1-1-2011

Laser-assisted single-molecule refolding

Rui Zhao
Wayne State University,

LASER-ASSISTED SINGLE-MOLECULE REFOLDING

by

RUI ZHAO

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2011

MAJOR: CHEMISTRY (Physical chemistry)

Approved by:

_____

Advisor   Date

_____

_____

DEDICATION

This dissertation work is dedicated to my wife Yichun Lin and my

parents.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor Dr.David Rueda for leading me to the single molecule world, for his patience and kindness, for his thought stimulating talk, and for his great mentoring.

Secondary, I would like to thank all the Rueda's Lab members for the great help during the past 5 years. Sanka Atapattu persuaded me to join the Rueda's Lab, and helped me a lot with the courses and the painful cumulative examination. Rajan Lamichhane, Krishanthi Karunatilaka and Zhuojun Guo teached me everything when I learn to work with RNA, DNA and proteins. Sharla wood always remembers our birthdays and brings us a cake. Amanda Solemn and Marcus revised my manuscripts many times. Myles Marshall carried part of the experiments in this thesis. All the lab members helped me in some ways. I also enjoyed the leisure talking during the coffee breaks. They made my 6-year Ph.D. life less stressful and much colorful.

Thirdly, I would like to thank my thesis committee members: Dr. Louis Romano, Dr. Chernyak Vladimir and Dr. Jianjun Wang for the precious time and helpful comments.

Last but not least, I thank all the friends I have met in Detroit, especially those I met in Central Alliance Church. They gave me a lot of help when I first came to Detroit and settle down here. I would like to thank Paster David Chang and his family, Dr. Lei Shen, Dr. Jingsong Wang, Mr. Jun Jiang and his family, Dr. Yucheng Chang and his family, Nurse Huili, Mr. Chengji Chen and his family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF ABBREVIATIONS

RNA	Ribonucleic acid
dsDNA	Double stranded Deoxyribonucleic acid
LASR	Laser-assisted single-molecule refolding
IR	Infrared
T-jump	Temperature jump
PEG	Polyethylene glycol
BSA	Bovine serum albumin
IR	Infrared
OSS	Oxygen scavenger system
CCD	Charge-coupled device
UV	Ultraviolet
FT-ICR	Fourier-transform ion cyclotron resonance
BME	β-mercaptoethanol
SubPop	Subpopulations
EMSA	electrophoretic mobility shift assay
RZ	Two way junction Hairpin ribozyme strand
PDMS	Polydimethylsiloxane
3D	Three-dimensional
EDTA	Ethylenediamine-N,N,N',N'-tetraacetic acid
TCSPC	Time-Correlated Single Photon Counting
HIV DIS	Human immunodeficiency virus dimerization initiation site

FRET	Fluorescence resonance energy transfer
MOPS	3-morpholinopropane-1-sulfonic acid
NMR	Nuclear magnetic resonance
NTP	Nucleoside 5'-triphosphate
PAGE	Polyacrylamide gel electrophoresis
TIRFM	Total internal reflection fluorescence microscopy
TRIS	2-amino-2-(hydroxymethyl)propane-1,3-diol

CHAPTER 1

Introduction

1.1 Non-coding RNAs

RNA molecules are biopolymers with various important biological functions. In the central dogma of molecular biology, RNA transfers the genetic information from DNA to protein. However, there are some transcribed RNAs that are not translated to proteins, namely non-coding RNAs The percentage of non-coding RNA increases from 1%-95% as the complexity of organism increases from Escherichia coli to human[1]. Non-coding RNAs perform important roles in living organisms.

In the 1970s, Thomas R. Cech and Sidney Altman discovered RNAs can act as enzyme just like protein; these RNAs were later referred to as ribozymes[2]. Most of the identified ribozyme (such as the hairpin ribozyme[3], group I ribozyme[2], and possibly the splicesome[4]) are involved in the hydrolysis of the phosphodiester bonds. The hairpin ribozyme is a small RNA that is crucial in the ringspot virus replication process. It cleaves the repeating –strand RNA from viral rolling cycle replication and ligates the unit RNA back to circular RNA for rolling cycle replication of + strand RNA[3]. Group I introns are large self-splicing RNAs found in many organism. They have to splice themselves out of the precursors (mRNA, tRNA or rRNA) to make active RNAs[2]. The ribosome is a megadalton size ribonucleoprotein complex, which makes the protein that is essential for life during translation. Ribosomal RNAs were found to catalyze the aminotransferase reaction to synthesize peptides[5].

Recent advances in RNA interference showed the ability of small RNA molecules (micro-RNA and small interfering RNA) to suppress specific genes, which make it a promising tool in biotechnology and medicine[6,7]. Also, the riboswitches can bind to small target molecules to turn gene expression on or off[8].

RNA molecules can also act as hereditary molecules, which, together with all the functions motioned above, make RNA the most versatile molecule compared to DNA and protein. These observations give rise to the 'RNA world hypothesis' of the origin of life[9].

1.2 RNA folding and hairpin ribozyme

For all the important RNAs mentioned above to play their biological roles, they have to fold into the precise 3-D structures. RNA folding from a linear chain to a functional structure can be challenging given the incredibly large configuration sampling space. It is clear now that RNA folds in a hierarchical way[10,11].
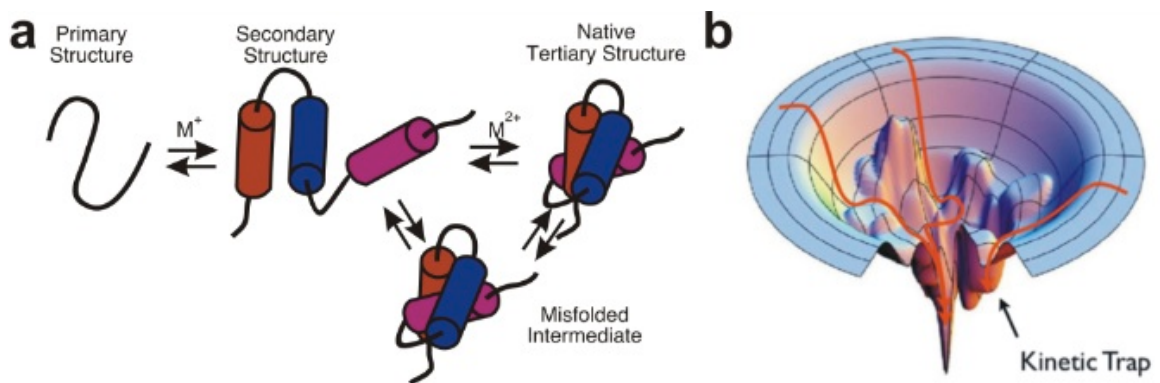
Figure 1.1 RNA folding problem **a.** Schematic plot of the RNA folding hierarchy. **b.** A folding potential energy surface with a kinetic trap. With kind permission from Springer Science+Business Media[10,12].[12]

Firstly, with the help of cation, RNA folds into secondary structure through base paring. Then tertiary interaction can be further formed with the assist of divalent cation, such as Mg2+. Due to the flexibility of RNA molecule, the folding potential energy

surface is particularly rugged and full of kinetic traps[13]. During folding, RNA molecules search along the potential energy surface. When correctly formed, the RNA molecule will find the global minimum, which corresponds to the native structure with biological function. However, sometimes RNA folds with non-native interactions and find a local minimum corresponding to intermediate states that may or may not be active. The conversion from a stable intermediate state to the native state is slow[14,15]. Normally, an enzyme (such as RNA chaperone protein[16,17], helicase[18,19] or cofactor[10,20]) is needed to help the RNA to refold into the native structure. Otherwise, kinetic traps can be deep enough to trap the RNA molecule and prevent further searching for global minimum native state. The RNA molecule will remain trapped in a local minimum, which may be characterized by structural fluctuations that are distinct from the native structure. The entrapped RNA molecule can remain in this characteristic structural dynamic for a long time, which was referred as memory effect in RNA folding. This memory effect has been mainly observed for a subset of ribozyme family, including 2-way minimal form hairpin ribozyme, 4-way junction hairpin ribozyme, and the Tetrahymena group I ribozyme[14,15,21].

1.3 Memory effect in RNA folding

1.3.1. Discovery of Memory effect in RNA folding

The hairpin ribozyme was found in tobacco ringspot virus satellite RNA, where it is involved in the viral rolling cycle replication process. After the repeating – strand RNA is replicated from the rolling cycle replication. The hairpin rirozyme encoded in the – strand RNA will form secondary structure as shown in figure 2 and catalyze a self cleavage reaction to slice the replicated RNA into unit copy. The natural form of hairpin

ribozyme has 4 helical arms, named A, B, C and D (Fig. 1.2a)[22]. From the crystal structure, arm D stacks on arm A and arm C stacks on arm B coaxially (Fig. 1.2b). The stacked arms can form perpendicular or parallel structure (undocked and docked structure) depending on the salt condition[23,24]. The hairpin ribozyme binds substrate in the undocked structure. But in the native form, the substrate is in *cis* and therefore bound during secondary structure formation. Then the hairpin ribozyme can dock and undock several times before the catalysis (Fig 1.2b). When docking, the two loops in arms A and B directly interact to form the catalytic core. Then, the cleavage and ligation reactions can take place. The cleaved hairpin ribozyme can dock and undock for many cycles before the product release or the ligation reaction[23]. The cleavage/ligation equilibrium favors ligation. As a consequence, the net outcome is favoring the formation of the cyclic –strand RNA copy[25].
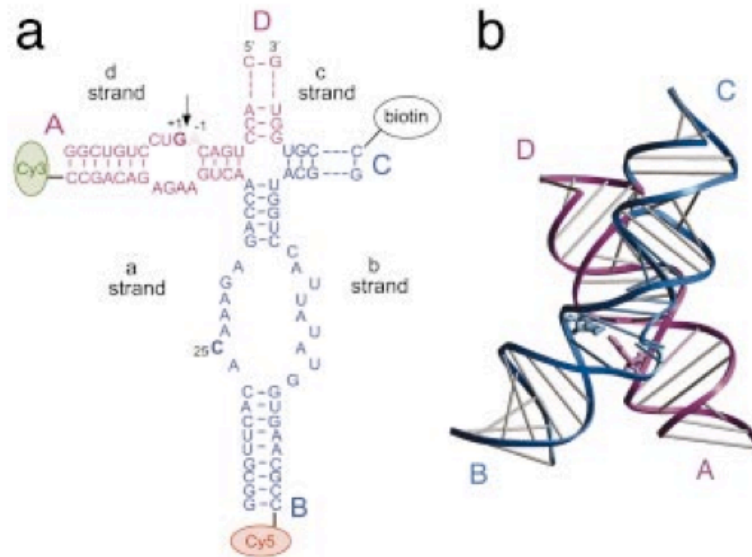
Figure 1.2 Four-way junction hairpin ribozyme. a. Secondary structure of the four-way junction hairpin ribozyme consists of 4 arms: A, B, C and D. Arm A and B have internal loops. The cleavage site is indicated by arrow. Fluorescent labeling and 5' biotin is added as shown for single molecule FRET study. **b.** Crystal structure of the four-way junction hairpin ribozyme shows stacking of arm C on B and arm D on A, respectively. Reprinted from Ref. [26] Copyright (2004) National Academy of Sciences, U.S.A.

The 2-way junction form of the hairpin ribozyme was artificially designed to study RNA enzymology as it can act as a multi-turnover enzyme[24]. In 2-way minimal form hairpin ribozyme, arm C and D were replaced by a C6A linker (Fig. 1.2a). This modification has decreased the docking rate by almost 3 orders of magnitude. Nevertheless, the catalytic core is intact and it performs the cleavage and ligation reaction.

Early ensemble studies identified heterogeneity in ribozyme catalysis: biphasic kinetics with fast and slow populations[27]. But not until Zhuang *et al*'s single molecule work of 2 way junction hairpin ribozyme did people directly obverse the heterogeneity in the structure dynamics, which corresponds to the heterogeneity to its catalysis[14].
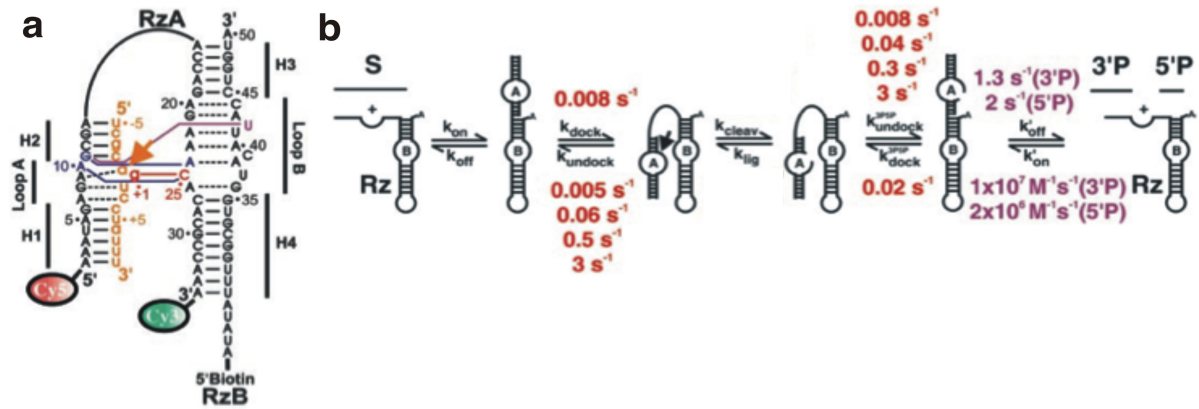
Figure 1.3 Two-way junction hairpin ribozyme. a. Secondary structure of two-way junction hairpin ribozyme. Solid line-Watson-Crick base pairs, dashed line-non-canonical base pairs, Purple line-U42 binding pocket, blue line-ribose zipper, red line-GC base pair. Fluorescent labeling and 5' biotin is added as shown for single molecule FRET study. **b.** Putative reaction pathway of hairpin ribozyme. Kinetic rate constant were calculated using single molecule dwell time analysis[28]. Reprinted with permission from AAAS[14].

At room temperature, docked and undocked structures are in equilibrium in buffered solution (50 mM TRIS-HCl ph 7.5 with 12 mM $MgCl_2$). The molecule can go through several docking and undocking cycles before the cleavage reaction takes place. After cleavage, the undocking rate is barely changed; however the docking rate is accelerated 2-3 fold.

Only one docking rate constant (0.008 s^{-1}) was determined, whereas there are 4 distinct undocking rate constants spanning 3 fold of magnitude (0.005-3 s^{-1}). The molecules tend to undock with the same rate constant for a long time with barely inter-conversion between different kinetic patterns (<5%, in 3 hours); this behavior was referred to as memory effect. This memory effect in the undocking dynamics will not be easily discernable in the ensemble bulking averaged experiment. Further modeling was applied to correlate the memory effect of docking heterogeneity with the cleavage reaction heterogeneity. With the perfect memory of undocking assumption, the model

fits with the experiment data very well. This again supports the notion of heterogeneity and memory effect in the hairpin ribozyme folding.

Similar memory effects were observed in the four-way junction hairpin ribozyme[26]. Compared to the two-way junction, the natural form hairpin ribozyme has four-way junction that allows docking to occur at 2-3 orders of magnitude lower magnesium concentration. In the dwell time scatter plot, both docking and undocking rate constants show heterogeneity with a 50-fold variation.

The stopped-flow kinetics study of 4-way ribozyme also displayed evidence for heterogeneous kinetics rates. At t=0, molecules in undocked structure are mixed with 0.5 mM Mg^{2+} and allowed to fold for various times. Both experimental and simulation results were best fitted by double exponential decay. The apparent kinetic rate constants for folding determined by stopped flow and single-molecule spectroscopy agree very well, varying by only 1.5 fold. For 4-way junction ribozyme, the heterogeneity of undocking persists even after catalysis.
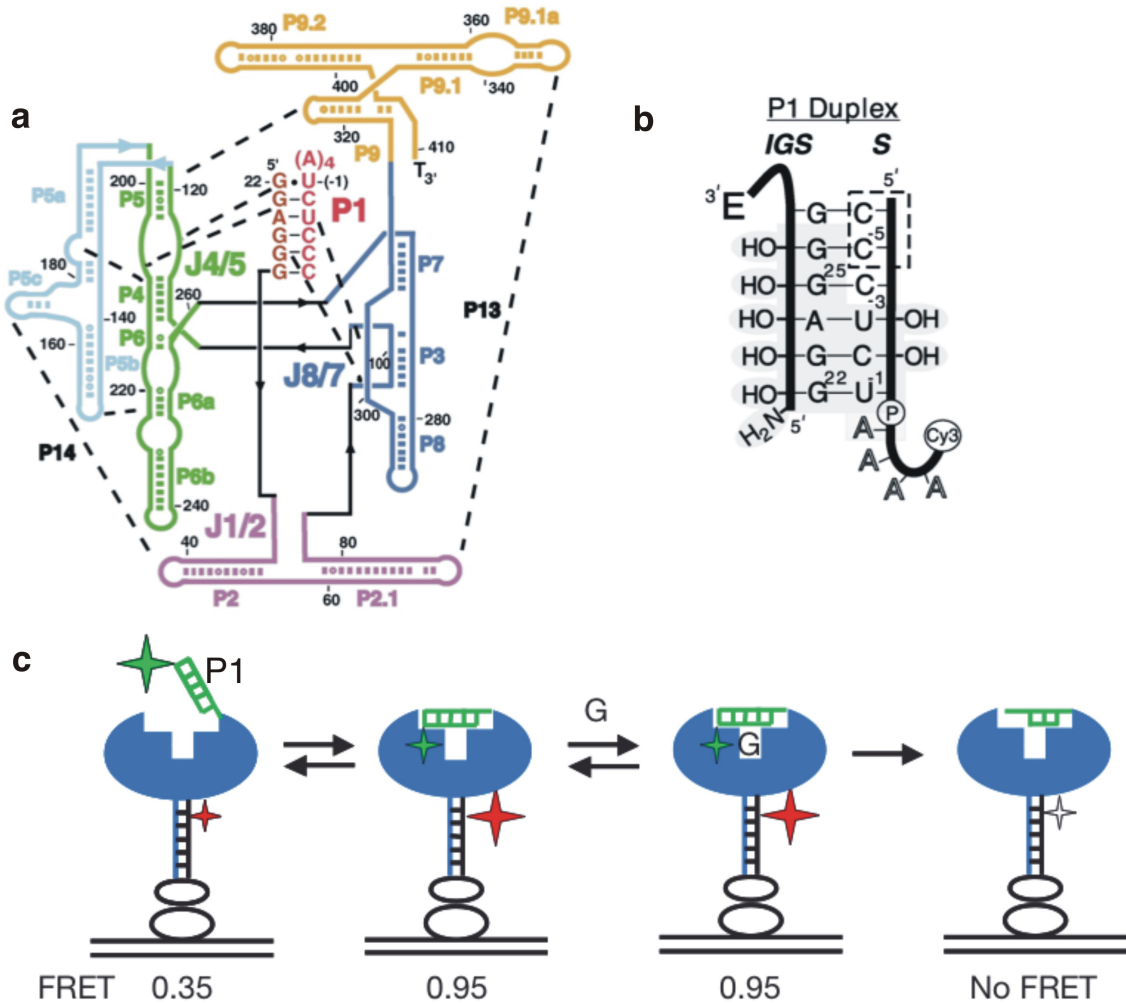
Figure 1.4 Tetrahymena group I intron ribozyme secondary structure and raction pathway. a. Secondary structure of Tetrahymena group 1 ribozyme. **b.** P1 domain with substrate binding. **c.** Tetrahymena group 1 ribozyme docking and cleavage of substrate. For FRET study, Cy3 is attached at 3' end of substrate, Cy5 is attached to 3' end of DNA tether and Biotin is attached to 5' end of DNA tether. Upon the substrate binding, emission from FRET donor will be observed. When the substrate docked into ribozyme, a high FRET will be observed. After cleavage and the product release, both donor and acceptor emission will disappear. Reprinted With permission from Elsevier [29]. Reprinted by permission from Macmillan Publishers Ltd: [Nature] ([15]), copyright (2010).

Heterogeneity in undocking was also observed in Tetrahymena group 1 ribozyme, tetrahymena group I ribozyme, a ~400-nucleotide RNA enzyme that cleave an oligonucleotide substrate. The catalytic core formation consists of two steps: firstly, in the multiple-turnover form of the ribozyme, the single stranded 5' end region of the

ribozyme binds a substrate forming the P1 duplex. Secondly, the P1 duplex docks into the ribozyme to form the active site. Guanosine binding is required to for the cleavage reaction. Single molecule FRET time experiment shows distinct docking /undocking between FRET state 0.35 and 0.95.

Docking equilibrium constants were calculated for each molecule; these K_{dock} varies by a factor of 300. The free energy (ΔG_{dock}) can be calculated as: ΔG_{dock} = RTln(K_{dock}). The ΔG_{dock} distribution of all the molecules is so broad that at least 6 sub-populations were needed to fit the data after correction for limited time windows. Subpopulations in the free energy histogram have the same rate constant for the cleavage reaction. The subpopulations barely inter-convert in 40 minutes (~9%), showing a memory effect in the docking kinetics.

1.3.2. The memory effect in folding is Intrinsic and long lived

1.3.2.1. The memory effect is not a surface immobilization artifact

The agreement of cleavage rate between the single molecule experiment and the bulk solution measurement indicates that surface immobilization did not change the functional properties of the hairpin ribozyme[14]. Hence the structural dynamics is probably not affected by surface immobilization methods (PEG-biotin and BSA-biotin).

