\to H}}\right)\right)}$$

Amplitude corrections were calculated by using k_{obs} (Figure 2.6 (top panel) - corrected for observation time window and photobleaching) using the above equation as shown below (plotted in Figure 2.6 – bottom panel).

<u>0 mM Theophylline</u>

Using 0.31/s as the H → L rate:

$$f\_1 = \left(\frac{572\left(\left(\frac{1}{0.31}\right)+\left(\frac{1}{0.65}\right)\right)}{572\left(\left(\frac{1}{0.31}\right)+\left(\frac{1}{0.65}\right)\right)+24.5\left(\left(\frac{1}{0.31}\right)+\left(\frac{1}{0.008}\right)\right)}\right) = \frac{2725}{2725+3141} = \mathbf{0.46\ H\to L}$$

$f\_2 =$ **0.54 L→H**

Using 0.1/s as the H→L rate:

$$f\_1 = \left(\frac{572\left(\left(\frac{1}{0.1}\right)+\left(\frac{1}{0.65}\right)\right)}{572\left(\left(\frac{1}{0.1}\right)+\left(\frac{1}{0.65}\right)\right)+24.5\left(\left(\frac{1}{0.1}\right)+\left(\frac{1}{0.008}\right)\right)}\right) = \frac{6600}{6600+3307} = \mathbf{0.66\ H\to L}$$

$f\_2 =$ **0.33 L→H**

0.5 mM Theophylline

Using 0.38/s as the H → L rate:

$$f\_1 = \left(\frac{449.5\left(\left(\frac{1}{0.38}\right)+\left(\frac{1}{0.42}\right)\right)}{449.5\left(\left(\frac{1}{0.38}\right)+\left(\frac{1}{0.42}\right)\right)+25\left(\left(\frac{1}{0.51}\right)+\left(\frac{1}{0.0058}\right)\right)}\right) = \frac{2228}{2228+6659} = \mathbf{0.25\ H{\rightarrow}L}$$

$f\_2 =$ **0.75 L→H**

Using 0.12/s as the H→L rate:

$$f\_1 = \left(\frac{449.5\left(\left(\frac{1}{0.12}\right)+\left(\frac{1}{0.42}\right)\right)}{449.5\left(\left(\frac{1}{0.12}\right)+\left(\frac{1}{0.42}\right)\right)+25\left(\left(\frac{1}{0.12}\right)+\left(\frac{1}{0.0058}\right)\right)}\right) = \frac{4791}{4791+6787} = \mathbf{0.41\ H{\rightarrow}L}$$

$f\_2 =$ **0.59 L→H**

2.5 mM Theophylline

Using 1.13/s as the H → L rate:

$$f\_1 = \left(\frac{240\left(\left(\frac{1}{1.13}\right)+\left(\frac{1}{0.4}\right)\right)}{240\left(\left(\frac{1}{1.13}\right)+\left(\frac{1}{0.4}\right)\right)+32.4\left(\left(\frac{1}{1.13}\right)+\left(\frac{1}{0.0081}\right)\right)}\right) = \frac{812}{812+3979} = \mathbf{0.16\ H{\rightarrow}L}$$

$f\_2 =$ **0.84 L→H**

Using 0.2/s as the H→L rate:

$$f\_1 = \left(\frac{240\left(\left(\frac{1}{0.2}\right)+\left(\frac{1}{0.4}\right)\right)}{240\left(\left(\frac{1}{0.2}\right)+\left(\frac{1}{0.4}\right)\right)+32.4\left(\left(\frac{1}{0.2}\right)+\left(\frac{1}{0.0081}\right)\right)}\right) = \frac{1800}{1800+4113} = \mathbf{0.30\ H{\rightarrow}L}$$

$f\_2 =$ **0.70 L→H**

10 mM Theophylline

Using 0.21/s as the H → L rate:

$$f\_1 = \left(\frac{603\left(\left(\frac{1}{0.21}\right)+\left(\frac{1}{0.95}\right)\right)}{603\left(\left(\frac{1}{0.21}\right)+\left(\frac{1}{0.95}\right)\right)+40\left(\left(\frac{1}{0.21}\right)+\left(\frac{1}{0.0007}\right)\right)}\right) = \frac{3519}{3519+57333} = \mathbf{0.05\ H{\rightarrow}L}$$

$f\_2 =$ **0.95 L→H**

Using 0.05/s as the H→L rate:

$$f\_1 = \left(\frac{603\left(\left(\frac{1}{0.05}\right)+\left(\frac{1}{0.95}\right)\right)}{603\left(\left(\frac{1}{0.05}\right)+\left(\frac{1}{0.95}\right)\right)+40\left(\left(\frac{1}{0.05}\right)+\left(\frac{1}{0.0007}\right)\right)}\right) = \frac{12708}{12708+57942} = \mathbf{0.17\ H\rightarrow L}$$

$f\_2 = \mathbf{0.83\ L\rightarrow H}$

These L→H and H→L values were plotted in Figure 2.6 – bottom panel and fitted with a modified Hill equation to obtain $Th_{1/2}$ values as discussed in Chapter Two.

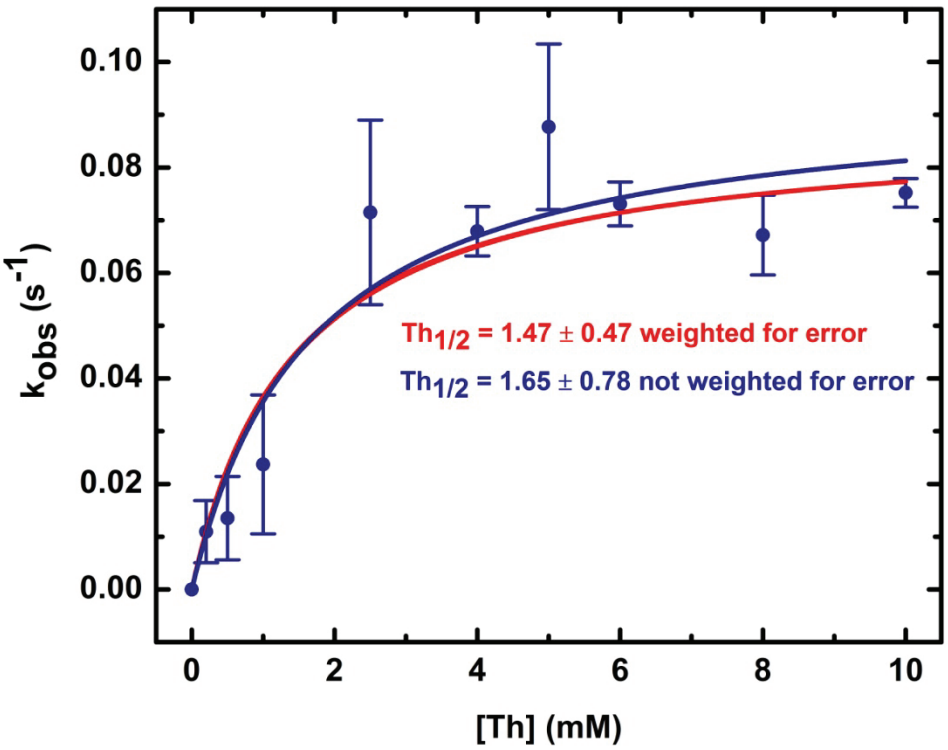


Figure A1.1 Weighted error analysis for the theophylline titration for the core domain of the hammerhead aptazyme.

APPENDIX TWO

SUPPLEMENTARY MATERIAL TO CHAPTER FOUR[6]

Choice of fluorophores on Spiders for fluorescence tracking

Ensemble steady-state FRET assays were done with the Tetra Methyl Rhodamine (TMR) labeled Spider (With four identical catalytic legs). Although the fluorescence signal increased when Zn^{2+} was added and decreased when EDTA was added, after a control experiment, it was found that Zn^{2+} had this effect on TMR irrespective of spiders and substrates. Subsequently, Cy3 was used as a fluorophore. With studies on Cy3B that claim this fluorophore is brighter than Cy3, Cy3B was used in all tracking experiments detailed in Chapter Four.

Cleavage assays of substrates

In order to confirm the catalytic abilities of substrates and Spider legs in the presence of Zn^{2+} and EDTA, cleavage assays were performed and visualized on a

[6]Choice of fluorophores and steady-state assays were performed by Chamaree de Silva, cleavage assays were performed by Dr. Anthony J. Manzo and Chamaree de Silva, substrate density experiments performed by Chamaree de Silva, Dr. Anthony J. Manzo and Xiaomu Guan, stage drift identification, slide preparation protocol and MatLab software for Gaussian fittings by Dr. Anthony J. Manzo and Matlab software for distances and velocities provided by Alexander Johnson-Buck.

polyacrylamide gel. It was shown that substrate is cleaved in the presence of various concentrations of Zn^{2+} and that the presence of EDTA stops cleavage.

Obtaining a suitable substrate density

Initially, a biotin-streptavidin surface to immobilize substrates was used similar to that of smFRET studies. By using a small fraction of substrates labeled with a fluorophore mixed with unlabeled substrate, and counting the number of visible (labeled) substrates on a field of view assuming homogeneous distribution of substrates, the total number of labeled and unlabeled substrates in a certain area was extrapolated. This was attempted with various substrate concentrations as well as various incubation time biotin, streptavidin and substrates. The closest distance between two substrates obtained by this method was about 100 nm, which is quite large for Spiders with a leg span of ~ 25 nm. It was clear that this method would not produce a suitable substrate density for Spider movement.

A commercially available slide preparation from Xantec Corporation (Germany) was attempted where quartz slides were sent to the company to coat with streptavidin. Xantec specified that their streptavidin is a covalent attachment to aldehyde activated carboxyl silane monolayer, resulting in an immobilization capacity of ~1 ng streptavidin/mm². This theoretically gives a value of ~ 10 nm distance between substrates, but it could not be achieved experimentally.

Silanization of slides to improve substrate density was attempted using an avidin surface instead of streptavidin. This method proved to be successful in obtaining a

substrate density for Spider movement. As stated above, both substrate and the Cy3-labeled substrate (non-cleavable) were loaded on the slide with a certain ratio between the two in order to calculate the substrate density.

First, the labeling efficiency of the Cy3 substrate was determined through UV-Vis spectrometry. The efficiency is important to correctly calculate the substrate density. Two absorption peaks were shown within the range of 200-800 nm. The sample is 1:1600 dilution in the cuvette (Figure A2.1). Based on Lambert-Beer law, the concentrations of both substrate and Cy3 were calculated, and the labeling ratio was obtained to be ~ 1:1.