To completely exclude the possibility that the heterogeneity of hairpin ribozyme docking and undocking originates from surface immobilization, a new immobilization method utilizing lipid bi-layer vesicle was adopted to eliminate the contact of the ribozyme with surface[30]. The RNA is encapsulated in small unilamellar vesicles (100 or 200 nm) at concentrations that make the encapsulation of multiple RNAs extremely rare. The hairpin ribozyme freely diffuses and folds inside the vesicle, as the diameter of

hairpin ribozyme is about 10-fold smaller. The diffusion coefficient of the encapsulated

RNA is ~10 fold smaller than that of the free ribozyme.
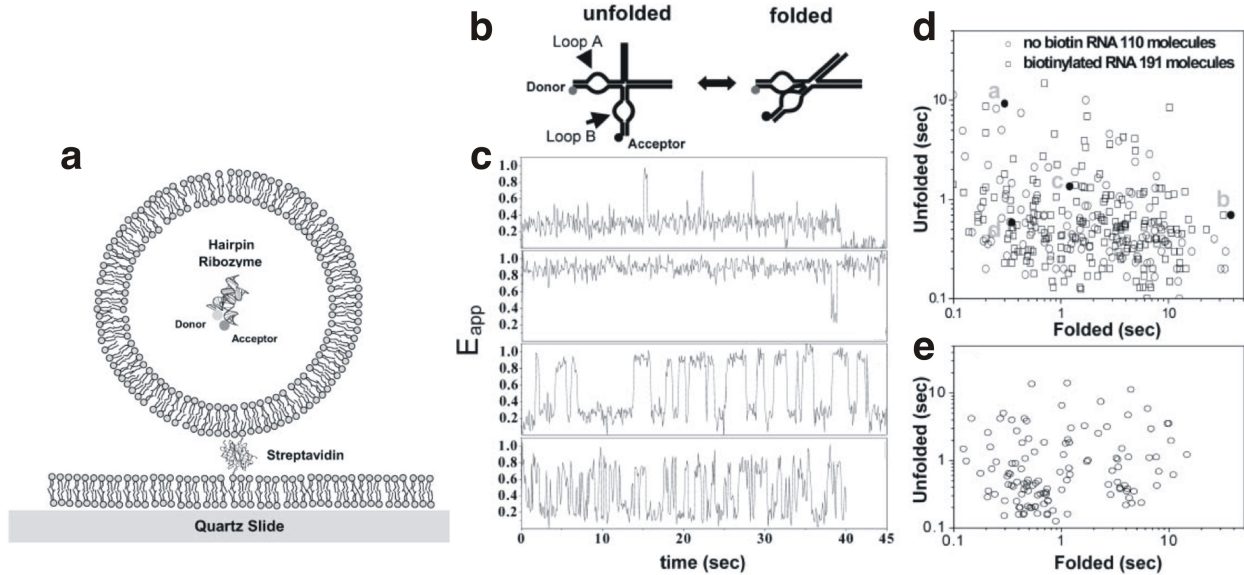
Figure 1.5 Single molecule FRET study of four-way junction hairpin ribozyme in a liposome. a. Drawing is to scale. The lipsome is about 30-nm in diameter. In actual single molecule FRET studies, the size of lipsome is normally ~100 nm or 200 nm. FRET donor and acceptor were labeled as described previously. Immobilization of the lipsome is through streptavidin-biotin interaction. **b.** Illustration of docking kinetic of hairpin ribozyme. **c.** single molecule FRET time courses of hairpin ribozyme. **d.** Scatter plot of averaged dwell time in docked and undocked states. Circle: Non-biotinylated RNA. Square: Biotinylated RNA. Solid dots: Four data points calculated from **c. e.** Scatter plot from surface immobilized RNAs**.** Reprinted from Ref. [26,30] Copyright (2004), with permission from Elsevier.

After the validating the single molecule FRET experiment in a ves[26]icle, they

characterized hairpin ribozyme folding kinetic inside 200-nm vesicles. The single

molecule FRET time course shows same apparent FRET efficiency for the docked and

undocked states and the same heterogeneity in the docking/undocking kinetics. As

seen in single-molecule experiments, the scatter-plot of averaged dwell time in

docked/undocked state shows no difference in the vesicle encapsulated and free

ribozyme molecules. The heterogeneity of the kinetics appears to be persistent throughout the experimental time window (minutes-hours).

These results indicate that the heterogeneity in hairpin ribozyme undocking is not an artifact of surface immobilization. Similar conclusions can be inferred for the 2-way junction hairpin ribozyme and tetrahymena group I ribozyme.

1.3.2.2. The memory effect is not a result of covalent differences

The memory effect could originate from covalent differences in the RNA from synthesis, such as a protection group; however this is unlikely because the memory effect of hairpin ribozyme is independent of synthesis methods: Tert-butyldimethylsilyl and ACE Chemistries. To completely exclude the possibility of covalent differences, high-resolution mass spectroscopy was adopted to measure mass of the subpopulations of hairpin ribozyme two-way junction[31]. High-resolution electrospray ionization hybrid quadrupole-Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry no difference in the major peaks for different populations. There were some small impurity peaks such as RZA without 5'end mono phosphate or 3'-end uridine of RzB, but nothing that could induce the molecular heterogeneity of folding.

1.3.2.3. The memory effect is not a result of secondary structure Difference

Difference is secondary structure could account for the memory effect; however this is extremely unlikely since strand re-annealing and displacement does not completely abrogate the memory effect. Nevertheless, to fully exclude the possibility of secondary structure contribution, a footprinting assay was performed to determine the secondary structure difference between the EMSA separated ribozyme[31]. RNase VI and Tb3+ were to cleavage the ribozyme at the desired sites. RNase VI cleaves RNA at the

double stranded regions. Tb3+ cleaves RNA at the magnesium-binding site and the single stranded regions at high concentrations. To eliminate the effect from the docking events, the footprinting assay was carried out at low salt conditions, in which the ribozyme resides primarily in the undocked state. The footprinting assay shows almost no difference in the secondary structure of the separated subpopulations. A single nucleotide in the slow undocking population is more single stranded than in the slow undocking population, suggesting a possible structural basis for the docking heterogeneity. However, low salt condition is purposely chosen in the footprinting assays to prevent docking. Tertiary interactions between two loops were not studied, which could be important for the memory effect.

1.3.2.4. Longevity of the memory effect

It is amazing to see the long lasting memory effect in ribozyme folding dynamics. At room temperature, only 9% of the Tetrahymena group I ribozyme losses the memory effect after 40 minutes[15] and 5% of the hairpin ribozyme losses the memory effect after 3 hours[14]. Walter's study on the two-way junction hairpin ribozyme shows an extremely long lasting memory effect in RNA folding[31]. In the EMSA experiment, non-cleavable mutant (A-1 2'O methyl) of two-way junction hairpin ribozyme was used to study folding independent of catalysis. The two major subpopulations of ribozyme were separated on a native gel and subsequently recovered for single-molecule experiment. Surprisingly, the dynamics of each subpopulation that was purified from the gel persist in single molecule experiments showing that the memory effect lasts for long time periods. Single molecule FRET study shows a major short-lived docked state (high FRET~0.8) one population, and a much longer docked state for the other population.

More strikingly, the longevity of this memory effect lasts through out the denaturing and re-annealing process. Walter et al, has run denaturing gels of the hairpin ribozyme eluted from the native gel. This process will put the RNA in urea at 50 °C for about 2 hours. The separated 2 strands of the ribozyme from both bands were then eluted and recombined again in all combinations. EMSA was carried out to separate subpopulations. For the combinations containing the RZB strand from the population that prefers the docked state, this subpopulation is preferentially populated (70% and 99% respectively). For all other strands, the re-annealed hairpin ribozyme redistributed between the subpopulations. The long lasting memory effect is unlikely due to secondary structure difference as during the denaturing gel, all the secondary structure was eliminated, and it should redistribute during the re-annealing process. The memory effect is independent of EDTA, which was used to chelate the Mg^{2+} from the binding site of the ribozyme. This indicates that Mg^{2+} was not involved in the memory effect of hairpin ribozyme folding.
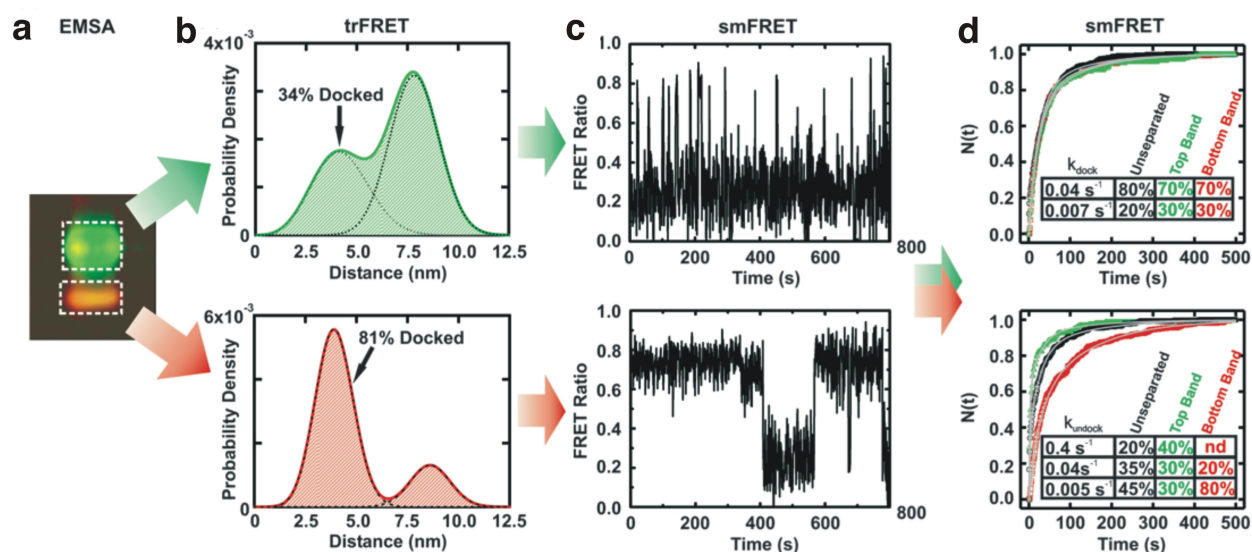
Figure 1.6 Enrichment of Sub-populations using EMSA. a. Subpopulations were separated by native FRET polyacromide gel in two bands: top band with low FRET and bottom band with high FRET. **b.** trFRET study using TCSPC shows two peaks corresponding to docked and undocked structures. Top band RNA favors undocked structure and bottom band favors docked structure. **c.** Exemplary FRET time course of the top band RNA and bottom band RNA. **d.** Comparison of kinetic rate of docking and undocking for top band, bottom band and unsperated hairpin. Reprinted from Ref. [31].

1.3.3. Transition state of docking

To further understand the origin of the memory effect of undocking heterogeneity, it is helpful to study the tertiary interactions formed in the transition state of ribozyme.

1.3.3.1. Docking transition state of hairpin ribozyme

To study the transition state conformation of the two way junction hairpin ribozyme[32], tertiary interaction information was obtained from crystallography[22]. Site directed mutations were made to the non-cleavable 2-way junction hairpin ribozyme to test the formation of tertiary interactions in the transition state. Φ-value analysis was adopted to evaluate the formation of tertiary interaction in the transition state. Φ was defined as: $\Phi = \Delta\Delta G^{\ddagger}_{dock}/(\Delta\Delta G^{\ddagger}_{dock} - \Delta\Delta G^{\ddagger}_{undock})$. An Φ value of 1 indicates the interaction disrupted by the mutation is formed before the rate-limiting step; in other words, the

interaction was completely formed in the transition state. Vise verse, an Φ value of 0 indicates the interaction disrupted by mutation was not formed in the transition state. $\Delta\Delta G^{\ddagger}$ was calculated as $\Delta\Delta G^{\ddagger} = RTln(k1/k2)$, where k1 and k2 are is the docking/undocking rate constants of wild type ribozyme and mutant, respectively. Mutational study shows that the ribose zipper interaction is at most partially formed under these conditions, as k_{dock} changes only 2.9 fold for a deoxy mutation at G11. Also, the U42 binding pocket is at most partially formed in the transition state, as k_{dock} changes only 1.7 fold for the dC12 mutant. The G+1:c25 base pair is not formed at all in the transition state, as k_{dock} does not change for A+1:C25 mutation.

Metal ion dependence of the docking and undocking rate constants was also determined from single molecule FRET study at varied magnesium and sodium concentrations. Both magnesium and sodium ions can accelerate docking as they can neutralize the negatively charged backbone and cancel the electrostatic repulsion. However the undocking rate constant is magnesium independent after 2mM [Mg^{2+}] when no sodium is present. The transition state is also stabilized by sodium ions, which are unlikely to have specific interaction with RNA. In the mixed titration experiment, magnesium and sodium show competition to binding to non-specific binding sites. Finally, they propose a theoretical model to visualize the transition state, which is different from the docked structure. In the transition state, divalent metal ions accumulate at the domain interface. Magnesium only titrations suggest that the transition state is most probably the contact model, in which the two domains are contacting (but not the same as docked state). At low [Mg] with or without Na, the

transition state is more likely a noncontact model, in which two domains are 5 Å apart, allowing up-taking of metal ions.

Other possible contributes to the RNA folding free energy barrier of docking induced additional non-electrostatic interactions, such as hydrophobic interactions. However, base stacking does not make a significant contribution as indicated by the lack of an effect of denaturant on the k_{dock}.

1.3.3.2. Transition state for Tetrahymena group I ribozyme

To study the transition state of Tetrahymena group I ribozyme docking, mutants shown to change the equilibrium for docking were chosen to study tertiary structure[29]. Almost all of the backbone modifications in the -1 to -3 and 22 to 24 regions resulted in a 0 Φ value, which indicates a transition state dissimilar to the docked state (or early transition state).

Φ analysis of guanosine binding shows no effect on k_{dock} suggesting that the ribozyme active site regions responsible for coupling are not involved in the transition state. Metal ion titration experiments show no effect of increased magnesium and sodium ion concentration on k_{dock} indicating no significant electrostatic difference between the undocked state and transition state. Destabilizing effect of Na on the K_{dock} is likely due to displacement of Mg that support docking. Urea denaturation experiments show only little change in solvent accessible surface area (ΔASA) in the transition state. Temperature dependence experiments show different signs of ΔS and $\Delta S^{\ddagger}$, indicating physical differences between the docked state and the transition state. All the evidence listed above indicates an early transition state.

Finally, a kinetic trap model was proposed for the docking transition state. They suggest that large cooperatively folded RNA such as the Tetrahymena group I ribozyme can have small local regions without native structure that forms kinetic traps. A kinetically trapped state in the undocked structure involving J1/2 and possibly J8/7 can best explain the dissimilarity between the transition state and the docked state.

1.3.4. Origin of memory effect

1.3.4.1. Possible Origin of memory effect

Zhuang *et al* suggest that the memory effect in the hairpin ribozyme originates from different conformation in loop A and B[14], which has different conformations in the NMR and crystallographic studies. Mutational study also shows lose of the memory effect when tertiary interactions in the loop are partially disrupted. If there were two different conformations for both loop A and B that each independently affect undocking, then the four combination of loops would result in four distinct undocking rates. Although no well-defined alternate structures were proposed, the meta-stable alternating loop conformation model is the prevailing model.

In addition, Ha's assay with 4-way junction without loops shows narrow distribution of averaged dwell time of U_P (undocked and arm A and B are proximal) and U_D (undocked and arm A and B are distal) states of four-way junction sample also suggest that the heterogeneity originates from the loop regions of ribozyme[26]. They proposed that the rate-limiting step for 2-way junction docking was the encounter of two domains, which is nearly three orders of magnitude slower than the transition rate from U_P to docked state. Thus only the rate of two domains encounter was observed, which shows a single exponential decay. Instead, in 4-way junction the rate-limiting step is the

transition from Up to docked state (rate constant ~ 3 s-1). And the presence of the junction accelerates the encounter of two arms (rate constant ~ 50 s-1). So the dwell time scatter plot shows the heterogeneous transitions from the U_P to docked state.

1.3.4.2. Magnesium's role in the memory effect of the Tetrahymena group I ribozyme

Magnesium ions don't have significant contribution to the memory effect of the hairpin ribozyme, however it plays important roles in the memory effect of the Tetrahymena group 1 ribozyme. Inter-conversions between subpopulations of Tetrahymena group 1 ribozyme have been observed by removing and restoring magnesium[15]. With magnesium, the inter-conversion between subpopulations is slow (9% after 40 min). After removal of magnesium using EDTA, tertiary interactions were disrupted but the secondary structure was intact. Now that the barrier separating the subpopulations has disappeared, the ribozyme will redistribute into subpopulations according to the free energy diagram after addition of magnesium. As shown in figure 1.7e, inter-conversions between subpopulations were observed. This result indicates that the memory effect in the Tetrahymena group 1 ribozyme is intrinsic and is mainly caused by different tertiary interactions that require magnesium ion.
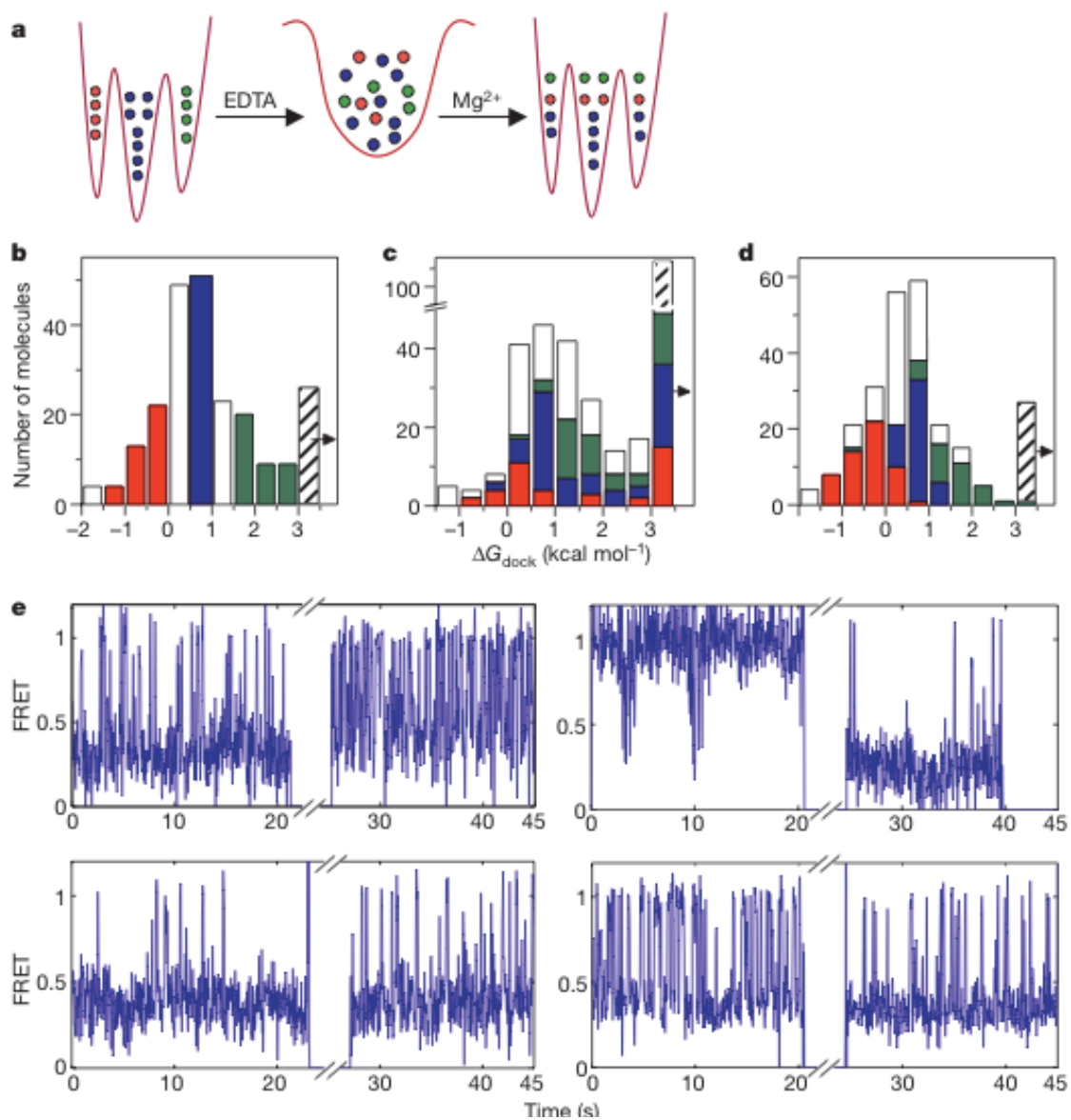
Figure 1.7 Magnesium's role in group I intron ribozyme folding heterogeneity. a. RNA folding free energy diagram changes with the magnesium removal and restore. **b.** Arbitrary categorize (colored bars) of subpopulation according to ΔG_{dock} value before magnesium removal. White bars were excluded in the analysis to minimize overlapping between subpopulations. Shadow bar on the right correspond to RNAs with $\Delta G_{dock} > 3$. **c.** Ribozyme redistributed into other subpopulation with different ΔG_{dock} after removal and restore of magnesium. Colors were matched with b. **d.** Without perturbation, Ribozyme slightly redistributed after 40 minutes. e. Exemplary FRET time course of ribozyme redistributed into other subpopulations. Reprinted by permission from Macmillan Publishers Ltd: [Nature] ([15]), copyright (2010).

1.4 Concluding remarks and future challenges

Single molecule study of RNA folding has uncovered the heterogeneous nature of RNA molecules in folding. A series of studies all indicate that the memory effect in RNA folding is intrinsic and not due to artifacts such as covalent difference from synthesis. And the longevity of the memory effect indicates a high free energy barrier to overcome to loss the memory effect. For the hairpin ribozyme, the memory effect most probably originates from structural heterogeneity in the loop region. A similar memory effect was found in the Sarcin-Ricin loop RNA, which shares structural similarity with the hairpin ribozyme in loop B. For the Tetrahymena Group I intron ribozyme, the memory effect originates from interactions involving magnesium ions.

For future study, it would be helpful to introduce inter-conversion between hairpin ribozyme subpopulations at the single molecule level to give unambiguous evidence to show that the memory effect is an intrinsic property of the hairpin ribozyme. Calculation of the free energy barrier height between subpopulations is will elucidate the molecular mechanism that causing memory effect. To achieve the aims mentioned above, we developed a new technique called 'Laser-Assisted Single-molecule refolding'.