Two samples were tested on the prepared avidin slides. Samples were prepared in standard buffer (10mM HEPES, 150mM NaCl, pH=7.0) and oxygen scavenger system (OSS). OSS was prepared by mixing "50uL" glucose oxidase with 100 µL T50 (50 mM Tris-HCl, 50 mM NaCl, pH=7.5). Afterwards, 12.5uL catalase was added to the mixture. 4 µL of this stock solution was mixed with 8 µL BME and 388 µL glucose (0.5556g D-(+)-glucose anhydrous dissolved in 5 mL buffer) to get the final diluted OSS.

Sample 1: 10 µM unlabelled substrate and 10 pM Cy3-substrate
Sample 2: 10 µM unlabelled substrate and 50 pM Cy3-substrate

Standard buffer was injected into the flow channel of the slide and incubated for 10 min to hydrate the surface. Then 80 µL of the sample was flown into the channel and incubated for 30min. Additionally, excess substrate was washed out with buffer multiple times. Finally OSS was injected into the channel and slides were imaged with TIRFM.

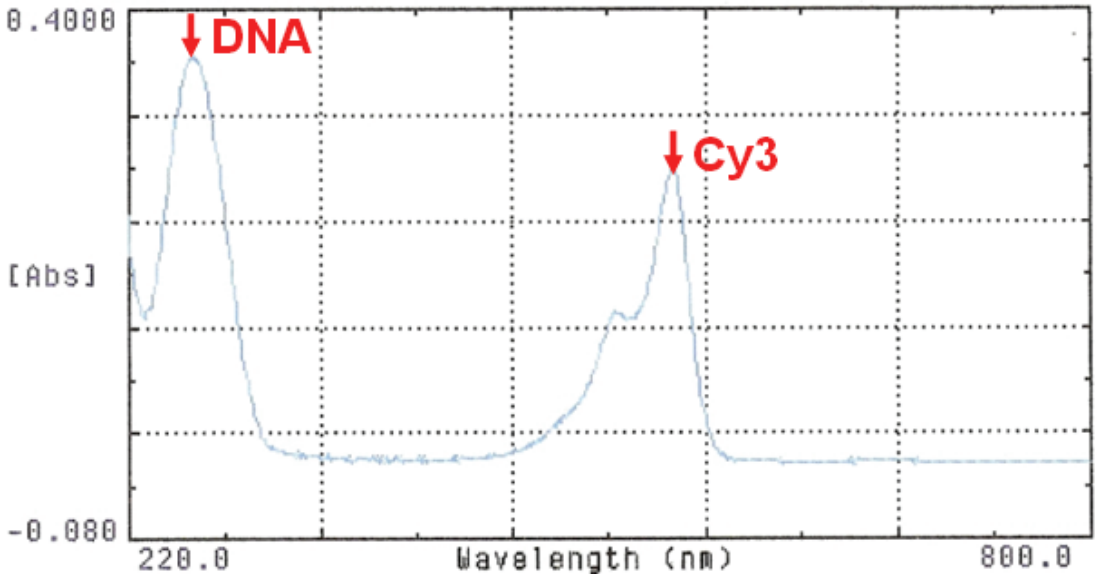


Figure A2.1 UV-Vis spectrum of the Cy3-labeled substrate.

Since the emission wavelength of Cy3 is lower than 610 nm, they can be observed in the donor channel (left). The spots in the right channel are due to background noise, which was subtracted using control experiments where slides were incubated with only unlabeled substrate. As expected, the slide incubated with Sample Two has more Cy3 fluorophores than Sample One (Figure A2.2).

DiaTrack 3.0 software was used to count the number of labeled substrates in a given field of view (Figure A2.3). The decay of the Cy3 fluorophore should be a first order reaction, therefore data points were fit with the equation $N = N_0 + dN \exp(-kt)$. The parameters were obtained, and the rate constant k was within the same order of magnitude for both samples. When data is extrapolates to t = 0, the initial number of particles for both slides can be obtained.

For the slide with 10 pM sample, N = 156, and for the slide with 50pM sample, N=869, which is approximately 5 times than that of 10 pM sample as expected. Using these values, the substrate density was calculated:

Sample One:

[156(10 μM/10 pM)] / [(256)(133 nm)((512)(133nm))] ~ 7 particles/100nm^2

Sample Two:

[869(10 μM/50 pM) / [(256)(133 nm)((512)(133 nm))] ~ 7 particles/100nm^2

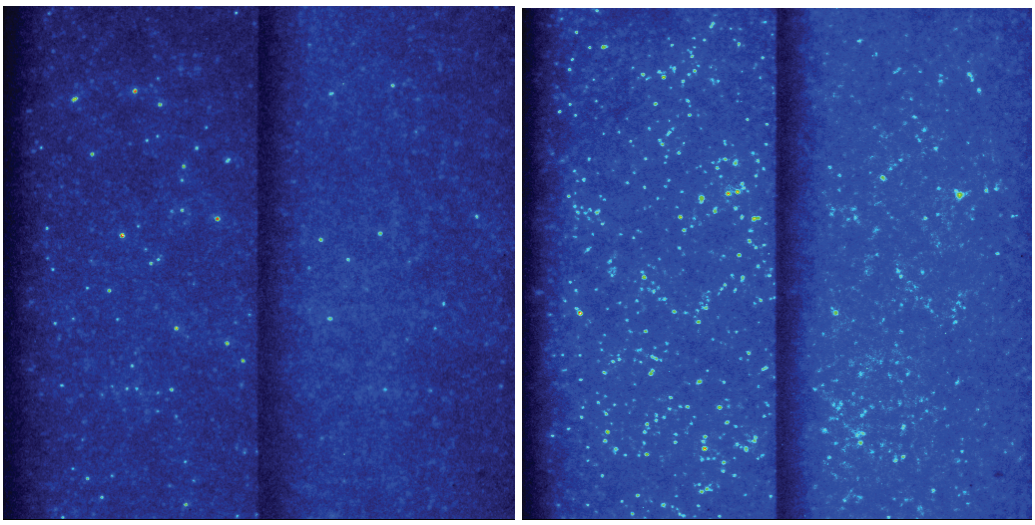


Figure A2.2 Substrate densities of (A) Sample One and (B) Sample Two.

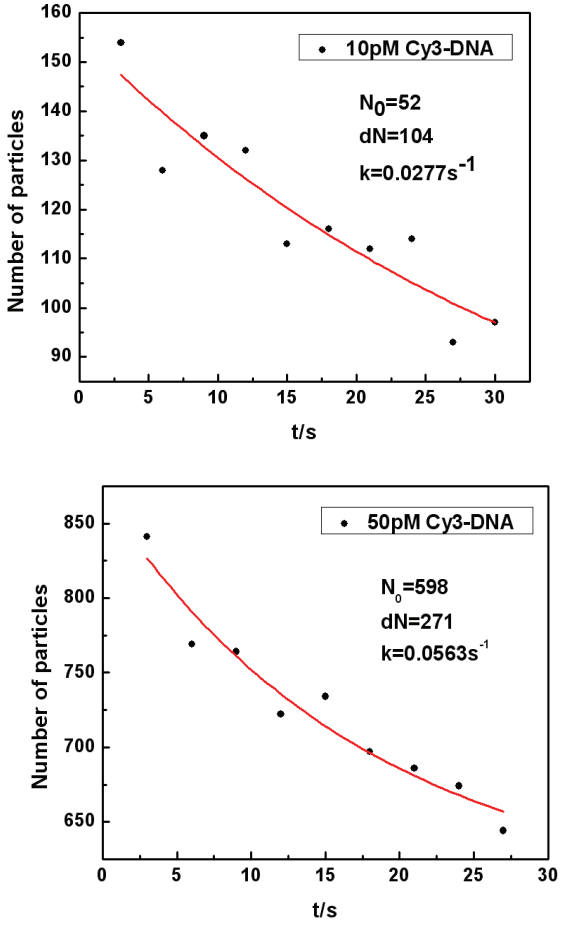


Figure A2.3 First order decay of the number of particles in (A) Sample One (B) Sample Two.

Therefore, the results are in good agreement with each other, and with further estimations it is clear that the distance between two substrates is about 3-5 nm, which is suitable for Spiders with 25 nm leg span.

Final slide preparation protocol

Cleaning slides

A pre-wash solution was made with 100mL autoclaved water, 20mL 35% ammonium hydroxide and 20mL hydrogen peroxide in 150mL beaker in the hood. The solution was heated on a hot plate in the hood for 20min covered, and stirred slowly with a stir bar. The slides were all immersed in the solution. Then the slides were rinsed with water, and put in 1M KOH, sonicated for 45 min. And then they were rinsed with water again. Finally the slides were flamed for at least 1min in the hood.

Silanizing slides

Preparing the silanization solution: 57mL acetone was mixed with 3mL APTES (3-Aminopropyltriethoxysilane) to obtain a 5% solution. The slides were immersed in 5% APTES solution for 1 hour at room temperature, and then both the container and the slides were washed by acetone, and dried for 1 hour at 80°C.

Attaching linkers

Preparing the PDC solution: 60mL DMF was mixed with 6.66mL pyridine, and 0.2g PDC (para-diisothiocyanatobenzene) was added to get a 0.2% solution. The slides were incubated in the PDC solution for 2 hours at room temperature. At the end they were washed with methanol twice and acetone once.

Attaching avidin

Some water was put at the bottom of the pipette box to make a humid environment. The slides were placed in the pipette box, and 70uL avidin solution was added at the center of each slide. The cover slips were carefully placed on top of the slides with avidin solution without making any bubbles. The slides were incubated for 2 hours at room temperature (Figure A2.4).

Preparing the flow channel

The slides were washed with water thoroughly and rinsed with 1 M NaCl + 40 mM NaOH for 1min. Then they were rinsed with deionized water again and dried with N_2. After that the tape was stuck onto the slides making a channel of 1.5-3 mm wide, and the cover slips were placed on the tape. Finally epoxy glue was used to seal the cover slips. Sliders were stored in 4˚C under dry vacuum conditions until use.

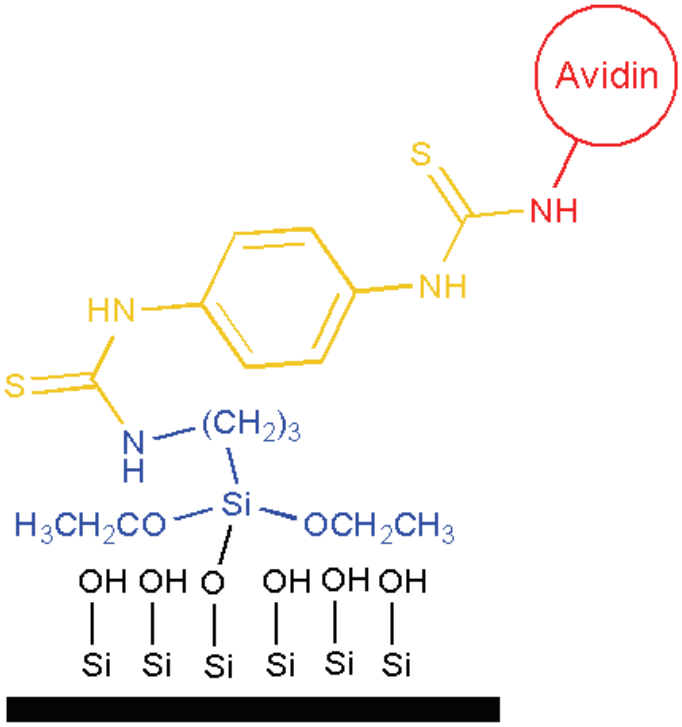


Figure A2.4 Slide preperation for Spider experiments.

Affect of BME on Spiders

It was later discovered that addition of BME completely inhibited Spider movement by SPR studies. Preliminary studies did not show any Spider movement, and BME included in OSS could have been the reason. Therefore, a different OSS which does not include BME (Chapter Four – Materials and Methods) was used.

Stage Drift control

By tracking (x,y) coordinates of Spiders labeled with Q-dots, it was clear that the stage-drift could be larger than the net Spider movement (Figure A2.5). Here, all three trajectories move in a similar fashion. Therefore, stationary particles were used as a drift correction mechanism in proceeding experiments as explained in Chapter Four.

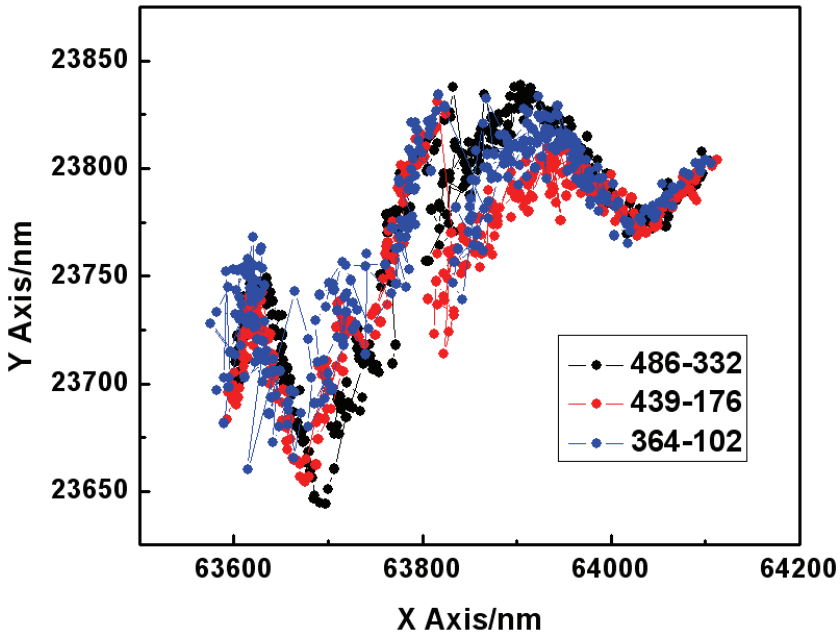


Figure A2.5 Movement patterns of three Spiders before drift correction.