CHAPTER 2

Methods

2.1 Förster Resonance Energy Transfer (FRET)

Förster resonance energy transfer (or fluorescence resonance energy transfer), named after the German scientist Theodor Förster, refers to the mechanism of energy transfer between two chromophores. In FRET, one donor chromophore is directly excited. Then it can either emit a photon in its emission wavelength or it can transfer the energy to a nearby (<10 nm) acceptor chromophore through a radiationless pathway. Excited acceptor will emit a photon in its emission wavelength.
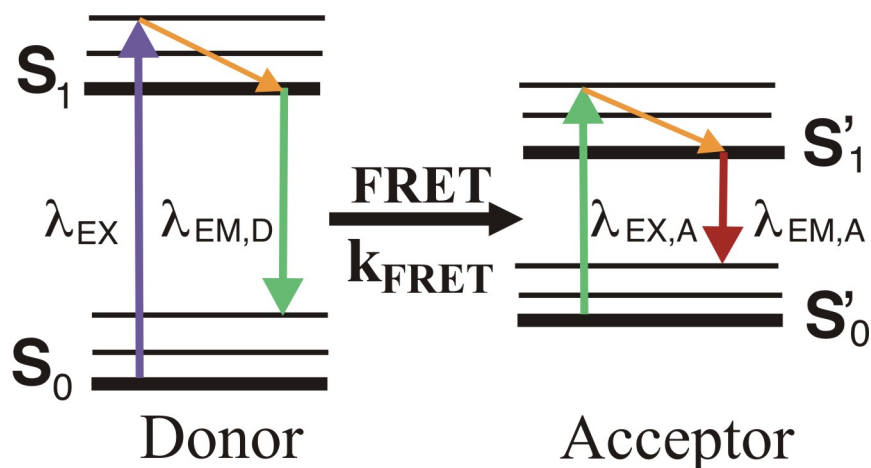
Figure 2.1 Jablonski diagram of FRET donor and acceptor. λ_{EX} is the excitation wavelength of donor fluorephore. $\lambda_{EM,D}$ is the emission wavelength of donor. $\lambda_{EX,A}$ is the excitation wavelength of acceptor fluorephore. $\lambda_{EM,A}$ is the emission wavelength of acceptor. k_{FRET} denotes the rate of energy transfer.

Since the distance between donor and acceptor (R) is smaller than the emission wavelength, the energy transfer is similar to near field communication. The energy transfer is through dipole-dipole interaction, which makes the FRET efficiency proportional to the inverse 6th power of distance R:

$$E_{FRET}(R) = \cfrac{1}{1 + \left(\cfrac{R}{R_0}\right)^6}$$

Eq(2.1)

where E_{FRET} is the FRET efficiency and R_0 is the Förster distance, where half of the energy is transferred from donor to acceptor. As shown in figure 2.2, FRET efficiency changes dramatically around R_0, which makes it a very powerful molecular ruler to measure the molecular structural fluctuation between 20 and 80 Å. FRET has been applied to study the structure of biopolymers such as RNA, DNA and protein and yields tremendous important information[28,33-37]. In addition, FRET is complementary tool to other structural methods such as crystallography and NMR.
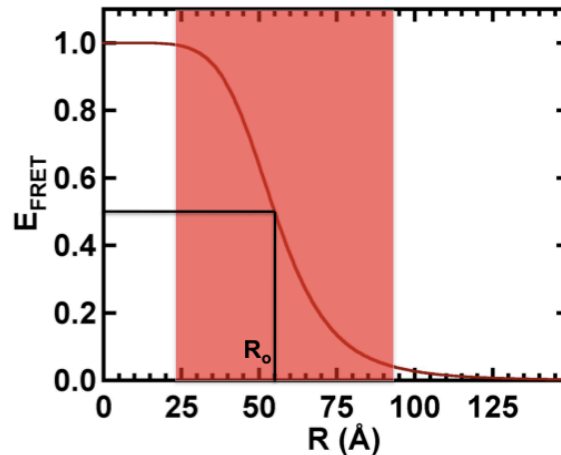
Figure 2.2 Distance dependence of efficiency of FRET for Cy3 and Cy5. R_0 is the förster distance, where half of the energy of excited donor was transferred to acceptor.

R0 depends on many factors as we can see in equation 2[38]:

$$R_0^6 = \frac{9000Q_0(\ln 10)\kappa^2 J}{128\pi^5 n^4 N_A}$$

Eq(2.2)

where Q_0 is the quantum yield of donor fluorescence alone, κ^2 is the dipole

orientation factor, J is the integral of spectrum overlap, n is the index-of-refraction of the medium and N_A is Avogadro's number.

In practice, overlap between the emission spectrum of donor emission and the acceptor excitation is required for energy transfer. One should chose the FRET donor and acceptor carefully so that they have as large an overlap as possible between donor emission and acceptor excitation. At the same time, they should have as small an overlap as possible between donor emission and acceptor emission for separation purposes when detecting the emission photons. Figure 2.2 shows that Cy3 and Cy5, a widely used FRET pair, meets the entire requirement above.
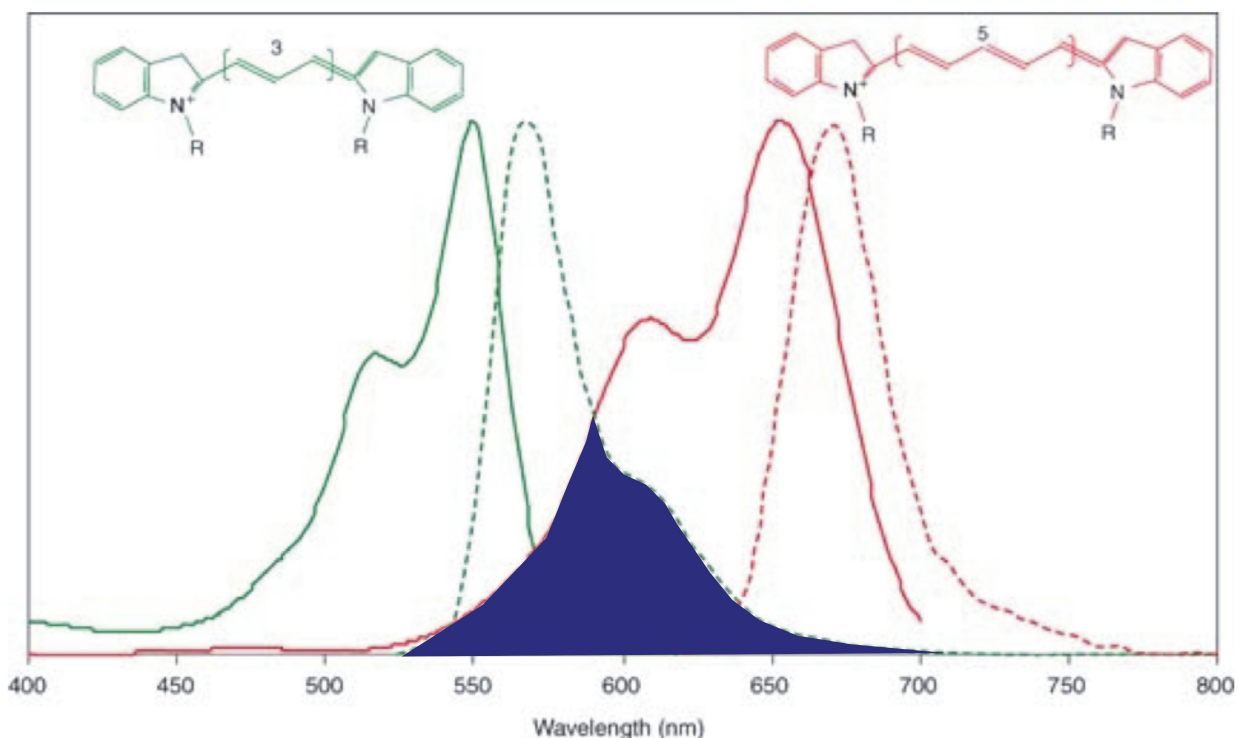
Figure 2.3 Spectrums of Cy3 and Cy5. Blue indicates the overlap between donor emission and acceptor excitation. Reprinted from Ref. [15,39].

Notably, κ^2 =2/3 when donor and acceptor dye molecules are freely rotating.

When designing FRET experiments, it is important to allow the dye molecules to freely rotate (i.e. Carbon 6 linker). Otherwise, R_0 will change and result in a false FRET change. The limitation of FRET includes the requirement of labeling with large organic dye molecules, which may interfere with molecular dynamics.

2.2　Total internal reflection

To achieve single-molecule fluorescence detection sensitivity, it is essential to minimize all kinds of background fluorescence. Total internal reflection fluorescence (TIRF) can be adopted for this purpose by only exciting a thin layer near the quartz-water interface[40,41]. The excitation laser is introduced to the quartz-water interface by a prism or an objective. When the incident angle of the excitation laser at the interface is larger than the critical angle, the laser beam will be totally reflected without a refracting beam going through the water. Snell's law can be used to calculate the critical angle.

$$\theta_c = \sin^{-1}\left(\frac{n_{sol}}{n_q}\right)$$

Eq(2.3)

Where θ_c is the critical angle and n_{sol} and n_q are the index of refraction of water and quartz slide, respectively.

However, there is still an evanescent wave penetrating the quartz-water interface for 100-200 nm. The penetrating wave intensity decreases exponentially as it goes beyond the interface. The penetrating intensity is calculated as:

$$I(z) = I_0 \exp\left(-\frac{z}{d}\right)$$

Eq(2.4)

$$d = \left(\frac{\lambda}{4\pi}\right)\left(n_q^2 \sin^2 \theta - n_{sol}^2\right)^{-1/2}$$

<div align="right">Eq(2.5)</div>

Where z is the distance in the solution from the quartz-water interface, I(z) is the evanescent wave intensity at z, I_0 is the laser intensity at the interface, λ is the wavelength of excitation laser and θ is the actual incidence angle.

Only the sample immobilized on the quartz slide is excited, and the freely diffusing fluorephores in the background remain unexcited. This way, the signal-to-noise level can be increased for single-molecule fluorescent detection.

There are two common ways to achieve TIRF: prism-based TIRF and objective-based TIRF.

i)   Prism-based TIRF: in this method, excitation laser is introduced on to the quartz-water interface using a quartz pellin-broca prism (CVI Melles-Groit, Albuquerque, NM) and an index of refraction-matched immersion oil is used to avoid the optical gap. The prism is placed on top of the quartz slide as shown in figure 2.4a. The incident angle has to be larger than the critical angle to achieve TIRF. Since d is normally fixed, the incident angle can be adjusted by changing the height h. The maximum height h_{max} corresponding to the critical angle is calculated to be:

$$h_{max} = d\sqrt{\frac{n_q^2 - n_{sol}^2}{n_{air}^2 - \left(n_q^2 - n_{sol}^2\right)}}$$

<div align="right">Eq(2.6)</div>

Where n_{air} is the index of refraction of air.

A lens (L) is used to converge the laser power. A three-dimensional micrometer

translation stage (Newport, Irvine, CA) was used to move the lens for better focusing and aligning. For pre-alignment, the excitation spot can be located with a lower magnification objective (20x, Olympus, center Valley, PA). A higher magnification water immersion objective (60x, Numerical Aperture=1.2, Olympus, Center Valley, PA) is used after the excitation spot is located in the center of field of view. The excitation spot can then be more precisely located by fine-tuning the lens position.
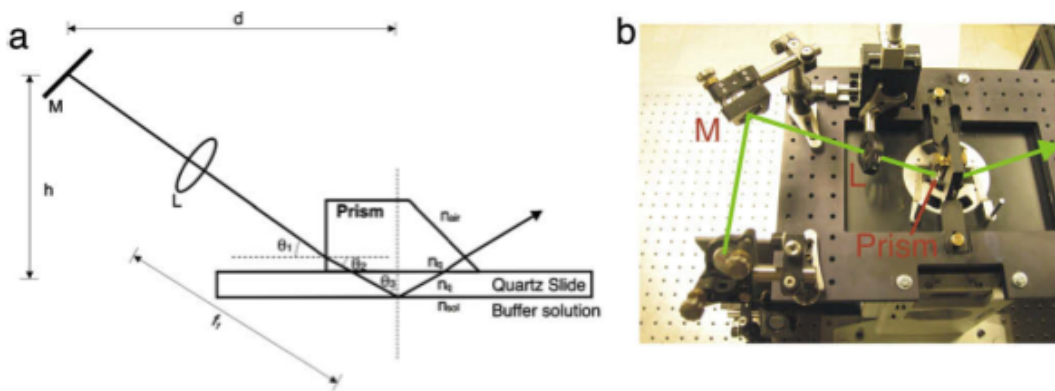
Figure 2.4 Principle of prism based TIRF. a. Schematic plot of prism based TIRF. M is the mirror. L is the lens. d and h is the vertical and horizontal distance between M and prism. f_1 is the focus length of L. n_{air}, n_q and n_{sol} are the Index-of-Refraction of air, quartz and solution, respectively. θs are the incidence angles. Excitation Laser path is denoted with black solid line with single arrow. **b.** Real image of prism based TIRF setup with schematic laser path. Reprinted from Ref. [28].

ii)   Objective-based TIRF: Compared to prism-based TIRF, objective-based TIRF is more difficult to achieve and maintain. The total internal reflection spot is correlated with the focus point, which makes TIRF very sensitive to mechanical vibrations. In objective-based TIRF, the excitation laser is introduced to the glass-water interface using an objective and an index of refraction matched immersion oil from bottom side of the assembled slides (Fig. 2.5b). A beam expander consisting of two lenses increases the side of the excitation spot. Two lenses with different focal lengths

(f1 = 5 mm and f2 = 400 mm, respectively) can expand a 0.4mm diameter laser beam to 25 mm when they are 405 mm apart. The expanded beam is then focused using a third lens to the back focus plane of the objective (60x, oil immersion, Numerical Aperture = 1.45, Olympus, Center Valley, PA)[42]. The key point to achieve objective TIRF is to have the laser beam come to the objective perfectly parallel to the vertical symmetric axis. Irises (Thorlabs, Newton, NJ) are helpful for this purpose. Then TIRF will be achieved when the beam is moved away from the center. For an objective of 1.45 numerical aperture, there is only a small range of distance to achieve TIRF. Too close to the center, the beam will be refracted. Too far away from center, the beam will be blocked. A signature of perfect TIRF is a bright totally reflected outcoming beam with the same size as the incoming beam and the absence of a refracted beam.

The procedure for objective-based TIRF alignment is as follows using the nomenclature in Fig 2.5:

1. Put a slide with fluorescent beads and water solution on the microscope and focus on the beads immobilized on the cover slip. This focal position of the microscope should not be changed throughout the alignment.

2. Without the beam expander and L3, align the laser beam to a near TIRF position using Mirrors M1 and M2, and dichroic mirror DM. M2 should be installed on a 2 dimensional (horizontal plane) translation stage to be moved parallel.

3. Put the beam expander between the laser and M1. Make sure the laser beam is going through the center of lenses, so that it does not change direction.

4. Put L3 between the microscope and M2, where the beam can be

focused on to the focal plane of the objective. Installing L3 onto a 2 dimensional translation (perpendicular to the beam path) stage can help to align the beam accurately to the TIRF position.
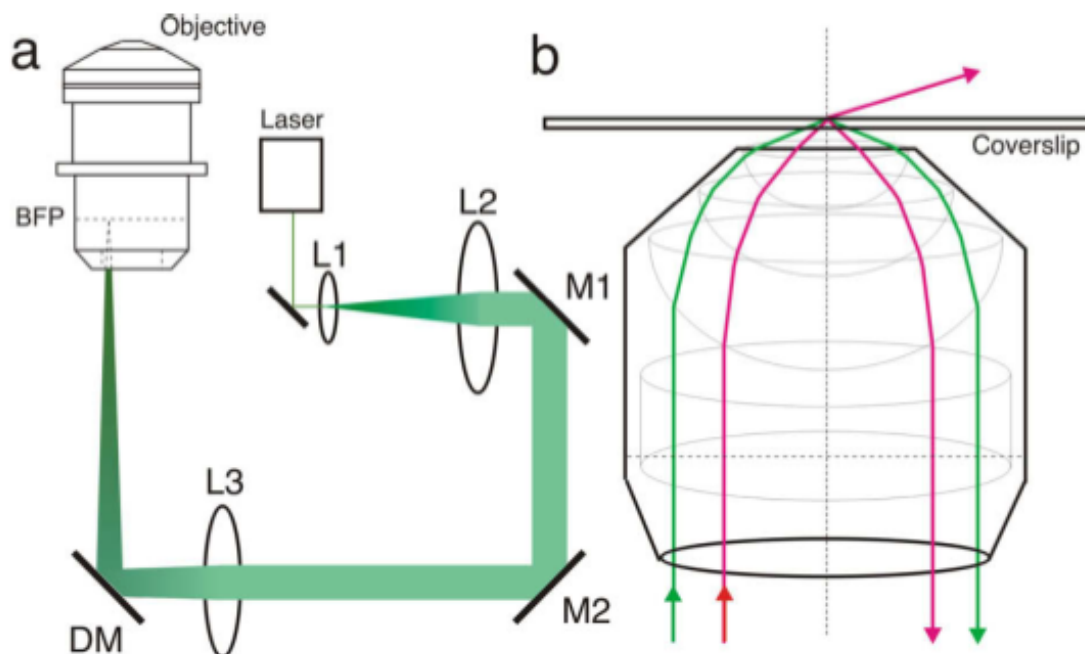
Figure 2.5 Principle of objective based TIRF. a. Schematic plot of objective based TIRF. Ls are lens. Ms are mirrors. DM is a dichroic mirror. BFP is the back focal plane. **b.** Light paths of excitation laser. Colors do not correspond to the laser wavelength. Reprinted from Ref. [28].

iii) Pros and cons: Compared to objective TIRF, prism based TIRF is easy to achieve and maintain. But the use of a prism limits the operator to samples that can be examined from the top. Since objective TIRF is sensitive to the mechanical vibrations, the focus drifts slowly (minutes) during the acquisition, which makes it a suitable for slow kinetic measurement.

2.3 Single molecule FRET detection

Over the past two decades, single-molecule experiments have yielded a wealth of answers to important biological problems. For instance, they have been used to

investigate mechanisms of molecular motors such as ATP synthase[36,43,44], myosin and kinesin[45,46], the DNA packaging engine of bacteriophage phi- 29[47,48] and the ribosome[49,50]. The most attractive aspect of single-molecule experiments is that they directly reveal the structural dynamics of individual molecules, which would otherwise be hidden in ensemble-averaged experiments[14,21,32,51-53]. There are many approaches that have been used in single-molecule microscopy, such as fluorescence correlation spectroscopy (FCS)[54], optical and magnetic tweezers[55,56], atomic force microscopy[57], confocal microscopy[58] and total internal reflection fluorescence microscopy (TIRFM)[34,59]. Each approach presents advantages and disadvantages that have been exploited to study a wide variety of biologically relevant questions. Here, we focus our attention on single-molecule spectroscopy using TIRFM excitation and FRET detection to study the folding kinetics of RNA molecules.

In a typical FRET study, FRET donor and acceptor fluorophores are used. There are many FRET donor and acceptor pairs available on the market including Cy3 and Cy5 as well as fluorescein and tetramethylrhodamine. In our studies, we have adopted Cy3 and Cy5 as our FRET donor and acceptor, as Cy3 and Cy5 have high quantum yields and give bright, stable single-molecule emission. Cy5 photobleaches occasionally, but it can be minimized using trolox in the reaction buffer. The FRET donor (Cy3) is excited using a 532 nm green laser. The excited Cy3 can transfer the energy to a nearby FRET acceptor (Cy5), so both emission from donor and acceptor can be detected.

The efficiency of FRET can be experimentally determined as: $FRET = I_A/(I_A+I_D)$, where I_A and I_D are the emission intensity of donor and acceptor. This apparent FRET is

good to estimate the distance information qualitatively. However, some groups have tried to measure absolute distance quantitatively from single molecule FRET measurement[60-62].

Our CCD camera is not color sensitive, so it is necessary to separate emission of donor and acceptor to two different channels and measure the emission intensity separately. To do that, a series of dichroic mirrors, lenses and mirrors are used to build a dual channel viewing system. First, the image from the side port of the microscope is cut by a slit to the size of half the camera at the focal distance to avoid smearing the image edge when out of focus. The image with mixed emission signal is then separated using a dichroic mirror that separates Cy3 and Cy5 emission (DM1, 635DCXR, Chroma, Rockingham, VT).
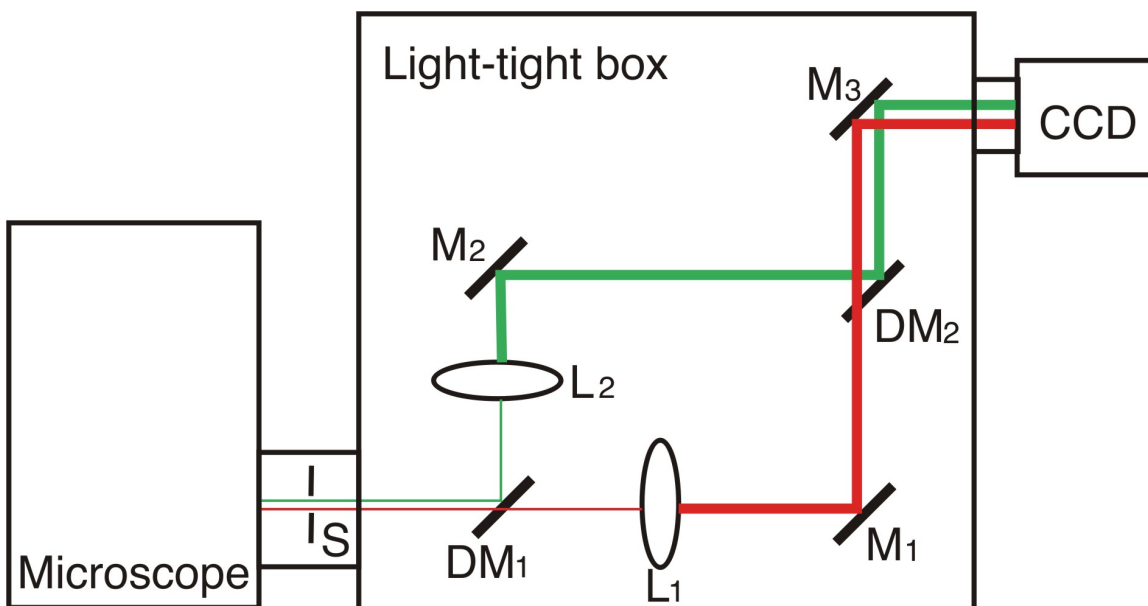
Figure 2.6 Schematic plot of the light tight box. Ls are lens. Ms are mirrors. DMs are dichroic mirrors. Original image from the microscope is separated into two colors, amplified by 2 and put side by side onto the CCD camera. Reprinted from Ref[28].