MatLab software for fitting Gaussian functions to PSFs of Spiders

Program "Tony_simplegift_tif_070503_findpeaks_v1p3.m"

```
%pbfitframe centroid-determination algorithm.  No distribution without
%permission from Prof. Sunney Xie, Harvard University.  Contact:
%xie@chemistry.harvard.edu.  Technical inquiries to Paul Blainey,
%blainey@fas.harvard.edu.  February 28, 2005

% Here it is modified by AJM, 060423

%This function fits a signal in an image with a 2-dimensional gaussian,
%reporting the fit parameters to the file xy.txt (description of output
%data below)

clear all;
close all;
close all hidden;
opengl neverselect;
% Input data file (.spe format)
g = ('C:\My Documents\Spider_4Cy3B_3+1\123108\123108_2.SPE');
nametoopen = g;

c=196 % for 60X
number_iterations=1; %number of iterations for search routine
fb_dim=1000; %fit box dimension (nm)
hold off;
format short;

%Set inital parameter values here
guessintensity = 2000; % counts
guesswidthx = 160/c;
guesswidthy = 160/c;    % pixels
guessoffset = 500;     % counts

% .spe reading routine added here, April 23, 2006 AJM
filename=g;
fh = fopen (filename, 'r');
d = fread (fh, 5, 'uint16');
exposure = fread (fh, 1, 'float');
fseek(fh,0,-1);
d = fread (fh, 2050, 'uint16');
xdim = d(22);
ydim = d(329);
```

```
dtype = d(110);
version = d(345);
frames = d(724) + 65536*d(725);

idata = fread(fh, xdim*ydim, 'uint16');
adata = zeros (xdim*ydim,1);
adata=idata;
adata = reshape (adata, xdim, ydim);
adatatemp = adata';
for jjj=1:512;
    for kkk=1:512;
        adata(jjj,kkk)=adatatemp(513-jjj,kkk);
    end;
end;

clear idata;
[A]=adata;
imnum=2; %reference for current image number
amax=max(A);
amin=min(A);
disp(amax);
disp(amin);

imindexval = frames;
NTOT=zeros(1,imindexval);
imindexval = 1;
        hold off;
        contour (A(:,:,1),4); axis equal; % Show image of 1st frame.

        disp('--->  select x0, y0')

        [X,Y] = ginput (1);

        xx = round(X);
        yy = round(Y);
        fitboxdimension = fb_dim/c;
        R = fitboxdimension/2;
        r = round(R);

        B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
        [x,y] = meshgrid(-r:r,-r:r);
        surf(x,y,B(:,:));

        [X,Y,button] = ginput(1);
        disp([button]);
 while ([button] ~= 115);
        disp('In loop');
        [X,Y,button] = ginput(1);
        method = [button];
switch lower(method);
   case {30}; % up arrow key
      disp('up');
      yy=yy-1;
```

```
   if ((yy-r) < 1) %Check to see if outside image dimensions.
       yy=yy+1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case {114}; % r key
   disp('up');
   yy=yy-1;
   if ((yy-r) < 1) %Check to see if outside image dimensions.
       yy=yy+1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case {28}; % left arrow key
   disp('left');
   xx=xx+1;
   if ((xx+r) > xdim) %Check to see if outside image dimensions.
       xx=xx-1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case {100}; % d key
   disp('left');
   xx=xx+1;
   if ((xx+r) > xdim) %Check to see if outside image dimensions.
       xx=xx-1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case {29}; % right arrow key
   disp('right');
   xx=xx-1;
   if ((xx-r) < 1) %Check to see if outside image dimensions.
       xx=xx+1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case {102}; % f key
   disp('right');
   xx=xx-1;
   if ((xx-r) < 1) %Check to see if outside image dimensions.
       xx=xx+1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case  {31}; % down arrow key
   disp('down');
   yy=yy+1;
   if ((yy+r) > ydim) %Check to see if outside image dimensions.
       yy=yy-1;
   end
   B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
   surf(x,y,B(:,:));
case  {99}; % c key
```

```
    disp('down');
    yy=yy+1;
    if ((yy+r) > ydim) %Check to see if outside image dimensions.
        yy=yy-1;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
 case {56}; % up key keypad
    disp('up 10');
    yy=yy-10;
    if ((yy-r) < 1) %Check to see if outside image dimensions.
        yy=yy+10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
case {105}; % i key
    disp('up 10');
    yy=yy-10;
    if ((yy-r) < 1) %Check to see if outside image dimensions.
        yy=yy+10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
 case {52}; % left arrow keypad
    disp('left 10');
    xx=xx+10;
    if ((xx+r) > xdim) %Check to see if outside image dimensions.
        xx=xx-10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
case {106}; % j key
    disp('left 10');
    xx=xx+10;
    if ((xx+r) > xdim) %Check to see if outside image dimensions.
        xx=xx-10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
 case {54}; % right arrow keypad
    disp('right 10');
    xx=xx-10;
    if ((xx-r) < 1) %Check to see if outside image dimensions.
        xx=xx+10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
case {107}; % k key
    disp('right 10');
    xx=xx-10;
    if ((xx-r) < 1) %Check to see if outside image dimensions.
        xx=xx+10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
```

```
    surf(x,y,B(:,:));
 case  {50}; % down arrow keypad
    disp('down 10');
    yy=yy+10;
     if ((yy+r) > ydim) %Check to see if outside image dimensions.
        yy=yy-10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
case  {109}; % m key
    disp('down 10');
    yy=yy+10;
     if ((yy+r) > ydim) %Check to see if outside image dimensions.
        yy=yy-10;
    end
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));
case  {43}; % - key
  disp([button]);
  fitboxdimension = fitboxdimension-2;
  if(fitboxdimension <= 2)
      fitboxdimension=2;
  end
  R = fitboxdimension/2;
  r = round(R);
  if (((yy-r) < 1) | ((yy+r) > ydim)) %Check to see if insode image.
      fitboxdimension = fitboxdimension+2;
      R = fitboxdimension/2;
      r = round(R);
  end
  [x,y] = meshgrid(-r:r,-r:r);
  B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
  surf(x,y,B(:,:));
case  {45}; % + key
  disp([button]);
  fitboxdimension = fitboxdimension+2;
  R = fitboxdimension/2;
  r = round(R);
  if (((yy-r) < 1) | ((yy+r) > ydim)) %Check to see if insode image.
      fitboxdimension = fitboxdimension-2;
      R = fitboxdimension/2;
      r = round(R);
  end
  [x,y] = meshgrid(-r:r,-r:r);
  B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
  surf(x,y,B(:,:));

case {3}; % right mouse button
    disp([button]);
    disp([X,Y]);
    oldfitboxdimension=fitboxdimension;
    fitboxdimension = fitboxdimension+4;
       if (fitboxdimension >= 44)
       fitboxdimension = 8;
```

```
 end

  R = fitboxdimension/2;
  r = round(R);

    if(yy-r < 0)
       fitboxdimension=oldfitboxdimension;
    elseif(yy+r > 512)
       fitboxdimension=oldfitboxdimension;
    elseif(xx-r < 0)
       fitboxdimension=oldfitboxdimension;
    elseif(xx+r > 512)
       fitboxdimension=oldfitboxdimension;
    end
  R = fitboxdimension/2;
  r = round(R);

   disp(r);
[x,y] = meshgrid(-r:r,-r:r);
disp([xx,yy]);
disp([x,y]);
B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
surf(x,y,B(:,:));
 case{1}; % left mouse button
    disp([button]);
    xold = xx;
    yold = yy;

    if (round(X) < 0)
    xx = xx - 4;
    end
    if( round(X) > 0)
    xx = xx + 4;
    end
    if(round(Y) < 0)
    yy = yy - 4;
    end
    if(round(Y) > 0)
    yy = yy + 4;
    end
        if(yy-r < 0)
        yy=yold;
    elseif(yy+r > 512)
        yy=yold;
    elseif(xx-r < 0)
        xx=xold;
    elseif(xx+r > 512)
        xx=xold;
    end
  B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
surf(x,y,B(:,:));

    case  {110}; % n key
```

```
       imnum=imnum+1;
    if(imnum > imindexval)
        imnum=imnum-1;
    end
     fseek(fh, 4100+(imnum-1)*2*xdim*ydim, -1);
    idata = fread(fh,xdim*ydim, 'uint16');
    adata = zeros (xdim*ydim,1);
    adata=idata;
    adata = reshape (adata, xdim, ydim);
    adatatemp = adata';
    for jjj=1:512;
        for kkk=1:512;
            adata(jjj,kkk)=adatatemp(513-jjj,kkk);
        end;
    end;
    clear idata;
    [A]=adata;

    R = fitboxdimension/2;
    r = round(R);
    [x,y] = meshgrid(-r:r,-r:r);
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));

case  {112}; % p key
       imnum=imnum-1;
    if(imnum == 0)
        imnum=imnum+1;
    end
    fseek(fh, 4100+(imnum-1)*2*xdim*ydim, -1);
    idata = fread(fh, xdim*ydim, 'uint16');
    adata = zeros (xdim*ydim,1);
    adata=idata;
    adata = reshape (adata, xdim, ydim);
    adatatemp = adata';
    for jjj=1:512;
        for kkk=1:512;
            adata(jjj,kkk)=adatatemp(513-jjj,kkk);
        end;
    end;
    clear idata;
    [A]=adata;

    R = fitboxdimension/2;
    r = round(R);
    [x,y] = meshgrid(-r:r,-r:r);
    B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
    surf(x,y,B(:,:));

otherwise;
   disp('Unknown key.');
   disp([button]);
end;
```

```
 disp([xx,yy]);
 end;  %End of while
  surfc(x,y,B(:,:));
  pause ; % pause

  for imindex = 1:imindexval;
  fseek(fh, 4100+(imindex-1)*2*xdim*ydim, -1);
  idata = fread(fh, xdim*ydim, 'uint16');
  adata = zeros (xdim*ydim,1);
  adata=idata;
  adata = reshape (adata, xdim, ydim);
  adatatemp = adata';
  for jjj=1:512;
      for kkk=1:512;
          adata(jjj,kkk)=adatatemp(513-jjj,kkk);
      end;
  end;

  clear idata;
  [A]=adata;
  hold off;
  if (imindex == 1
      guessx = 0;
      guessy = 0;
  else
  end;

  guessintensity = max(max(A));
  guesswidthx = 160/c;
  guesswidthy = 160/c;

  guessoffset = mean(mean(A));
  disp('guessoffset is: ');
  disp(guessoffset);

  bestcoeffs = zeros(1,6);
            B = double(A((yy-r):(yy+r),(xx-r):(xx+r)));
contour (B); axis equal;
      [x,y] = meshgrid(-r:r,-r:r);
      coeffs = zeros(number_iterations,6);

          initparams = double(bestcoeffs);
          initparams(1) = guessintensity;
          initparams(2) = guessx;
          initparams(3) = guessy;
          initparams(4) = guesswidthx;
          initparams(5) = guesswidthy;
          initparams(6) = guessoffset;

      for iteration_number = 1:number_iterations;
          best_values = pbgausscomplete51(initparams,x,y,B);

          coeffA = zeros(1,6);
```

```
                coeffA(1,1) = best_values(1,1);
                coeffA(1,2) = best_values(1,2);
                coeffA(1,3) = best_values(1,3);
                coeffA(1,4) = best_values(1,4);
                coeffA(1,5) = best_values(1,5);
                coeffA(1,6) = best_values(1,6);

                guessintensity = coeffA(1,1);
                guessx = coeffA(1,2);
                guessy = coeffA(1,3);
                guesswidthx = coeffA(1,4);
                guesswidthy = coeffA(1,5);
                guessoffset = coeffA(1,6);

                coeffs(iteration_number,1) = guessintensity;
                coeffs(iteration_number,2) = guessx;
                coeffs(iteration_number,3) = guessy;
                coeffs(iteration_number,4) = guesswidthx;
                coeffs(iteration_number,5) = guesswidthy;
                coeffs(iteration_number,6) = guessoffset;

                coeffs(iteration_number,1) =
0.001*coeffs(iteration_number,1);
                coeffs(iteration_number,2) =
c*coeffs(iteration_number,2);
                coeffs(iteration_number,3) =
c*coeffs(iteration_number,3);
                coeffs(iteration_number,4) =
c*coeffs(iteration_number,4);
                coeffs(iteration_number,5) =
c*coeffs(iteration_number,5);
                coeffs(iteration_number,6) =
0.001*coeffs(iteration_number,6);
            end

            bestcoeffs = coeffs(number_iterations,:);
            bestcoeffs(1,1) = 1000 * bestcoeffs(1,1);
            bestcoeffs(1,2) = (1/c) * bestcoeffs(1,2);
            bestcoeffs(1,3) = (1/c) * bestcoeffs(1,3);
            bestcoeffs(1,4) = (1/c) * bestcoeffs(1,4);
            bestcoeffs(1,5) = (1/c) * bestcoeffs(1,5);
            bestcoeffs(1,6) = 1000 * bestcoeffs(1,6);

            surf(x,y,B(:,:));
            [z1,z2] = meshgrid(-r:r,-r:r);
            B_fit = bestcoeffs(1)*exp(-0.5*( (((z1-
bestcoeffs(2)).^2)/bestcoeffs(4).^2) + (((z2-
bestcoeffs(3)).^2)/bestcoeffs(5).^2))) + bestcoeffs(6);
            hold off;
            surf(x,y,B(:,:));
            camlight left; lighting gouraud
            hidden off;
            hold on;
```

```
            surf(z1,z2,B_fit,'FaceColor','yellow','EdgeColor','none');
            camlight left; lighting gouraud
            view(-15,40)
            pause(0.005);

            amax=max(B);
            amax=max(amax);
            peakintensityval(imindex) = amax;

            partsum=sum(sum(B_fit-bestcoeffs(6)));
            NTOT(imindex)=partsum;

            chi=0;
                  for cx=1:(2*r+1);
                for cy=1:(2*r+1);
            chi=chi+((B(cx,cy)-B_fit(cx,cy))^2)/(B_fit(cx,cy));
                end;
            end;
            chi=chi/( (((2*r+1)*(2*r+1))-6));
            chi=chi/6.0;

            disp(imindex);
            disp(chi);

            absolute_bestcoeffs = zeros(1,8);
            absolute_bestcoeffs(1,1:6) = bestcoeffs;
            absolute_bestcoeffs(1,2) = (coeffs(1,2) + (c * xx));
            absolute_bestcoeffs(1,3) = (coeffs(1,3) + (c * yy));
            absolute_bestcoeffs(1,4) = coeffs(1,4);
            absolute_bestcoeffs(1,5) = coeffs(1,5);
            absolute_bestcoeffs(1,7) = (c * xx);
            absolute_bestcoeffs(1,8) = (c * yy);

            xy = absolute_bestcoeffs(1,1:6);
            totxy(imindex,1:6)=xy;

  end
xval = 1:1:imindexval;
xpixel = xx  + round(guessx);
ypixel = yy  + round(guessy);
outfilename = strcat('C:\My Documents\Spider_4Cy3B_3+1\123108\',
'123108_2_', num2str(xpixel),'_',num2str(ypixel),'.dat');

disp(outfilename);
fid = fopen(outfilename, 'wt');
for iterationnum = 1:imindexval;
fprintf(fid, '%6.2f %12.8f %12.8f %12.8f %12.8f %12.8f %12.8f
%12.8f\n', xval(iterationnum), NTOT(iterationnum),
totxy(iterationnum,1:6));
end
fclose(fid)
save;
close;
```