Image size in both the red and green channels was amplified using 2 fold using lenses (L1 and L2). Then the images in both channels were recombined side-by-side using dichroic mirror (DM2). Finally the image of separated channels will come to the CCD camera sensor chip, where the emission intensity of both donor and acceptor will be digitalized and further stored on a computer. Two color separation systems with adaptors to microscope and CCD camera is also commercially available (Optosplit, Andor, South Windsor, CT)

2.3.1 CCD camera for single-molecule detection

The fluorescence emission image of single molecules is captured using a high quantum yield amplified CCD camera (IXon+, DV-897E, Andor, South Windsor, CT). The sensor chip is air cooled to -80 to minimize the thermal noise. The signal is also

amplified by electron multiplication gain at its highest value to maximize the sensitivity of camera. Together with the TIRF excitation and light-tight of the optical path, the camera is able to detect a low number of photons from single fluorophores. To collect enough photons for each molecule in each frame, the time resolution of the camera is typically set as 33 frames per second. Higher time resolution (up to 3 ms in practice) capture can be achieved by reducing the number of vertical lines of acquisition. Higher excitation laser intensity should be used to compensate for the decrease in the number of photons in each frame associated with a shorter accumulation time. Similarly, slow kinetic measurements can be achieved using increased exposure time and decreased excitation laser intensity.

The digitalized image signal on the CCD camera is then sent to and stored on a computer for further analysis using an acquisition card controlled by labview (National Instruments).

2.3.2 protocol of conducting single molecule FRET using TIRFM

Preparation of microscope slides with microfluidic channels
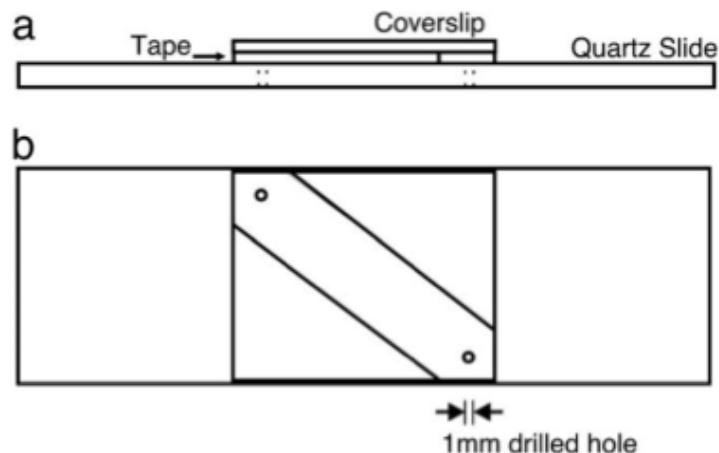
Figure 2.7 Schematic plot of assemble sample slides. a. Side view of the assembled slide. Coverslip was attached to the quartz slide with double sticky tape to make the sample channel. **b.** Top view of the assembled slides. Reprinted from Ref. [28].

The normal microscope slide must be modified to make a fluidic channel to introduce the samples. Two holes were drilled in the quartz slides using a hand drill (Dremel 300-N, Racine, WI) and drill bits (1.0 mm diameter, Kingsley North, Norway, MI) on a work-station (Dremel 220-01, Racine, WI). Then the slides were cleaned thoroughly using the following protocol:

1. Boil slides for 30 min in distilled water. Then remove the epoxy and sticky tape from previous experiment (if any) using a razor blade.

2. Scrub slides vigorously by hand with thick detergent power (Alconox, VWR, West Chester, PA) solution for 2 min. Then completely rinse off the soap from the slide with double-distilled water (ddH$_2$O).

3. Rinse and Scrub the slides with ethanol and ddH$_2$O successively for 1 min each. Then repeat it again.

4. Boil the slide in the mixture of 20 mL hydrogen peroxide (30%), 20 mL ammonium hydroxide (30%) and 100 mL ddH$_2$O for >20 min.

5. Stop boiling, rinse the slides with ddH$_2$O and dry them in the flame of a Bunsen burner. Drive the water from one side of the slide to the other using the flame to avoid formation of deposits on the slide.

After the cleaning of slides, they should be assembled immediately.

1. Put two pieces of double-sided sticky tape (scotch) parallel onto both side of the drilled holes approximately 4 mm away from the two-hole line. A second layer of double-sided sticky tape can be applied to increase the channel depth. Each layer is about 100 μm thick.

2. Carefully place a coverslip onto the sticky tape and cover the two holes evenly. Gently push the coverslip against the sticky tape to stick it on the slides.

3. Cut along the four edges of the coverslip using a razor blade to remove the excess sticky tape.

4. Seal the edge of the microfluidic channel with epoxy (5 min epoxy, Devcon, Danvers, Ma)

This finished microfluidic channel is 9 mm wide and 200 μm deep. The sample can be injected into the channel using trimmed 200 μL pipette tips. The slides can be reused after this protocol more than 12 times.

2.3.3 Sample preparation

The labeled and purified RNA/DNA sample[33] must be immobilized them on the quartz slide (coverslip) for TIRFM study. The protocol is as follows:

1. Heat anneal sample in 10 μL buffered solution with desired salt condition

(typically 1 μM labeled sample, 50 mM TRIS-HCl pH 7.5, 50 mM NaCl and variable [$MgCl_2$]). The buffer condition can be changed according to the system being studied.

2. Serially dilute the sample in standard buffer to a final concentration of 10-25 pM. Keep the sample on ice away from light to avoid photobleaching of the dyes.

3. Inject 50 μL of biotin-linked bovine serum albumin (B-BSA, Sigma-Aldrich, St, Louis, MO) solution (1 mg/ml in 50 mM TRIS-HCl, pH 7.5, 50 mM NaCl) into the microfluidic channel. Let it incubate for 10 min. Put two drops of solution on the two holes to avoid drying out.

4. Inject 85 μL of T50 buffer (50 mM TRIS-HCl, pH 7.5, 50 mM NaCl) into the channel two times to flush out the unimmobilized B-BSA.

5. Inject 85 μL of 0.2 mg/mL streptavidin (Invitrogen, Carlsbad, CA) solution (in T50 buffer). Let it incubate for 10 min.

6. Inject 85 μL of standard buffer two times to flush out the unimmobilized streptavidin.

7. Inject 85 μL of 10-25 pM sample solutions. Let it incubate for 5 min.

8. Inject 85 μL of standard buffer to flush out the unimmobilized sample.

Now the RNA/DNA sample is surface immobilized and ready for TIRFM study. However, the oxygen molecules in the solution can cause the dye molecules goes to the 'dark' state during the laser excitation, which will interrupt the FRET and give false results. To avoid that, an oxygen scavenger solution (OSS) is introduced to the fluidic channel. The protocol to prepare OSS is as follows:

1. Dissolve ~25 μL (by powder volume) of Glucose Oxidase (Sigma-Aldrich, St. Louis, MO) and 6.2 μL of Catalase (Roche Diagnostics, Indianapolis, IN) in 50 μL of T50 Buffer. This solution should be stored at 4 ˚C cannot be used for more than 3 days.

2. Add 2 μL of the above solution into 198 μL of reaction buffer with 10% (by weight) dextrose (Fisher Scientific, Pittsburgh, PA) and 2 μL 2-mercaptoethanol (Fisher Scientific, Pittsburgh, PA). Let it incubate for 5 min.

3. Inject the OSS solution into the microfluidic channel with the immobilized samples.

    In addition to this OSS solution, an alternative has been developed by the Puglisi's group to avoid possible interference of this enzyme system to some specific biological system[63,64]. Also, the Ha's group has found that trolox, a vitamin E analogue, helps preventing blinking and photobleaching of the dyes[65]. Finally, sealing of the two holes on the slide with epoxy can prevent the oxygen in the air from dissolving into the solution. This can prevent the pH value change of the solution from the OSS reaction and extend the lifetime of the sample-immobilized slide to more than 3 hours.

2.3.4 Data mining for single-molecule kinetics

The saved image array containing kinetic information of single-molecules needs to be converted to readable emission intensity time trajectories. For this purpose, a homemade program written in IDL (ITT Visual Information Solutions) is used to search for the molecules and export the intensity for each frame. Accurate mapping of the same molecules in both channels are calibrated in advance using the image of

fluorescent bead molecules, as they are bright. When calibrating, thirty frames images are averaged to give a high quality picture; coordinates of three pairs of molecules are manually chosen in both channel as an initial guess for the alignment mapping. After background subtraction, the program will search the whole image for nice Gaussian peaks with certain criteria to differentiate good molecules from noise and artifacts. All the molecules in both channels will be matched using the polynomial optimization method. The mapping correlation matrix between two channels is exported and saved for the analysis of single-molecule experiments.

When analyzing the single molecule fluorescent image, the first 100 frames are averaged and background subtracted to give high signal-to-noise (s/n) ratio image. The two channels are overlapped using the mapping information from the previous step. Then the program will search through the overlapped image for high quality Gaussian peaks as good molecules. Once a good molecule is found, the x and y coordinates of the center of Gaussian peak are saved as the position of the peak for future use. Once all the good molecules are found in the overlapped image, the program integrates (with background subtraction) the Gaussian peak for all the good molecules in each frame as the emission intensity of molecules. Then the fluorescent emission time trajectories are exported into an ASCII file for future analysis.

2.3.5 Thermodynamic analysis

For a single molecule FRET time trajectory as shown in figure 2.8, we can calculate the free energy difference between two FRET states according to the ratio of times spent in both states. The dwell time of a state, for instant, state A, starts from a transition to state A and ends at transition to state B. To get a good statistics, at least 30

molecules are necessary, and 100 molecules are generally considered desirable. In most cases, the addition of more than 100 molecules does not significantly change the histograms or dwell times.
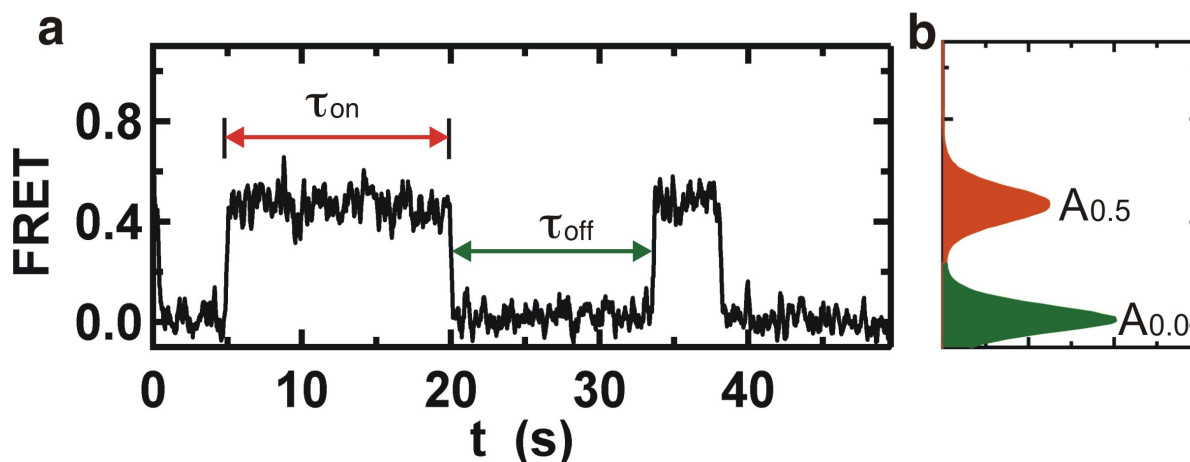
Figure 2.8 Histogram analysis of a exemplary FRET time course. a. A FRET time course with two states at around 0 and 0.5. τ_{on} is the time spent on one event on 0.5 state. τ_{off} is the time spent on one event on 0.0 state. **b.** Histogram of FRET showing two Gaussian peaks at 0.0 and 0.5. A0.5 and A0.0 denote the area from integration for the two peaks. Reprinted from Ref. [28].

Assuming different states were in thermodynamic equilibrium within the experimental time window, the equilibrium constant can be calculated as

$$K_D = \frac{A_{0.0}}{A_{0.5}}$$

Eq(2.7)

where A_0 and $A_{0.5}$ are the integrated events in each state. The Gibbs free energy between two states can be calculated as:

$$\Delta G^0 = -k_B T \ln(K_D)$$

Eq(2.8)

where k_B is the Boltzmann constant and T is the temperature.

2.3.6 Kinetic analysis

Not only thermodynamics parameters can be determined from the single molecule FRET time trajectories. The most exciting aspect of single molecule assay is that we can determine the apparent kinetic rate of the reactions between different states. Assuming the reactions between two states are classic first order uni-molecular reversible reactions with rate constants $k_{forward}$ and $k_{backward}$, which are with the unit of s^{-1}, $k_{forward}$ and $k_{backward}$ will be the inverse of the lifetime in the starting and ending state. The lifetime of starting and ending states can be determined by fitting the histogram of the dwell time distribution in these states to a single exponential decay. If there were multiple reactions going on between the same two states, the fit will give multiple exponential decays, which are corresponding to different reaction rates constants[66]. More complex system with more FRET states, can be analyzed uisng the HaMMy software developed in the Ha's group[67]. The free energy can also be calculated using the rate constant from the fit.

$$K_D = \frac{k_{backwards}}{k_{forwards}}$$
Eq(2.9)

If the reaction is not a unimolecular reaction, the equilibrium constant, apparent rate constant and free energy can be modified with the concentration correction of the other reagents or products.

2.4   Temperature jump (T-jump)

T-jump is a useful technique in studying chemical kinetics of biological reactions such as DNA, RNA and protein folding[68-70]. The idea is to change the temperature in a

very short time (ns-μs, has to be faster than the chemical relaxation process) to break the chemical equilibrium. The raised temperature remains unchanged before the heat diffuse away. When the chemical reactions shift to establish the new equilibrium, the observable changes of the molecular dynamics can be monitored as a way to estimate the kinetic parameters of the reactions in the system under investigation. There are two common ways to achieve the sudden rise in temperature. One is to use the discharge of a capacitor; the current going through the conductive media (buffered salt solution) is converted to thermal energy. This raises the temperature in a few microseconds. The other relatively newer way is to use an infrared laser (wavelength ~1400 nm), which will be absorbed by the water OH overtone. In theory, this method can achieve a T-jump quickly, in only 10 ps, which is the 'dead time' of this T–jump experiment. The laser-generated heat will diffuse away gradually (typically 5-50 μs), which determine the slowest limit in the kinetic detection. After this time, the chemical equilibrium shifts again due to the temperature change. The time scale of the DNA, RNA and protein secondary structure formation is in the μs range, which makes T-jump an ideal tool to study the kinetics of biopolymers folding.

The detection methods can be very different for different biopolymers. But the basic idea is to use a probe associated with structural fluctuation, such as absorbance change at the 280/260 nm during DNA/RNA folding/unfolding, the fluorescence emission change of tryptophan in proteins or the FRET changed between two labeled chromophores.

2.5 Laser-assisted single-molecule refolding

Single-molecule Arrhenius and Eyring analyses of folding rate constants can be used to determine energy barriers, but, for technical reasons, are limited to temperatures below ~50°C, which may not be high enough to melt many nucleic acid folds. Furthermore, Mg2+ ions, which play an important role in RNA folding, can cleave the backbone of RNA molecules at high temperatures (>50°C) making it impractical to study deep kinetic traps in their presence.

Here, we report a new single-molecule technique, termed laser-assisted single-molecule refolding (LASR), that enables characterization of the activation barriers of deep kinetic traps by combining traditional single-molecule detection with infrared (IR) laser-induced temperature jumps (T-jumps). An IR laser is used to excite the second overtone of the water OH stretch (~1400 nm) and thus rapidly increase the solvent temperature.
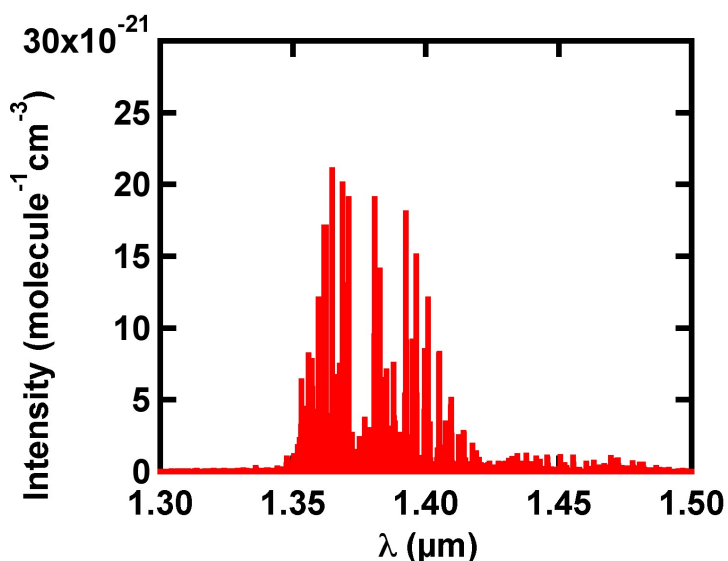
Figure 2.9 Water absorbance in the near IR region. Reprinted from Ref. [28,71].

The sudden T-jump can put an RNA molecule above the transition barrier. The energy regions $>k_BT$ of the folding potential then become accessible. Biopolymers suddenly escape from a kinetically trapped state. They can redistribute to other states (including the native state) during relaxation (Fig. 2.10).
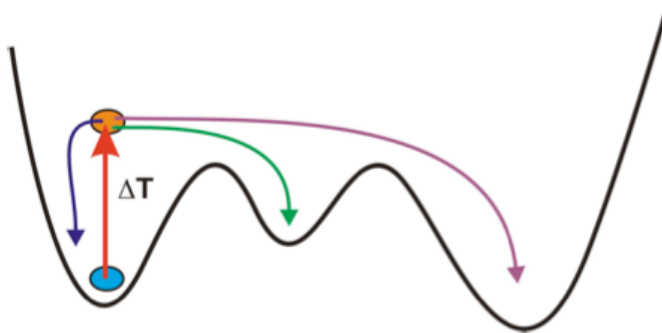
Figure 2.10 Schematic plot of LASR help biopolymer escape from kinetic trap on a potential energy surface.

Our T-jumps are transient (but reversible) temperature increases, and thus differ from traditional ones. We demonstrate the application of LASR to measure single-

molecule melting curves of double-stranded DNAs (dsDNAs) and to characterize a model kinetic trap comprised of two complementary hairpin RNAs that can form a stable loop-loop interacting complex (kissing complex). Then we applied LASR to study the memory effect in the hairpin ribozyme undocking.

In all of our studies, we are monitoring FRET changes. Due to the time resolution of the CCD camera, kinetic process below 33 ms cannot be resolved in real time. Real kinetic study faster than 33 ms is not feasible. Instead, we monitor the FRET state change before and after the T-jump and use assumptions from transition state theory to extract kinetic barrier information of the molecular reaction. A nonlinear amplified fiber laser, which generates a broadband wavelength (450-1800 nm, 2W), was used as an IR laser source. The desired wavelength (1300-1400 nm) was selected as water absorbs efficiently in this range[72]. Small portions of red visible light (650+ nm) were kept for alignment purposes. The final effective IR laser power in the water solution is estimated to be ~100 mW taking account the energy loss from reflection and the water absorption efficiency. A focal lens (focus length=100mm) is used to converge the laser beam to a small observation volume. (Fig 2.11a).
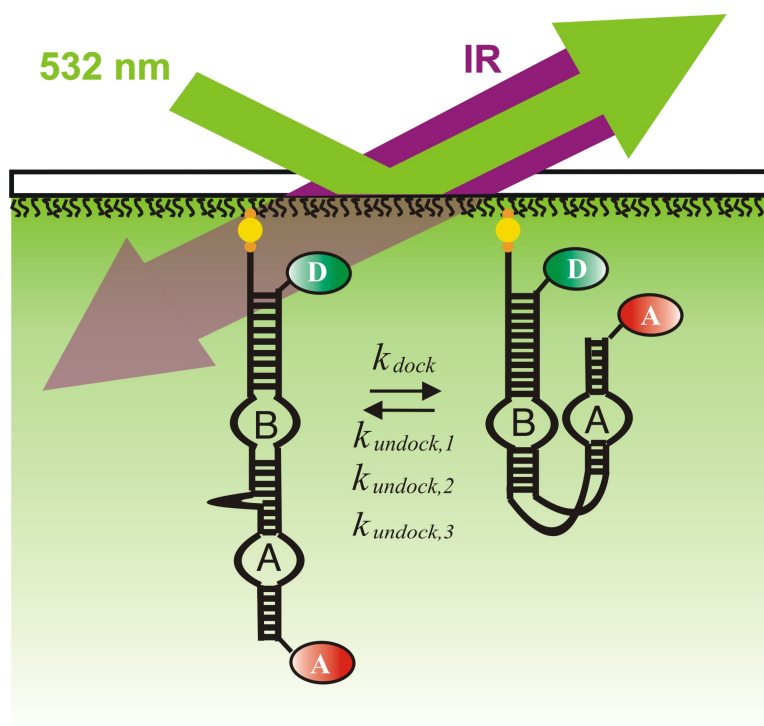
Figure 2.11 Schematic plot of LASR. Excitation laser (Green) and IR laser were incorporate into the same observation volume via prism. RNA molecules were surface immobilized by biotin (orange dot) streptavidin (yellow circle) bridge. FRET is monitored by single molecule FRET detection.

A computer-controlled shutter is used to control the duration of the temperature jump. To eliminate the time effect, all the temperature jump experiments were carried out using 1 second heating.

2.6   Temperature calibration

Once Laser-induced temperature jump is operational, the magnitude of the temperature jump must be calculated. This is very important because accurate temperature calibration is necessary for quantitative and reproducible LASR experiment. More importantly, we can get the temperature dependence of the reaction yield, which can be used to evaluate the energy barrier of the reaction.

To accurately measure the temperature in such a small volume can be

challenging.