Matlab software for calculating distances and velocities of Spiders

```
%Program for calculating displacement by AJB

close all;
clear all;
rawdata = load('C:\My Documents\Spider\S133403.DAT');
rawframes = size(rawdata, 1);

framerate = input('How many frames per minute?');
   timeBeforeZinc = 0;
if timeBeforeZinc == 0
       initialStart = 1;
       initialEnd = 1;
       for n = 1:rawframes
           rawdata(n,3) = rawdata(n,3)/framerate;
       end
       end
rawdisplacement = zeros(rawframes,3);
rawdisplacement(:,1)=rawdata(:,3);

rawinitialXpt = mean(rawdata(initialStart:initialEnd,1));
rawinitialYpt = mean(rawdata(initialStart:initialEnd,2));
%For displacement from the previous frame
for i = 1:rawframes-1
    rawdisplacement(i,2) = ((rawdata(i+1,1)-
rawdata(i,1))^2+(rawdata(i+1,2)-rawdata(i,2))^2)^(1/2);
end
%For final displacement between first and last frames
finaldisplacement = ((rawdata(rawframes,1)-
rawdata(1,1))^2+(rawdata(rawframes,2)-rawdata(1,2))^2)^(1/2);
%Velocity from frame to frame
for i = 1:rawframes-1
    rawdisplacement(i,3)= rawdisplacement(i,2)/(rawdata(i+1,3)-
rawdata(i,3));
end
framevelocity = mean (rawdisplacement(1:rawframes-1,3));
disp (strcat('frame velocity = ',num2str(framevelocity)));
disp (strcat('final displacement = ', num2str(finaldisplacement)));
totalrawdistance = sum (rawdisplacement(:,2));
disp (strcat('total raw distance = ', num2str(totalrawdistance)));
```