In traditional T-jump experiment, the magnitude of the temperature jump is normally calibrated with the temperature dependence of the absorbance of a chromophore. So the first method we have tried is to characterize the temperature dependence of the emission intensity of a fluorophore beads (0.2 μm modified red microspheres, Invitrogen, Carlsbad, CA) in bulk as a calibration curve (Fig. 2.12). As we can see, the fluorescent emission of fluorophore beads decrease with the temperature linearly before 70 °C. After 70 °C, the emission intensity decreases faster in a polynomial fashion.
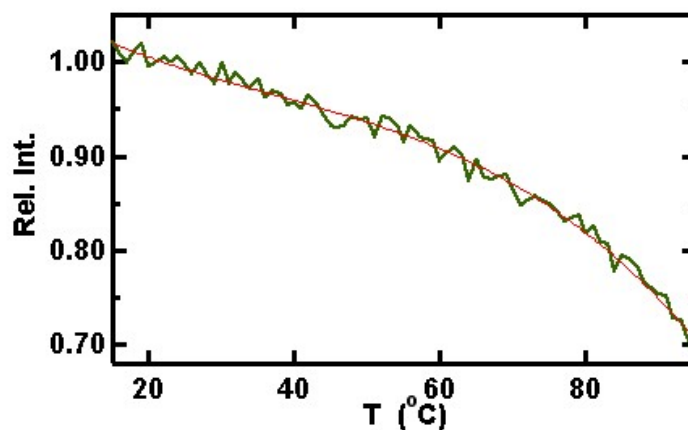
Figure 2.12 Temperature dependence of normalized fluorosphere emission from bulk measurement.
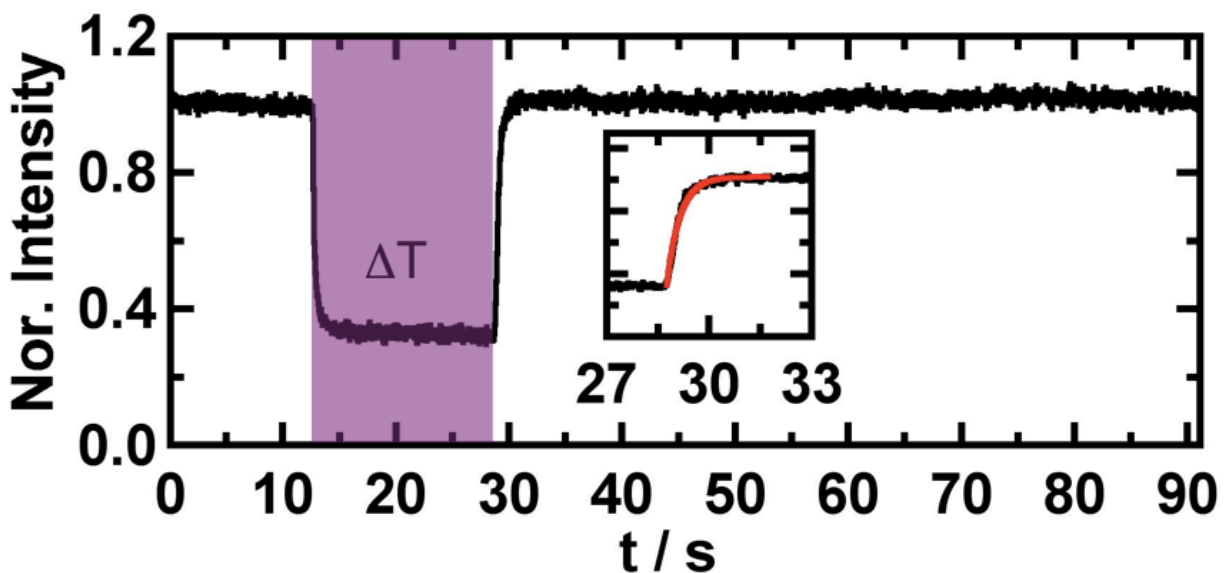
Figure 2.13 Normalized emission intensity from a single fluorosphere particle. Insert is the exponential fit of the temperature relaxation after 14 s.

Then single molecule T-jump experiments were carried out for the single immobilized fluorophore beads on a prism-based TIRFM. The normalized emission intensity of a single fluorophore bead was plotted against time. As shown in figure 2.13, the emission intensity decreased as soon as the IR laser (red region) was on. After ~2 s, the emission intensity reached a plateau, which corresponds to the maximum temperature point. As soon as the IR laser was turned off, the emission returned to the previous value, corresponding to a return to the room temperature. The relaxation process took place in 100 ms on our setup. The maximum temperature calculated from the single trace in figure 2.13 is ~150 °C. This value is higher than would be physically possible. We proposed several possible explanations. The major contribution of the error be from the distortion of the image origin due to the change in the index of refraction of the water solution inside the reaction channel. Simply put, the decreased emission intensity is partly due to the shift of the Gaussian peak out of the original

integration area. We have tried to correct this effect by relocating the position of the Gaussian peak during the temperature jump. The corrected temperature jump was calculated to be 110 $^{\circ}$C, which was closer to the true value. The image distortion effect can be minimized if objective TIRF were adopted because in objective TIRF, the fluorescent emission does not go through the solution. There is another minor error that may originate from the damage of the fluorophore beads at high temperatures. As this method turns out to be an inaccurate way of measuring temperature, we will not describe this in detail.

Other groups (D. Nesbitt) tried to use the temperature dependence of the lifetime of the excited electrons on the fluorophore to calibrate the temperature jump. Single fluorophore emission at different temperatures was measured using time correlated single photon counting (TCSPC). The final temperature can be controlled by the input laser intensity. Then a correlation between the final temperature and input laser power can be obtained. This is a very challenging and unstable way to calibrate the temperature. The calibrated temperature can be wrong by more than 10 $^{\circ}$C.

In our lab, we have adopted a gold wire microsensor to calibrate the temperature of the system. The gold wire microsensor was fabricated by depositing gold onto a glass substrate using a homemade masker (figure 2.15). A thin layer (1 micron) of PDMS was applied on top of the sensor to insulation. The temperature dependence of the resistance of the gold wire sensor can be pre-determined in a temperature-controlled oven (figure 2.14).
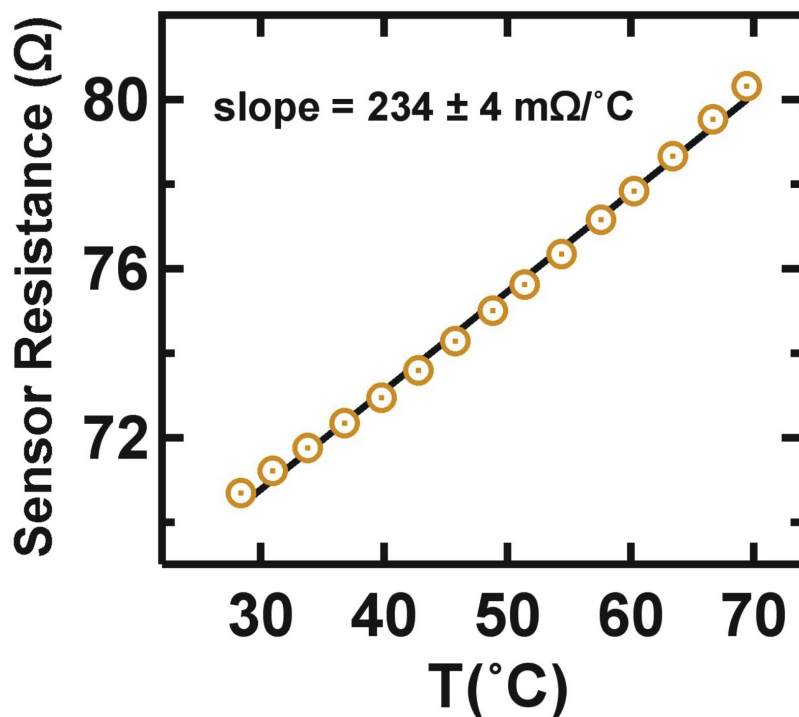
Figure 2.14 Temperature-resistance coefficient of micro-sensor measured in a temperature-controlled oven. Reprinted from Ref. [73].

During calibration of the micro-sensor, gold wire sensor was placed in a temperature-controlled oven. The temperature was increased with a step of 3°C and 5 minutes of equilibrating time. The resistance of the gold wire sensor is monitored using a multi-meter (Agilent). As expected, the resistance of the gold wire sensor increases linearly with temperature. We can safely extrapolate the dependence curve to 100 °C, as the linearity of the gold should hold at low temperature (<400°C).

Then the glass substrate with the sensor was cut and assembled to make a microfluidic channel (Fig. 2.15). The temperature calibration process should be carried out every time before the T-jump experiment with the corresponding buffer solution in the microfluidic channel.
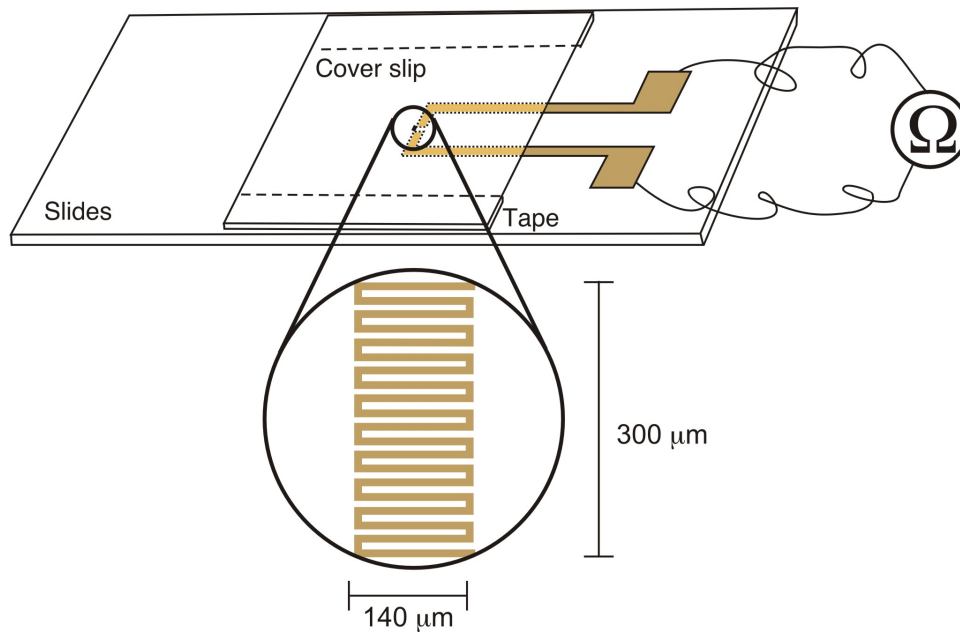
Figure 2.15 Temperature calibration using gold wire micro-sensor. Circle magnified of the sensor with a size of 140x300x1 μm. Reprinted from Ref.[73].

When doing temperature calibration of the buffered solution, the sensor slide was placed on the microscope. The IR laser was focused onto the sensor with a duration controlled by a shutter. The resistance change can be monitored using multi-meter, which can later be converted to the temperature change.
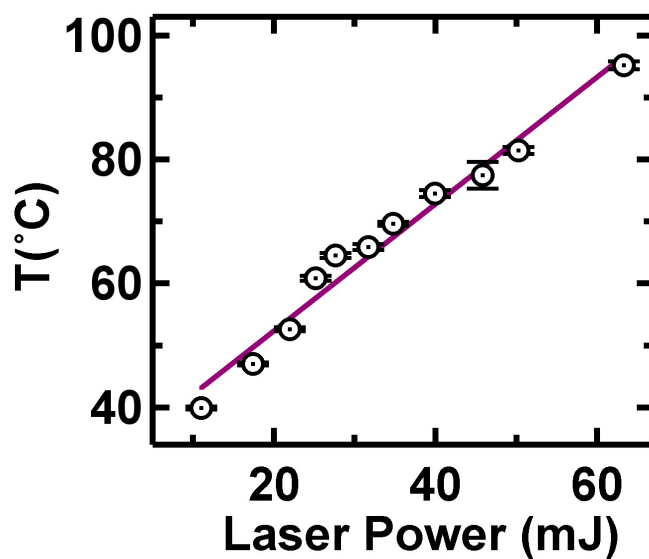
Figure 2.16 Final temperature calibration curve. Laser power was calculated from the original input power and attenuation from ND filter. Reprinted from Ref. [73].

In figure 2.16 shows a temperature calibration curve of the system. As we can see, the final temperature of the observation volume increases linearly with the input laser power. With this temperature calibration curve, the desired temperature range can be achieved using combination of ND filters.
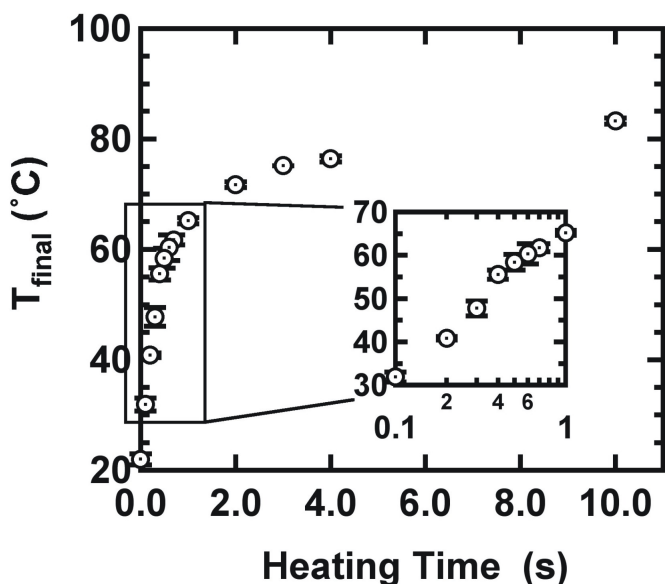
Figure 2.17 Temperature calibration curve with varied heating duration. Insert is the plot of first second with time in logarithm scale.

The temperature calibration can also be achieved by varying the duration of heating. As we can see in figure 2.17, the final temperature increases quickly in the first 2 second. After 2 seconds, it increases slowly as the heat input and dissipation reaches equilibrium. The whole process can be described as a double exponential function. This calibration method has complications from the varying heating duration, so it is not used for our study.

2.7 Concluding remarks and future challenges

We have developed a now technique combing single-molecule FRET study with laser-induced T-jump to study the kinetically trapped RNAs. We have accurately calibrated the temperature of the T-jump in small observation volume (nano-liter) with 1˚C accuracy. Due to the time resolution of the CCD camera, real time kinetic study is not feasible. However, we can use assumption from transition state theory to get barrier information from temperature dependence of the kinetic rate constants. This technique

is readily applied to study the RNA, DNA and protein folding.

(Chapter 2 is adopted from paper Zhao, R. & Rueda, D. RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **49**, 112-117, doi:S1046-2023(09)00103-0[28])

CHAPTER 3

Laser-induced single-molecule melting of DNA helix

3.1 Introduction

To validate the new technique we developed, we designed three DNA duplexes with predicted melting temperature ranging between 50 and 70 ˚C (Mfold[74]). The three DNA duplexes have 18, 21 and 24 base pairs. They are designed to form a stable DNA duplex as heterodimers, but not homodimmer (table 2). Each strand of the DNA will form a hairpin structure in the absence of the complementary strand. The actual melting temperature of the DNA duplex was experimentally determined in bulk using UV melting experiments (T_m = 48 ± 1, 61 ± 1 and 70 ± 1˚C, respectively, Fig. 3.2, black curves). Then LASR was applied to melt the DNA duplex to a give single-molecule melting curve to allow comparison with the bulk measurements to validate LASR.

3.2 Methods

Table 3.1DNA duplex sequences used in this study.

Name	Description	Sequence[1,2]
dsDNA1	Short duplex	5'-ATC TAT AAA AAT ATA GAT-3' 3'-TAG ATA TTT TTA TAT CTA-5'
dsDNA1'	Fluorophore labeled short duplex	5' Cy3-ATC TAT AAA AAT ATA GAT dT(Cy5)TTT-3' B 3'-TAG ATA TTT TTA TAT CTA-5'
dsDNA2	Intermediate duplex	5'-ATC ATC TCT CTC TAA GAT GAT-3' 3'-TAG TAG AGA GAG ATT CTA CTA-5'
dsDNA2'	Fluorophore labeled Intermediate duplex	5' Cy3-ATC ATC TCT CTC TAA GAT GAT dT(Cy5)TTT-3' B 3'-TAG TAG AGA GAG ATT CTA CTA-5'
dsDNA3	Long duplex	5'-ATT GCG ATA GAG AGA GAT CGC AAT-3' 3'-TAA CGC TAT CTC TCT CTA GCG TTA-5'
dsDNA3'	Fluorophore labeled Long duplex	5' Cy3-ATT GCG ATA GAG AGA GAT CGC AAT dT(Cy5)TTT-3' B 3'-TAA CGC TAT CTC TCT CTA GCG TTA-5'

3.2.1 UV Melting Experiments: UV-Melting experiments were carried out on a Beckman Coulter spectrophotometer (DU800, Brea, CA). Equal amounts (0.2 μM) of both strands of the DNA duplex were mixed in standard buffer (50 mM TRIS-HCl pH 7.5, 10 mM $MgCl_2$) and heated to 90 °C for 2 min and incubated at room temperature for 20 min. The formation of extended duplex was confirmed by gel shift assay (Data not shown). The absorbance of the DNA duplex at 260 nm with a 1 cm cuvette was measured at room temperature. The mixed solution was then diluted till the absorbance reaches ~1 OD. The absorbance at 260 nm was then monitored every half minute (0.5-second average) as the temperature increased from 15 to 95 °C (1°C/min rate) to give a UV melting curve (Fig. 3.2), which can be analyzed with using computer program

(MeltWin V3.5 Build 010401) to obtain melting temperatures.

3.2.2 Single Molecule Melting: For single molecule FRET study, one strand of the DNA duplex was labeled with 3' end biotin, 5' end Cy5 and a Cy3 near the 3' end. 10 μL of solution A containing 1 μM of each strand and 1 μL of BME in standard buffer condition were heated at 90 °C for 45s and incubated at room temperature for 20 min. Then solution A was sequentially diluted to a final concentration of 10 pM. Then the DNA duplex was immobilized on the quartz slides surface using biotin-streptavidin bridge. IR laser induced temperature jump was applied to the immobilized DNA duplex. When the temperature was high enough to completely denature the duplex, the unimmobilized strand could dissociate and diffuse away. The immobilized strand folded back to the hairpin structure, where donor and acceptor become close and give high FRET. The percentage of the dissociated DNA duplex was calculated at different temperatures to give a single-molecule melting curve. The melting temperature can be obtained by fitting to the equation 10.

$$f(T) = f_0 + (f_1 - f_0)\frac{T^n}{T_c^n + T^n} \qquad \text{Eq (3.1)}$$
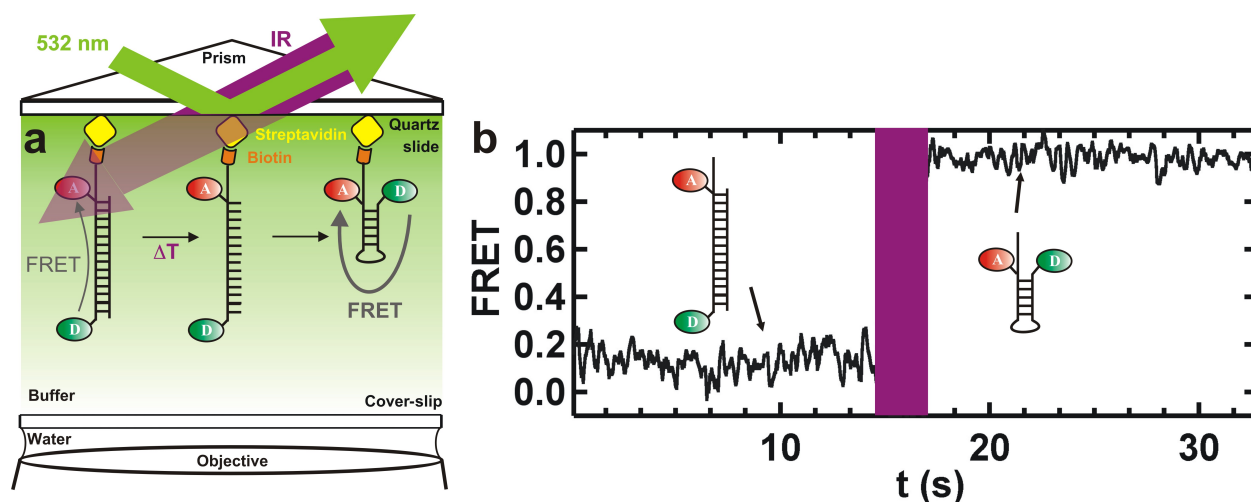
Figure 3.1 Schematic plot of LASR applied to DNA duplex melting. a. DNA duplex was form from pre-heat annealing then immobilized as described. The immobilized strand is labeled with Cy3 and Cy5. After the dissociation of unimmobilized strand, the DNA will fold into hairpin structure. **b.** Typical FRET time course of dissociation of DNA duplex after the T-jump. Reprinted from Ref. [73].

3.3 Results and discussion

In figure 3.1b, the molecule appears at 0.15 FRET before the T-jump (purple), which corresponds to the extended duplex structure. This structure is stable at room temperature. After the T-jump, the FRET ratio is 1.0 and remains static for a long time, corresponding to the dissociation of the DNA duplex and the formation of the stable hairpins. This proves that we can drive the dissociation reaction of the DNA duplex and monitor it at the single-molecule level. The experiment was also purposely designed to change the FRET ratio from low to high to eliminate the complication of the photo-bleaching effect of the dye molecules.

Also we can see in figure 3.2, the UV melting curve and the single molecule melting curve overlap very well with each other for all three DNA duplexes. The fitted melting temperature from the single-molecule melting curves are T_m = 51 ± 1, 61 ± 1 and 70 ± 1°C, respectively. Both the UV melting experiment and the single molecule

melting experiment are done in the same salt conditions. The nice overlapping of the two melting curves strongly supports the accuracy of the temperature calibration of our technique.
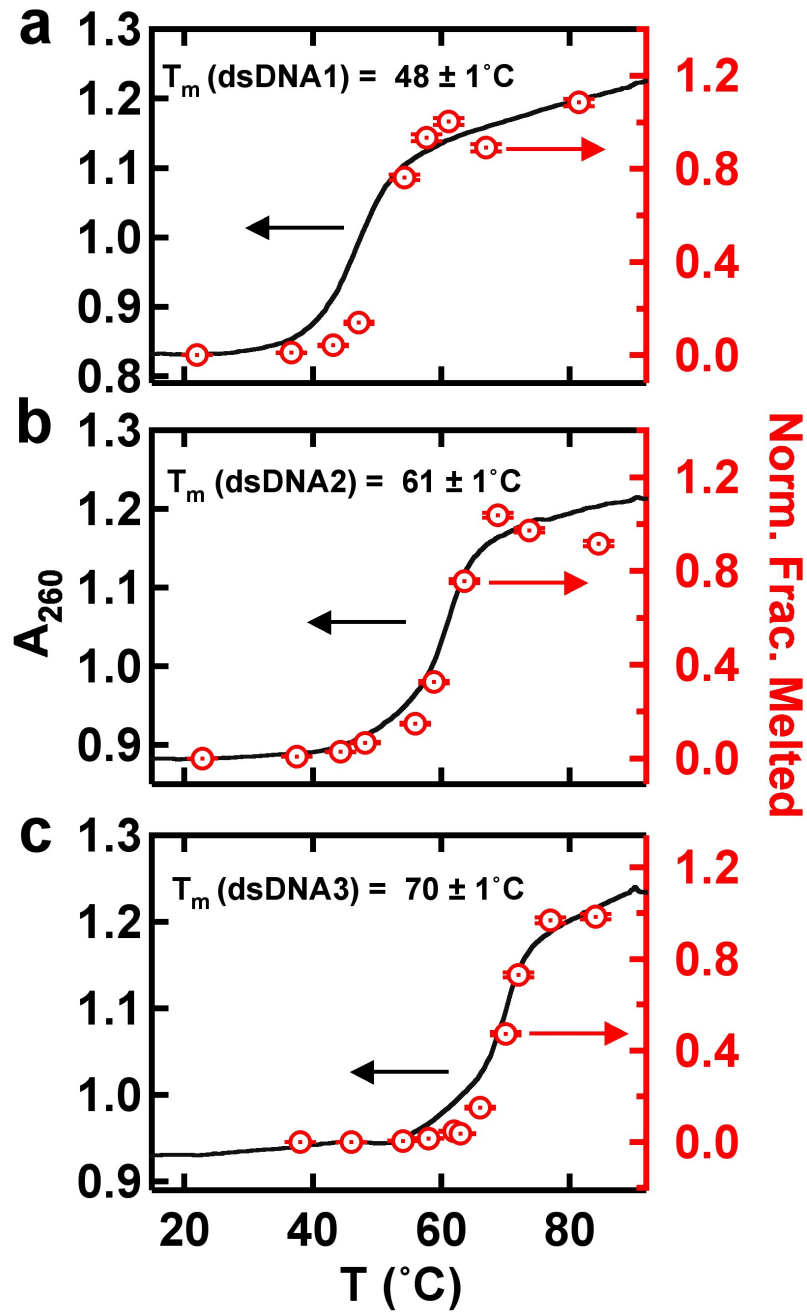
Figure 3.2 Melting curve of three DNA duplexes from UV melting (black) and LASR (red) experiments. Both experiments at done in the same conditions: 50 mM TRIS-HCl pH 7.5, 10 mM MgCl$_2$. Reprinted from Ref. [73].

3.4 Conclusion

UV melting is a bulk measurement the may be composed of multiple reactions that cannot be distinguished by simply monitoring the absorbance at 260 nm. The Van't Hoff analysis of the UV melting curve assumes a two-state model and equilibrium between the states; this could have introduced some error for the melting temperature. Luckily, in the gel-shifting assay, a major band is shown for the DNA duplex formation, which indicates that the bulk measurement more or less represents the simplified microscopic reactions.

In the single-molecule melting assay, we are directly observing the dissociation of the DNA duplex. This should be a more accurate measurement of the melting temperature. However, the 1 s heating time may not be long enough for the system to reach equilibrium. Instead, it is a measurement of the activation energy barrier. Since all the molecules that overcome the kinetic barrier will proceed to the dissociated duplex, the percentage of the dissociated molecules is actually the percentage of the molecules that overcome the kinetic barrier. The enthalpy distribution of the molecules is a Maxwell-Boltzmann distribution. The single-molecule melting curve can be fit to determine the activation energy barrier.

As discussed above, the two experiments may give different result in the ideal case since they are two different types of measurement. However, practically the UV melting experiment was carried out with a temperature-increasing rate of 1 °C/min. At this rate, for long DNA duplexes with long length (>18 bp), the equilibrium assumption in the Van't Hoff analysis is no longer true at low temperatures (<50°C).
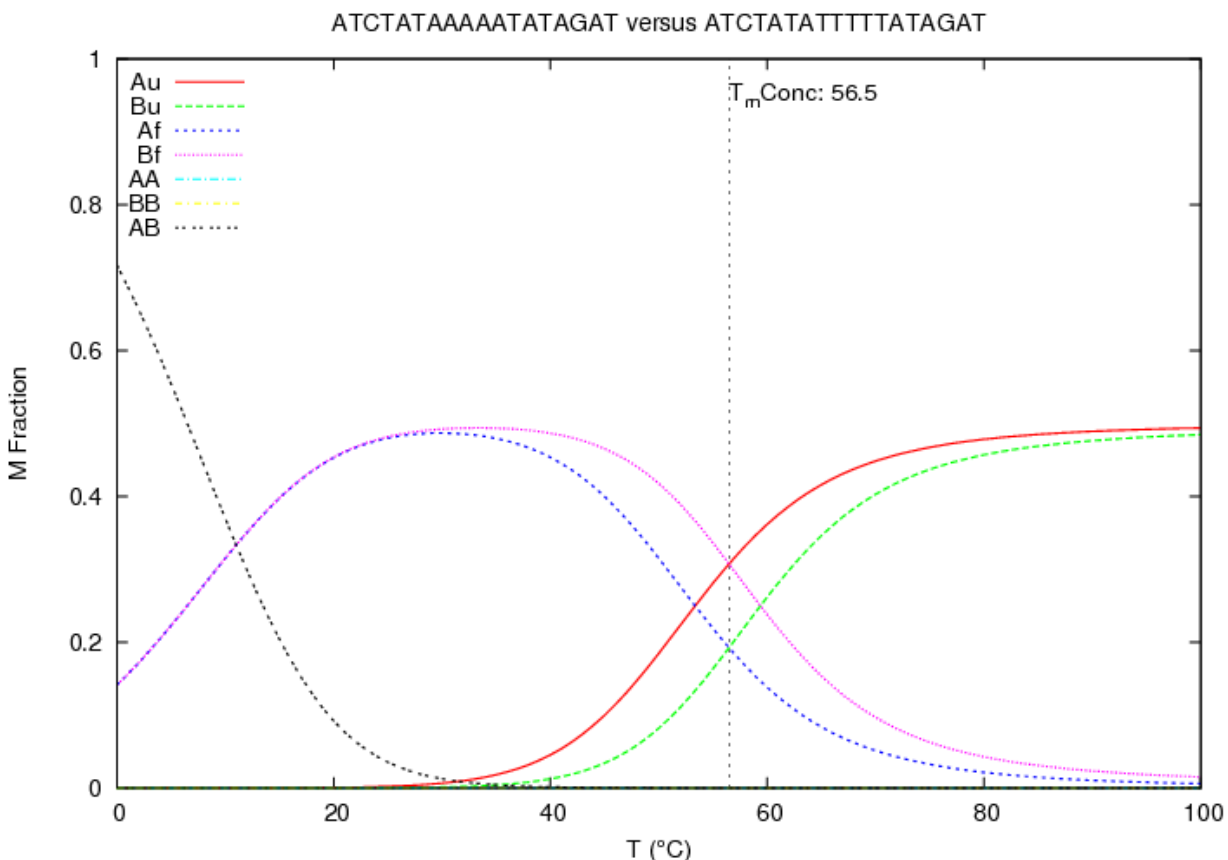
Figure 3.3 Melting curve of DNA duplex 1 predicted by Mfold[74]. Au and Bu are unfolded single strand A and B monomer. Af and Bf are folded single strand A and B monomer. AA, BB and AB are dimers. Melting curves for all specie are indicated by colors.

The predicted melting temperature for 10 nM DNA duplex 1 is below 20 °C (Fig 3.3), which is much lower than measured value (Fig 3.2 top). However, in single-molecule FRET experiments, we don't see dissociation of the DNA duplex in our experimental time windows (300 s) at 40°C. In conclusion, both UV and single-molecule melting curves are actually measuring the height of activation energy barrier for the dissociation of the DNA duplex. The UV melting experiment has a longer equilibrating time (1 min), which shifts the melting curve to the lower temperatures (Fig. 3.2). This shifting effect is more obvious for DNA duplexes with lower melting temperature, as the

activation energy is lower and so easier to overcome.

In summary, we have obtained both the melting curve from UV melting experiments and single-molecule experiments and they overlap very well with each other. This demonstrates the accuracy of the temperature measurement of our technique and applicability to macromolecular systems.

(Chapter 3 is adopted from paper *Zhao, R. et al. Laser-assisted single-molecule refolding (LASR). Biophys J **99**, 1925-1931, doi:S0006-3495(10)00889-1*[73])

CHAPTER 4

Single molecule characterization of HIV DIS RNA

4.1 Introduction

In the HIV viral replication cycle, during the viral encapsidation process, two identical HIV viruses have to form a dimer via the dimerization initiation sequence (DIS). Before the dimerization reaction takes place, two DIS RNAs can interact with each other in the loop region to form a stable intermediate state, known as the kissing complex (Fig. 4.1 left)[75,76]. This intermediate structure is so stable that it can be crystallize it[77]. Figure 4.1a shows the crystal structure of the HIV DIS kissing complex. As we can see, the kissing complex is formed through six base pairs, which forms a mini A type helix. Magnesium ions (yellow balls) are also important to stabilize this structure. There are different mechanisms proposed for the dimerization reaction, which may or may not go through the kissing complex structure. This process is very important as it provides a potential target to stop HIV viral replication as a cure for AIDS. However, the mechanism of the dimerization reaction is still not fully understood. Here we are elucidating dimerization reaction of the HIV DIS RNA using single molecule fluorescence detection. By examining one molecule at a time, we can monitor the dimerization reaction in real time. These experiments have yielded mechanistic insight into kissing complex-mediated dimerization.
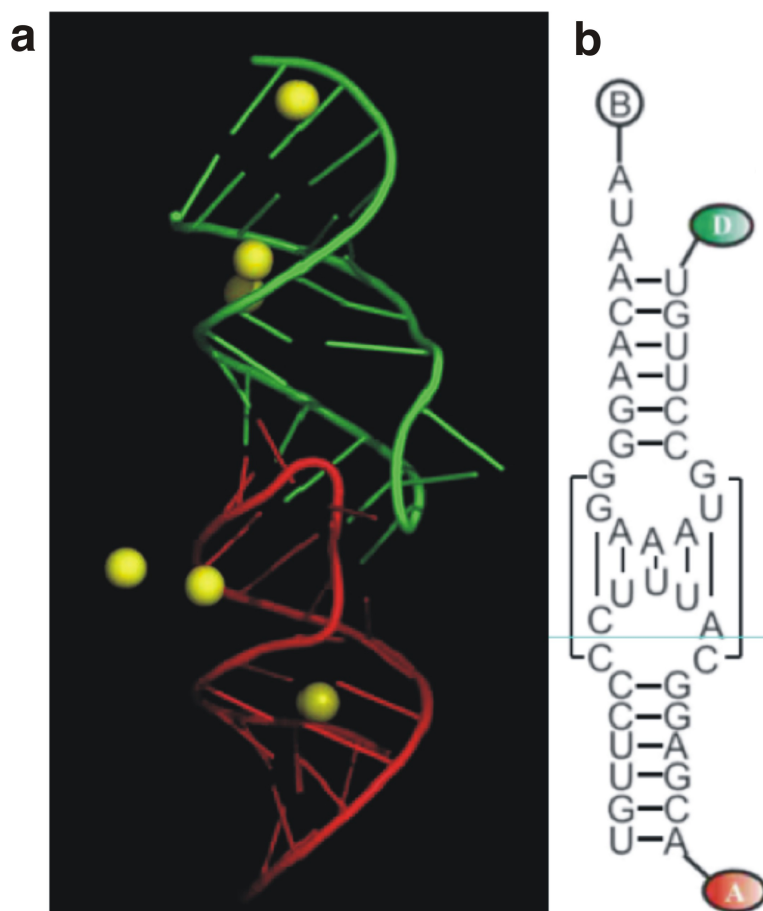
Figure 4.1 Structure of kissing complex a. 3D structure of the HIV DIS RNA kissing complex[77]. Green and red RNA strand are identical in sequence. Yellow balls are Mg2+ that is important for stabilization of the kissing hairpin **b.** Secondary structure of kissing complex used in this study. B: 5' biotin. Green and red ovals: FRET donor and acceptor fluorophores. Lines indicate the base pairs in the loop.
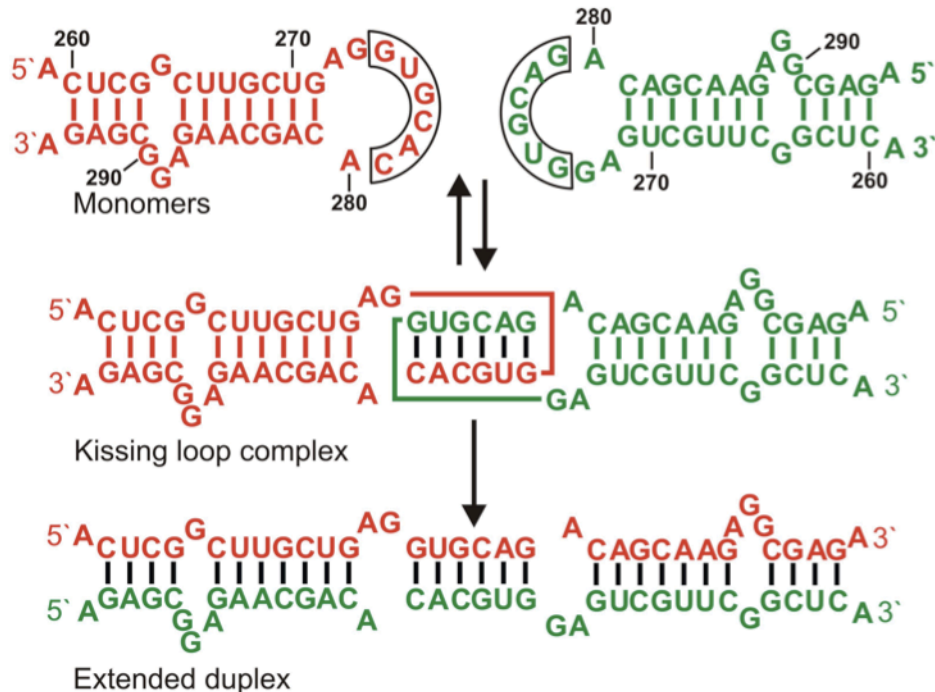
Figure 4.2 Reaction pathways for the formation of HIV DIS RNA kissing complex and extended duplex. Adopted from Ref. [78].

4.2 Methods

For simplicity and to prevent the formation of the homodimer, we used a engineered model substrate based on DIS RNA. This RNA was designed and well characterized in Dr. Andrew Feig's Lab (Wayne State University, Chemistry). The kissing complex is stabilized by 7 base paring interactions between the two hairpins, but does not form homo-dimers. The stems are complementary and can form an extended duplex. For our single molecule fluorescence studies, the top hairpin (HPII) was labeled with a 5'-biotin and 3'-FRET donor (Cy3). The bottom hairpin (HPI-BP) has been labeled with a 5'-Cy5.

HPII was immobilized on the quartz slide via a biotin-streptavidin bridge. Excess HP1-BP was then flushed into the channel and allowed to freely diffuse in the

microfluidic channel. It can form a kissing complex or extended duplex with HPII (Fig. 4.2). The conformational fluctuations can be monitored using FRET. When only HP-II was present, we only see a peak at 0 FRET (Fig 4.3D top).

4.3 Results and Discussion

When two hairpins without kissing loop were present, we still only see 0 FRET (Fig 4.3d top) as there were no kissing complex or extended duplexes formed. When complementary hairpin RNAs HP-II and HPI-BP were preannealed before the immobilization, then we see two peaks at 0 and 1.0 FRET (Fig. 4.3d third panel). When the kissing complex was formed, as the distance between the dyes molecules on the ends of the hairpins decreased, we will see an increase in FRET (0 to 0.5) and vice verse.

As shown in figure 4.3b, bottom, the FRET oscillates randomly between 0 and 0.5. This trace allows real time monitoring of formation and dissociation of the kissing complex. When the extended duplex was formed, FRET donor and acceptor are close to each other, which would give a high FRET (Fig 4.3c FRET=1). This dimerization reaction does not occur often at room temperature in our observation time window (minutes). Only 3 out of >600 molecules show transition to the extended duplex conformation and it is not clear whether or not the kissing complex is an obligatory intermediate.

The formation of the kissing complex and the extended duplex can be demonstrated by the following control experiments. The histogram from the experiments that include only donor labeled hairpin (HPII) has a single peak at low FRET (~0). The histogram of the control experiment with hairpins that cannot form neither a kissing

complex nor and extended duplex has only one peak at FRET=0 as no kissing complex or extended duplex are formed. The pre-annealed complementary hairpins give two peaks, at FRET = 1 and 0 respectively, which corresponds to the formation of the extended duplex and those with photo-bleached Cy5 or donor only RNA hairpin. In contrast, when HPII is immobilized and 35 nM HPI is free diffusing acceptor labeled hairpins, three peaks appear in the histogram: FRET = 0, 0.5 and 1. The 0 FRET peak could be donor only hairpin or the Cy5 photo-bleached extended duplex. The 1 peak is the extended duplex, which accumulates with time. The 0.5 FRET peak was assigned for the kissing complex as the FRET value agrees with the distance. In addition, when we increase the length of one of the stems we see the peak shift to lower FRET value (0.35) (Fig 4.4b). The FRET value further drops to 0.2 (Fig. 4.4d) when we increase the length of both stems.
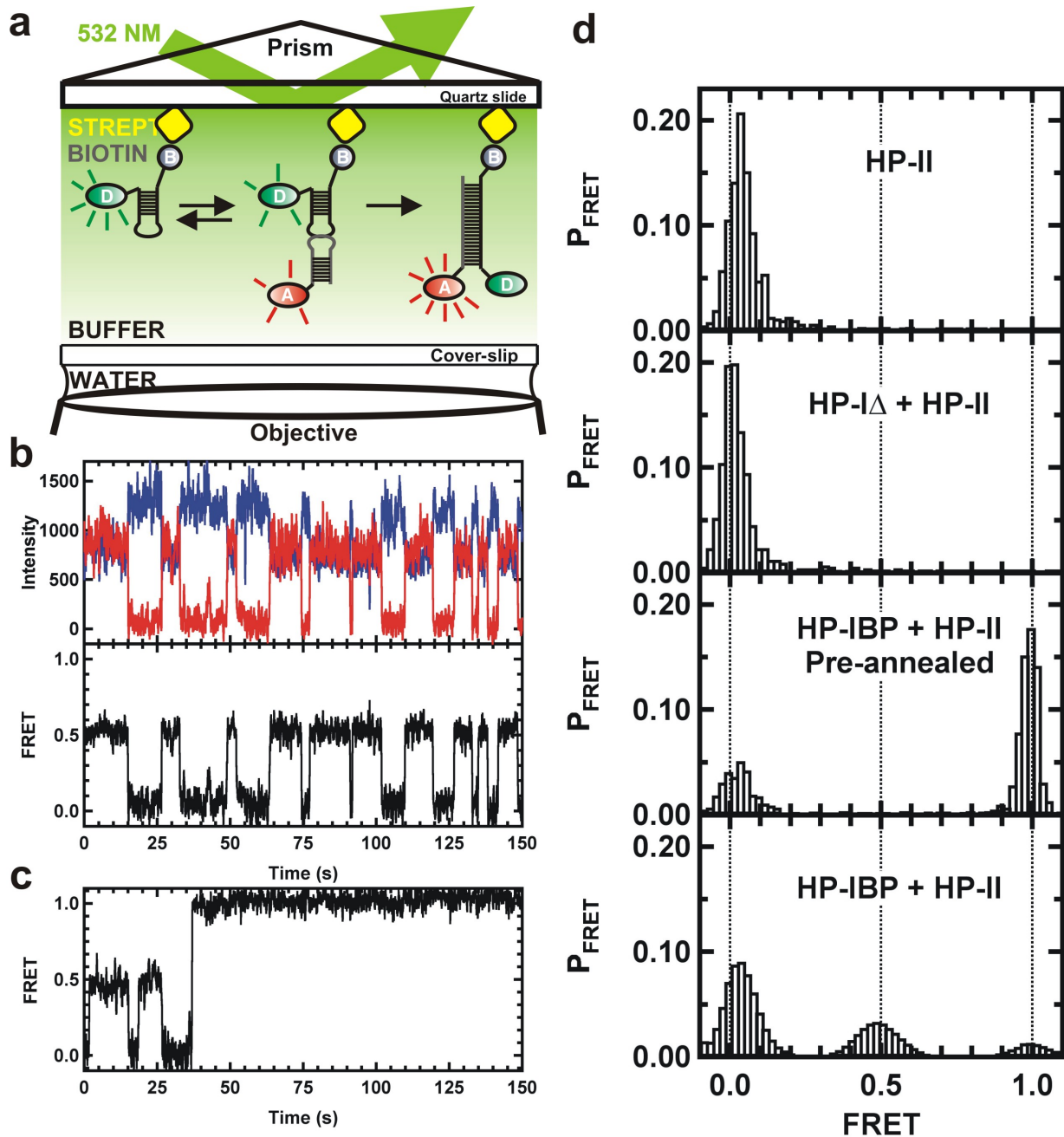
Figure 4.3 Single-molecule study of kissing complex a. Schematic plot of single-molecule FRET study of kissing hairpins. Donor labeled RNA hairpin was immobilized and 35 nM of acceptor labeled RNA were in the background. **b.** Top: Donor and acceptor emission intensity of single molecule kissing complex formation and dissociation. Bottom: FRET time course calculated from top panel. **c.** One single-molecule FRET time trajectory shows a transition to the 1.0 FRET state. **d.** FRET histograms of the RNA samples as indicated in the annotation.
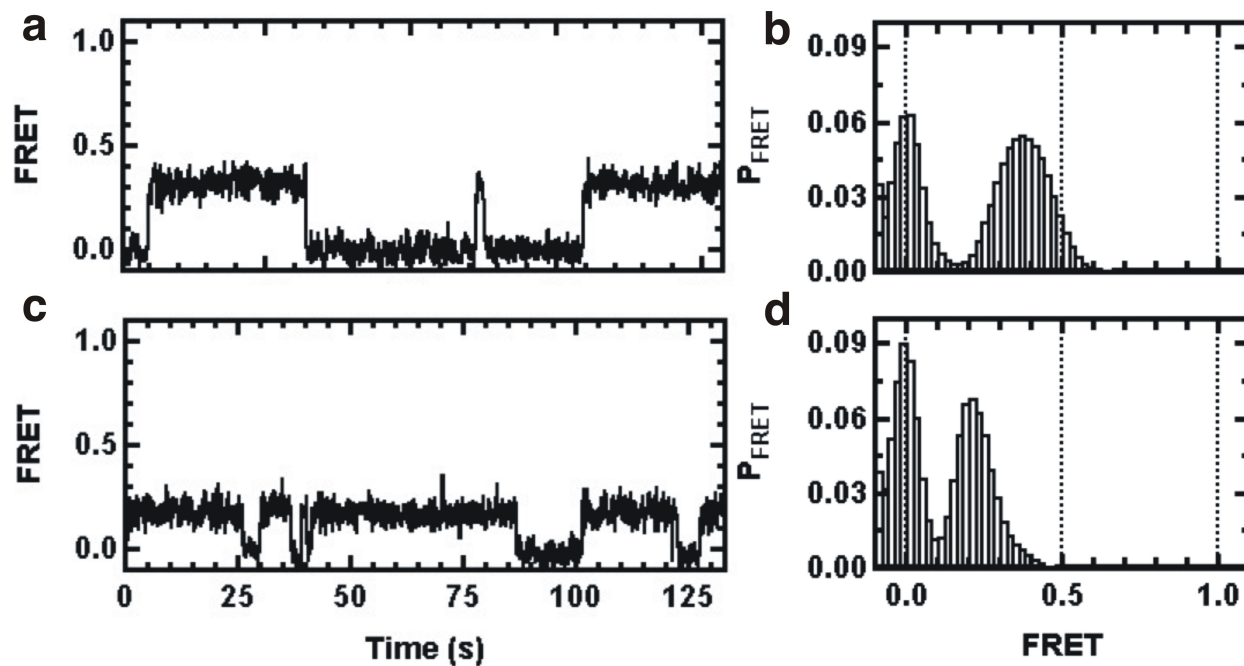
Figure 4.4 Control experiment of mutants with longer stems HP1D and HP2D. a. Single molecule FRET time course of HP2+HP1D. **b.** FRET histogram of HP2+HP1D. **c.** Single molecule FRET time course of HP2D+HP1D **d.** FRET histogram of HP2D+HP1D.

To characterize the stability of the loop-loop kissing interaction we have designed different mutations in the loop region for to obtain the association and dissociation rate of the kissing complex and hence the free energy of the formation of kissing complex. For the parent hairpins, the dwell time analysis of 'on' and 'off' states yields association and dissociation rates as 3.42×10^6 $s^{-1}M^{-1}$ and 0.18 s^{-1}, respectively. The free energy of the kissing complex formation was calculated as -9.9 kcal/mol.
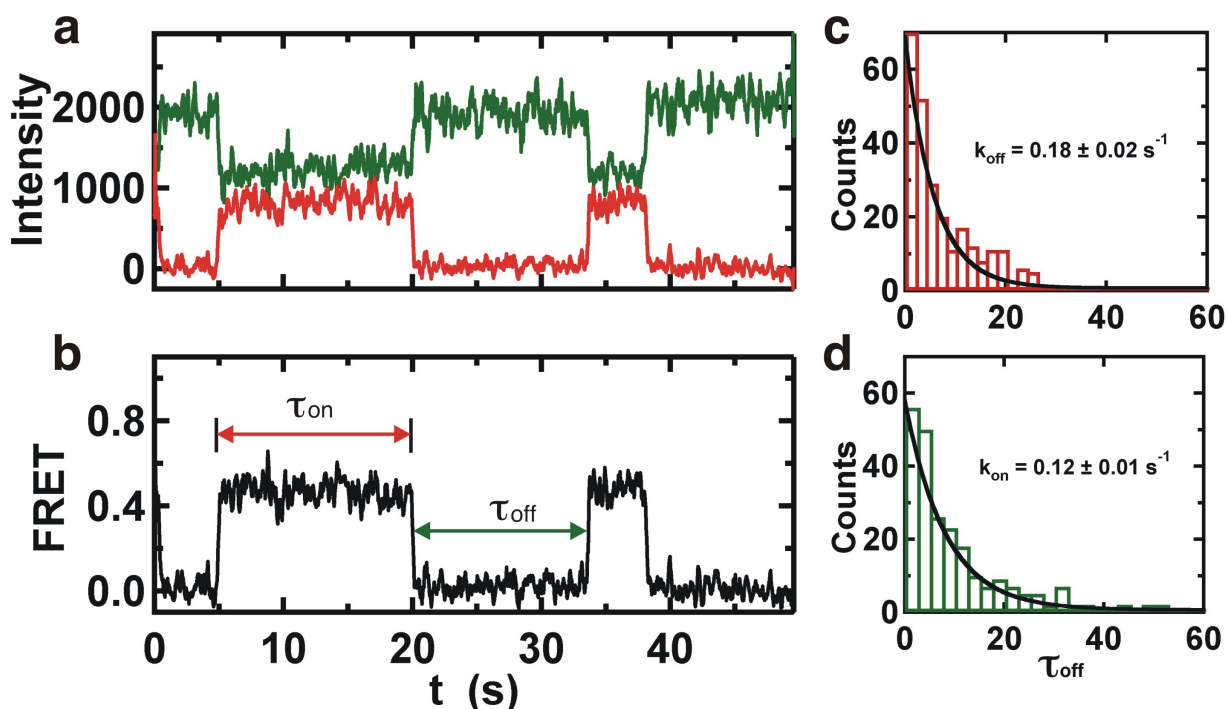
Figure 4.5 Dwell time analysis of HP2+HP1BP to calculate k_{on} and k_{off}. a. FRET donor and acceptor emission intensity of HP2+HP1BP. **b.** FRET time course of HP2+HP1BP. **c.** Single exponential fit to the histogram to calculate k_{off}. **d.** Single exponential fit to the histogram to calculate k_{on}.

In the first mutation, we changed the UUU in the middle of the loop to CCC. This way, we break three base pairs. We don't see kissing at all for the mutation, which indicates that four base pairs are not enough to stabilize this kissing interaction. In the second mutation, we change the AU base pair in the middle of the loop to a CG base pair. By doing this, we are adding one extra hydrogen bond to the kissing loop interaction. As we can see in figure 4.6a, there is an obvious increase in the lifetime in the kissing complex structure (FRET = 0.5). The histogram also shows an increased peak height for the kissing complex. The dwell time analysis gives the same association rate k_{on}, as the parent hairpins because it is diffusion limited. The dissociation rate decreased from 0.18 to 0.07 s^{-1}, presumably due to the stabilization from the extra hydrogen bonds. The free energy of the kissing complex formation was calculated to be

10.5 kcal/mol. A net stabilization effect of 0.6 kcal/mol from the one extra hydrogen bond is a reasonable number at room temperature.
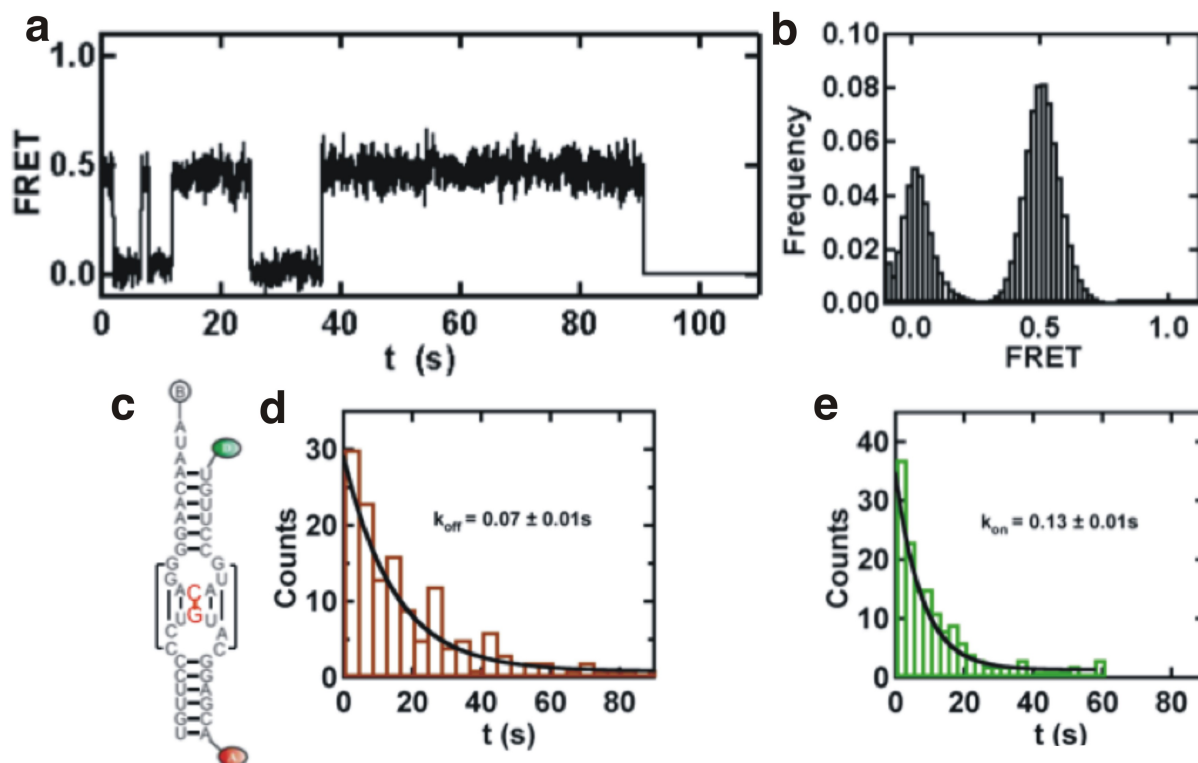
Figure 4.6 Single molecule FRET study of GC mutant. a. FRET time course of GC mutant. **b.** Histogram of GC mutant. **c.** Secondary structure of GC mutant. **d.** Dwell time analysis to calculate k_{off}. e. Dwell time analysis to calculate k_{on}.

In the third mutation, we changed the GC base pair in the loop near the 5' end of the acceptor labeled hairpin RNA to a GA mismatch. By doing this we have reduced up to 3 hydrogen bonds in the kissing interaction. As we can see in figure 4.7, the molecules stays in the 0.5 FRET state for only a very short time, which means the kissing complex is very unstable. The FRET histogram shows a dramatically decreased peak near 0.5 FRET state. The dwell time analysis yields association and dissociation rates as 0.16 and 6 s^{-1}, respectively. Again the association rate is not changed, as it is diffusion limited. However, the dissociation rate increased by ~50 fold. The free energy

of the kissing complex formation was calculated as -8.0 kcal/mol. The GA mismatch has destabilized the kissing complex by 1.9 kcal/mol, which indicates all three base pairs are disrupted.
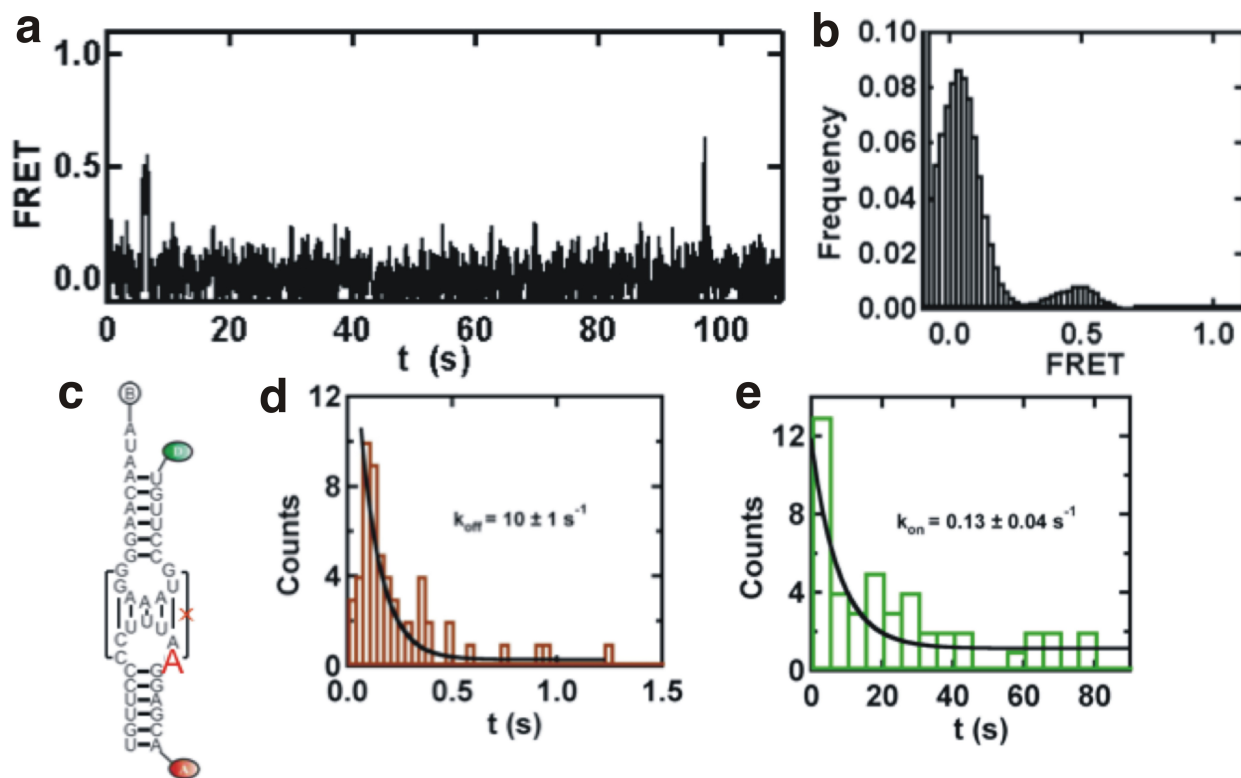
Figure 4.7 Single molecule FRET study of CA1 mutant. a. FRET time course of CA1 mutant. **b.** Histogram of CA1 mutant. **c.** Secondary structure of CA1 mutant. **d.** Dwell time analysis to calculate k_{off}. **e.** Dwell time analysis to calculate k_{on}.

In the forth mutation, the GC base pair in the loop near the 3' end of the acceptor labeled hairpin RNA was mutated to a GA mismatch. Similarly, up to three hydrogen bonds were broken in the kissing interaction. Also the molecules stay in the 0.5 FRET state for a short time and the 0.5 FRET peak in the histogram shrink to nearly 0. The dwell time analysis yields association and dissociation rates as 0.13 and 10 s^{-1}, respectively. The association rate is not changed much. But the dissociation rate increased by 30 fold. Only slight directionality for the formation of the kissing complex

was shown here (30 fold VS 50 fold). The free energy of the kissing complex formation is determined as -7.6 kcal/mol. The GA mismatch mutation near the 3' end destabilizes the kissing complex slightly more than the other end. The kinetic parameters are summarized in the following table.
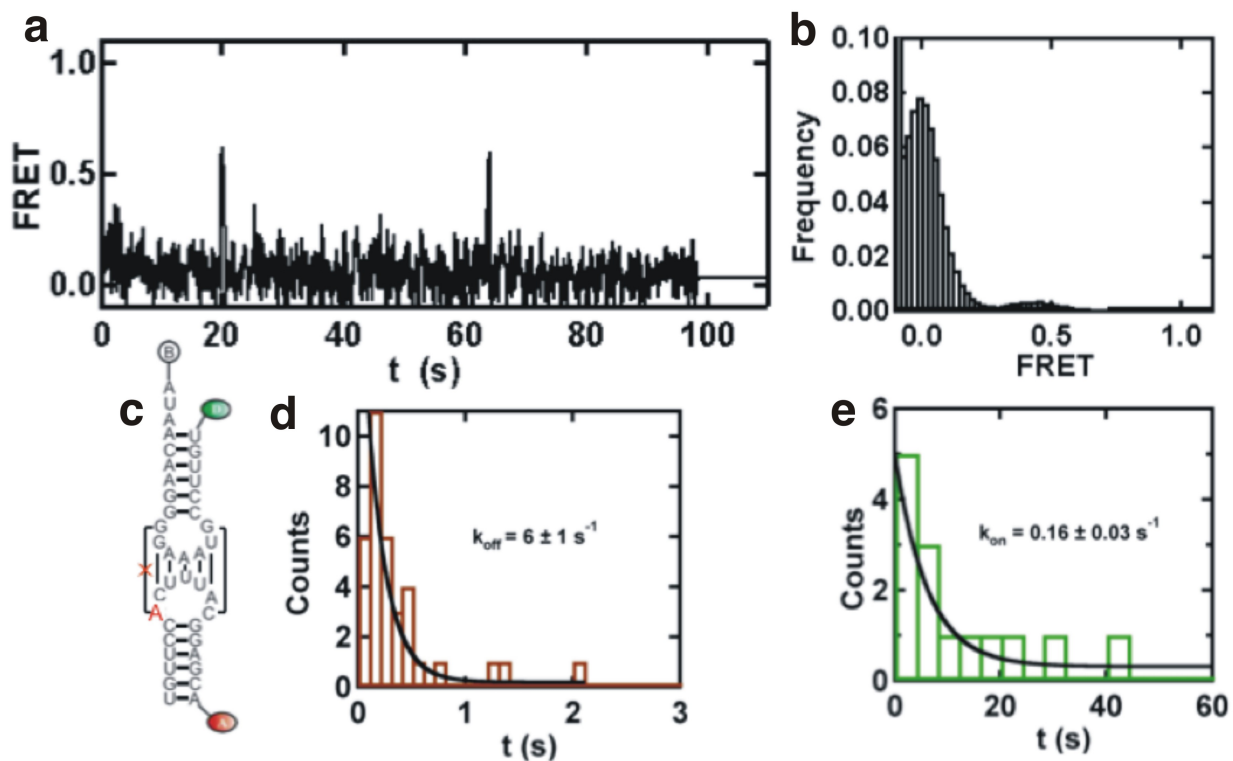
Figure 4.8 Single molecule FRET study of CA2 mutant. a. FRET time course of CA2 mutant. **b.** Histogram of CA1 mutant. **c.** Secondary structure of CA12 mutant. **d.** Dwell time analysis to calculate k_{off}. **e.** Dwell time analysis to calculate k_{on}.

Table 4.1 Kinetic information of the kissing complex and mutants.

Hairpins	k_{on} (s^{-1})	k_{off} (s^{-1})	ΔG^0_{298} (kcal/mol)	$\Delta\Delta G^0_{298}$ (kcal/mol)
WT	0.12± 0.01	0.18± 0.02	-9.9± 0.2	-
A3→C3	No kissing	No kissing	-	-
AU→GC	0.13± 0.01	0.07± 0.01	-10.5± 0.2	-0.6± 0.3
5'C→A	0.16± 0.03	6± 1	-8.0± 0.2	1.9± 0.3
3'C→A	0.13± 0.04	10± 1	-7.6± 0.3	2.3± 0.4

4.4 Conclusions

Single molecule study has allowed resolution of hairpin, kissing complex, and extended duplex that would not be possible in bulk experiments. This has allowed more detailed dissection of kissing complex formation. For a stable interaction with nanomolar binding affinity, the molecules do not remain stably bond all for long time, especially at low concentrations. This mutational study shows that 7 base pairs are necessary to stabilize this kissing interaction in the salt condition: 50 mM TRIS pH 7.5, 10 mM Mg^{2+}. The seven base pairs in the loop region are all involved in the kissing interaction and there is only slightly directionality. However, the dimerization reaction was barely observed at room temperature in our experimental time windows. External force is needed to drive the dimerization reactions.

CHAPTER 5

Laser-Assisted Refolding of HIV DIS RNA

5.1 Introduction

To become active, biological macromolecules such as proteins, RNA or DNA, must fold from a linear chain into a native three-dimensional structure following pathways determined by their folding potential energy surfaces. However, these potential energy surfaces can be rugged and rife with local energy minima that confine the biopolymer into long-lived intermediate states called kinetic traps[13]. Kinetic traps can impede biological function, because they prevent the formation of the native state. In vivo, kinetic traps can be resolved by refolding the biopolymer with the help of chaperones, helicases and other cellular folding co-factors[17,19,20,79]. To understand the macromolecule refolding problem, it is necessary to characterize kinetic traps by determining their activation barriers, but this can be very challenging because the barrier may be significantly higher than thermal energy at room temperature (k_BT).

Single-molecule microscopy-based approaches are ideally suited to study kinetic traps, because they enable their identification even in the presence of other conformations in solution. Single-molecule Arrhenius and Eyring analyses of folding rate constants can be used to determine energy barriers but, for technical reasons, are limited to temperatures below ~50˚C, which may not be high enough to melt many nucleic acid folds. Furthermore, Mg^{2+} ions, which play an important role in RNA folding[10,11], can cleave the backbone of RNA molecules at high temperatures (>50˚C) making it impractical to study deep kinetic traps in their presence.

Here, we report a new single-molecule technique, Laser Assisted Single-Molecule Refolding (LASR), which enables characterization of the activation barriers of deep kinetic traps by combining traditional single-molecule detection with infrared laser (IR) induced temperature jumps (T-jump)[69,80-82]. An IR laser is used to excite the second overtone of the water OH stretch (~1400 nm)[71], and thus, rapidly increase the solvent temperature. As a result, kinetically trapped biopolymers in solution receive sufficient thermal energy to overcome the refolding activation barrier. Our T-jumps are transient, but reversible, temperature increases and thus differs from traditional ones. We demonstrate the application of LASR to measure single-molecule melting curves of double stranded DNAs (dsDNAs) and to characterize a model kinetic trap comprised of two complementary hairpin RNAs that can form a stable loop-loop interacting complex (kissing complex).

5.2 Methods

5.2.1 RNA labeling and purification. RNA samples were purchased from the W.M. Keck synthesis facility at Yale University with a 5'-biotin (B) and a 3'-amino linker for Cy3 labeling (HPI and HPIII, table 5.1) or with a 5'-Cy5 (HPII and HPIV, table 5.1). All RNAs samples were purified and labeled as described[83].

5.2.2 LASR experiments. Temperature Jumps were generated using a broadband supercontinuum fiber laser (SC450, Fianium, Eugene, OR). IR radiation was selected using a long-pass filter (RG665, Schott, Cambridge, UK) and introduced through the total internal reflection prism, but at an angle more acute than the total internal reflection angle. A small portion of red light was kept for alignment purpose. When not used, the IR laser was blocked using a computer driven shutter. The shutter

was opened for 1 s for heating. Neutral density filters were used to attenuate the IR laser power, and control the jump temperature. The CCD camera was blocked during heating to avoid damaging the CCD chip. Multiple experiments at various jump temperatures were performed to obtain the LASR curves (Fig. 5.2), and then fit to the following hyperbolic equation

$$f(T) = f_0 + (f_1 - f_0)\frac{T^n}{T_c^n + T^n}$$
(Eq. 5.1)

where, $f(T)$ is the fraction reacted at a temperature T, f_0 and f_1 are the initial and final fractions reacted, respectively, T_c is the critical temperature (refolding or dissociation) and n is a cooperativity factor.

5.2.3 Temperature Calibration. The jump temperature was calibrated using a home-built micrometer-size gold sensor. The sensor was fabricated on a microscope slide and assembled into a micro-fluidic channel in standard buffer (Fig. 2.15). The experimental geometries for the gold wire and the sample were maintained almost identical, and therefore, only small differences in thermal properties are expected. IR laser radiation was introduced through the total internal reflection prism, but at an angle more acute than the total internal reflection angle. The sensor's electric resistance was measured with a multimeter (34401A, Agilent, Santa Clara, CA). With the sensor on top of the microscope objective, the laser was aligned by maximizing the sensor's electrical resistance using a fine-tune translation stage with micrometer accuracy. Then, the IR laser was blocked using a computer driven shutter. The shutter was opened for 1 s to heat the sensor, and its electrical resistance was recorded. Different neutral density (ND) filters were used to attenuate the IR laser power, and control the jump temperature. The

IR laser intensity in the H_2O absorption bands was determined by measuring the laser intensity before and after a 1 cm water cuvette. The sensor's electrical resistance increase was then converted to temperature using figure. 2.14. The jump temperature should be calibrated prior to each LASR experiment.

5.2.4 Determination of Activation Barrier from LASR experiments. To determine transition state parameters, we begin with the assumption that, during the temperature jump, the kissing complex (KC) is in equilibrium with an activated complex ($TS^\ddagger$) to form the extended duplex, similarly to Eyring's Transition State theory:

$$KC \rightleftarrows [TS^\ddagger] \rightarrow ED \qquad \text{(Eq. 5.2)}$$

The equilibrium constant between the KC and $TS^\ddagger$ can be defined as

$$K^\pm = \frac{\left[TS^\pm\right]}{\left[HP\right]} \qquad \text{(Eq. 5.3)}$$

Once the reactant reaches across the transition state, it proceeds to form the extended duplex. Therefore, we can estimate [TS] as the fraction of molecules that form the extended duplex at a given jump temperature (f_{ED}), whereas [HP] can be estimated as the fraction of molecules that did not form the extended duplex (1 - f_{ED}). The equilibrium constant can be rewritten as:

$$K^\pm = \frac{f_{ED}}{1 - f_{ED}} \qquad \text{(Eq. 5.4)}$$

Thus, (Eq. 4) can be rearranged as:

$$f_{ED} = \frac{K^\pm}{1 + K^\pm} \qquad \text{(Eq. 5.5)}$$

From the LASR melting curves (Fig. 5.2), f_{ED} is determined experimentally as:

$$f_{ED} = \frac{f(T) - f_0}{f_{max} - f_0} \qquad \text{(Eq. 5.6)}$$

where $f(T)$ is the fraction reacted at a temperature T, f_0 is the fraction reacted at low temperature and f_{max} is the maximum fraction reacted. The fraction reacted, $f(T)$, can be expressed as:

$$f(T) = f_0 + (f_{max} - f_0)f_{ED} = f_0 + (f_{max} - f_0)\frac{K^{\pm}}{1 + K^{\pm}} \qquad \text{(Eq. 5.7)}$$

The constant, $K^{\ddagger}$, can be expressed in terms of transition state Gibbs free energy as

$$K^{\pm} = \exp\left(-\frac{\Delta G^{\pm}}{RT}\right) \qquad \text{(Eq. 5.8)}$$

The temperature dependence of $\Delta G^{\ddagger}$ can be obtained using the Gibbs-Helmotz equation,[84]

$$\Delta G^{\pm} = \Delta H_r^{\pm}\left(1 - T/T_r\right) \qquad \text{(Eq. 5.9)}$$

where $\Delta H_r^{\ddagger}$ is the activation energy barrier for the refolding reaction and T_r is the refolding temperature obtained using (Eq. 1). Finally, the fraction reacted during a LASR temperature jump, $f(T)$, can be expressed as

$$f(T) = f_0 + (f_{max} - f_0)\frac{e^{-\frac{\Delta H_r^{\pm}}{RT}(1 - T/T_r)}}{1 + e^{-\frac{\Delta H_r^{\pm}}{RT}(1 - T/T_r)}} \qquad \text{(Eq. 5.10)}$$

This expression was used to fit the LASR curves in Fig. 5.2 to determine the activation energy barrier of extended duplex formation. A similar expression can be derived for the dissociation reaction.

5.2.5 Determination of Activation Barrier Parameters by Eyring Analysis.

The dissociation activation barrier can be obtained by measuring the temperature dependence of the kissing complex dissociation rate constants (k_{off}) and using Eyring analysis[85]. The kinetic rate constant was obtained by fitting the distribution of dwell times in the kissing complex to a single exponential decay, as previously described[28]. Using a microscope stage temperature controller, the rate constants k_{off} was determined at temperatures ranging from 15 to 23 ˚C. At higher temperatures, the number of molecules forming the kissing complex decreased dramatically. The resulting rate constants were linearized in an Eyring plot (Fig. 5.5) and fit to Eyring's equation):

$$\ln\left(\frac{k_{off}h}{k_B T}\right) = -\frac{\Delta H_d^{\pm}}{R}\frac{1}{T} + \frac{\Delta S_d^{\pm}}{R} \qquad \text{(Eq 5.11)}$$

to obtain the activation energy barrier $\Delta H_d^{\ddagger}$.

Table 5.1RNA sequences used in this study.

Name	Description	Sequence[1,2]
HPI	RNA hairpin 1	5' B-AUA ACA AGG GGA <u>A</u>AU GCC UUG U-3' Cy3
HPII	RNA hairpin 2	5' Cy5-ACG AGG CAU <u>U</u>UC CCC UUG U-3'
HPIII	RNA hairpin 3	5' B-AUA ACA AGG GG**C CC**U GCC UUG U-3' Cy3
HPIV	RNA hairpin 4	5' Cy5-**UGU UC**G CAU UUC CC**G AGC A**-3'

[1] B = biotin, Cy3 and Cy5 are linked to the nucleic acid by a 6-carbon amino linker.
[2] The underlined bases in HPI and HPII were also mutated to G and C, respectively.
[2] Bold bases in HPIII and HPIV are modifications from HPI and HPII, respectively.

5.3 Result and Discussion

5.3.1 LASR applied to study a model kinetic trap

We then sought to apply LASR to study the dimerization reaction of two complementary hairpin RNAs that can form a stable kinetic trap, the kissing complex (Fig. 5.1a). Loop-loop, or kissing, interactions are ubiquitous tertiary structure stabilizing elements in RNA[77,86-89]. Two RNA hairpins (HPI and II, Table 5.1) were designed to have complementary loops that form stable hairpins, to form stable extended duplexes and to prevent formation of homodimers. HPI was labeled with a 3'-Cy3 and a 5'-biotin for surface immobilization, and HP2 was labeled with a 5'-Cy5. The labels are designed to yield 0 FRET for the free hairpins, intermediate FRET when the kissing complex forms, and high FRET in the extended duplex conformation (Fig. 5.1a). This kissing complex behaves as an ideal model for kinetic traps, because formation of the extended duplex is very slow compared to kissing complex formation. Before a LASR T-jump, surface immobilized HPI can readily bind and dissociate HPII in solution (35 nM HPII, 50 mM TRIS-HCl, 10 mM $MgCl_2$) to form the kissing complex, as

evidenced by the presence of random excursions to 0.45 FRET in the single-molecule time trajectories (Figs. 5.1b-e). We have previously determined the kissing complex formation and dissociation rate constants at room temperature (k_{on} and k_{off}) to be 0.12 and 0.17 s^{-1}, respectively (chapter four). Only 3 out of >600 observed trajectories exhibit a transition to the extended duplex conformation in the absence of T-jumps, as expected for this model kinetic trap.
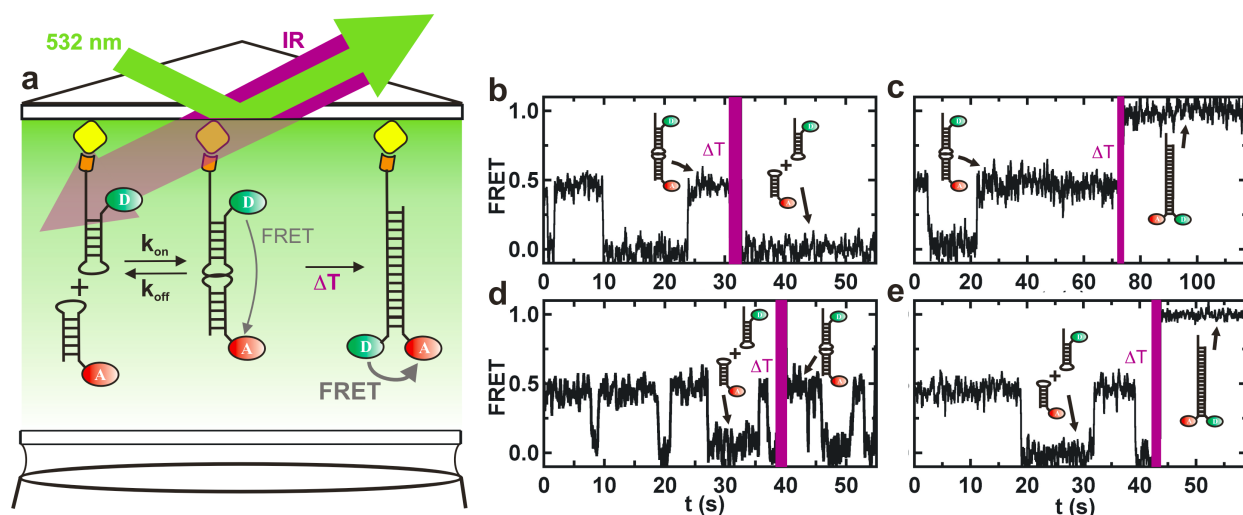
Figure 5.1 LASR study of kissing hairpins. a. Schematic plot of LASR study of RNA kissing hairpin. Single molecule FRET study was carried as previously described. Purple arrow indicates the IR laser, which is wider than the excitation laser. **b-e.** Characteristic single-molecule time trajectories of LASR induced kissing complex dissociation, extended duplex formation, kissing complex formation and extended duplex formation from free hairpins, respectively. Reprinted from ref [73]

Following a T-jump to 67°C, we observed four distinct behaviors (Fig. 5.1b-e). Among 212 molecules in the kissing complex conformation immediately prior to the jump, 25 ± 3% refold into the extended duplex conformation (high FRET, Fig. 5.1b), while 62 ± 5% dissociate to the free hairpin conformation (0-FRET, Fig. 5.1c), and the remaining molecules stay in the kissing complex conformation. We then varied Tjump from 28 - 72°C to determine the temperature dependence of the kissing complex

dissociation and refolding reactions (Figs. 5.2a-b). Our data show that the fraction of molecules that dissociates or refolds increases with increasing Tjump, as expected. A hyperbolic fit to the resulting refolding curves (see Methods) yields dissociation and refolding temperatures T_d = 47 ± 2 ˚C and T_r = 54 ± 2 ˚C, respectively[28]. The refolding temperature is higher than the dissociation temperature, because the activation barrier to form the extended duplex is higher than that for dissociation. It is noteworthy that the refolding curve is more cooperative (steeper slope near the transition temperature) than the dissociation curve. This likely indicates that the number of hydrogen bonds broken and formed during extended duplex formation is larger than those broken during dissociation.
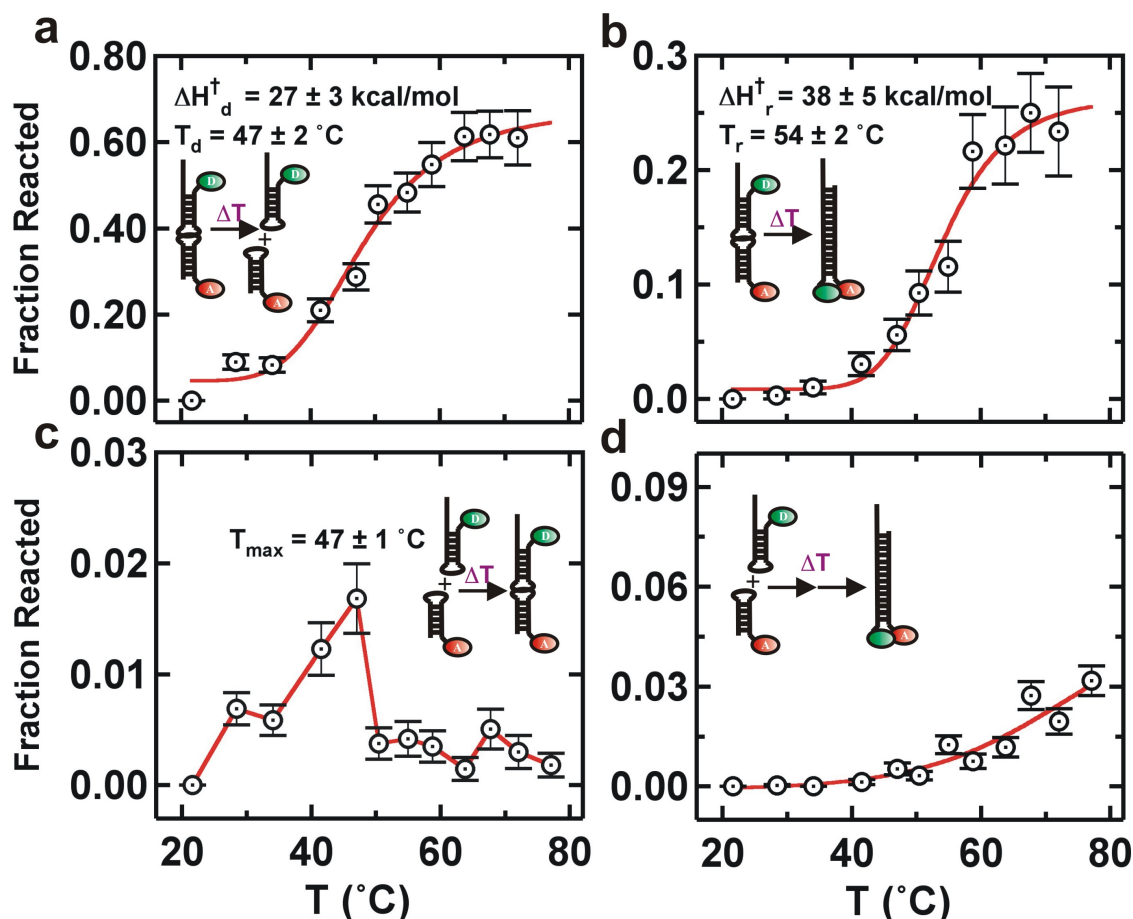
Figure 5.2 Single-molecule melting curves. a. Single-molecule melting curve and fitting for enthalpy of dissociation of kissing complex. **b.** Single molecule melting curve and fitting for enthalpy of dimerization of kissing complex. **c.** Single molecule melting curve of formation of kissing complex. **d.** Single-molecule melting curve of dimerization starting from free hairpin state. Reprinted from ref [73]

During the T-jump, a small fraction (1-2%) of the molecules initially in the free state (FRET = 0) form the kissing complex (0.5 FRET, Fig. 5.1d). This fraction increases between 28 and 47°C, and decreases to <1% at higher temperatures (Fig. 5.2c). Interestingly, the maximum of this refolding curve (T_{max}) coincides with the dissociation temperature (T_d), because above this temperature, kissing complex dissociation is predominant (Fig. 5.2b).

During T-jumps above 40˚C, we also observe a small fraction of molecules ($\leqslant$ 3%) that transition from the free state (FRET = 0) to the extended duplex (high FRET, Fig. 5.1e), and this fraction increases with Tjump (Fig. 5.2d). These molecules may reach the extended duplex conformation by two competing pathways that we cannot distinguish (Fig. 5.3c): (i) through the kissing complex intermediate or (ii) by melting and annealing HPI and HPII during the 1s T-jump. To test these two possible pathways, we designed an RNA hairpin (HPIII, Table 5.1) that cannot form the kissing complex with HPII (non-complementary loops) but is still capable of forming the extended duplex, albeit with a 3 nucleotide internal bulge. Before the LASR T-jump (Fig. 5.3a, t < 30s), single-molecule trajectories of surface-immobilized HPIII in the presence of HPII in standard conditions show no evidence of kissing complex formation, as expected. Following the LASR T-jump (Fig. 5.3a, t > 30s, Tjump = 70˚C), however, a small fraction of molecules transition to the extended duplex conformation (high FRET). Because HPII and HPIII cannot form the kissing complex, these molecules can only reach the extended duplex through pathway 2 (Fig. 5.3c). Therefore, the fraction of HPIII molecules that forms the extended duplex represents an estimate of HPI molecules that form the extended duplex through the melting pathway. The fraction of HPI molecules reaching the extended duplex is larger than HPIII at all temperatures (compare Figs. 5.3b and 5.2d), and their difference represents an estimate for HPI molecules that form the extended duplex through pathway 1. This enables determination of the fraction of HPI molecules reacting through pathway 1 at each temperature (Fig. 5.3d). Between 40 and 80˚C, this fraction is constant at ~80%, indicating that the kissing complex pathway is largely favored over the melting pathway over this temperature range. This result

shows that LASR can be used to estimate the yields of competing reaction pathways, a quantity that would be challenging to measure otherwise.
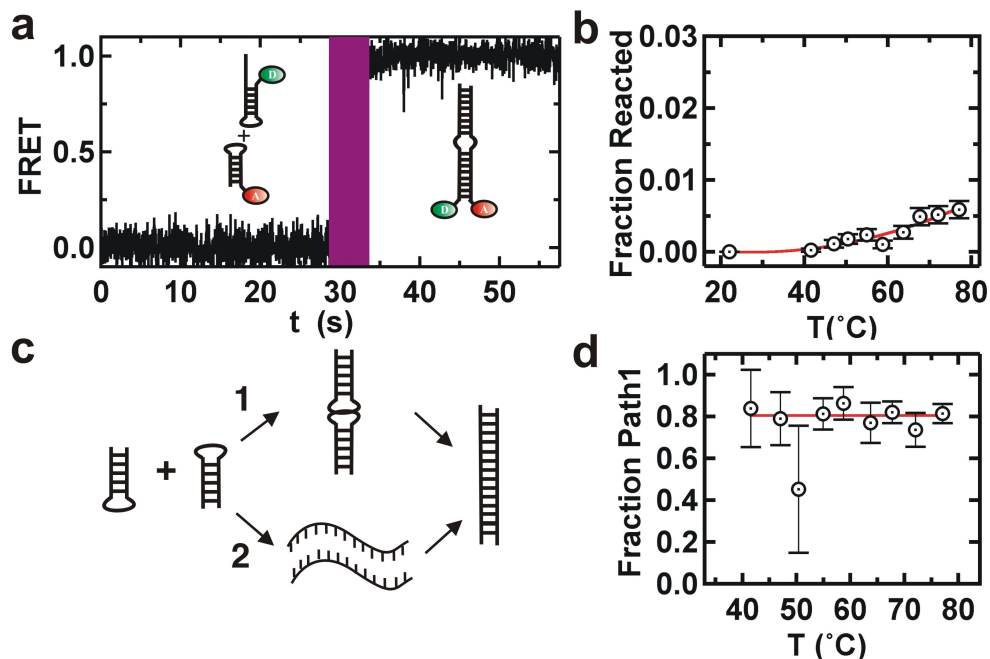
Figure 5.3 LASR applied to non-kissing mutant. a. FRET time course of non-kissing mutant. Purple bar indicates the T-jump. **b.** Single-molecule melting curve of dimer formation. **c.** Two possible competing reaction pathways. **d.** Ratio of two competing reaction pathways. Reprinted from ref [73]

A fourth RNA hairpin (HPIV, table 5.1) was designed to form a kissing complex with HPI (complementary loops) but not to form a stable extended duplex (non-complementary stems).   Before the LASR T-jump, single-molecule trajectories of surface immobilized HPI in the presence of HPII in standard conditions clearly show the formation of the kissing complex, as expected (Fig. 5.4a, t < 30 s). Following a T-jump to 75°C, we observed 25% of the molecules in the kissing complex immediately before the jump refold into a high FRET conformation. However, this conformation differs from the extended duplex (Figs. 5.1b-e) in two ways. First, its average FRET ratio is slightly lower than that of the extended duplex (0.95 in stead of 1.0, compare Fig. 5.2b and Fig.

5.4a). Second, its FRET ratio drops to zero after a few seconds, whereas the extended duplex remains stable in the high FRET for the duration of the experiment. Based on these observations and the sequences of HPI and HPIV, we assigned this new species to a transient RNA duplex that is stabilized by up to 9 base pairs from the loops of HPI and HPIV (Fig. 5.4 and Table 5.1). This experiment enables us to determine the stability of this transient RNA duplex by measuring the dwell time distribution in the high FRET state (Fig. 5.4b). An exponential fit to this distribution reveals a dissociation rate constant k_{diss} = 0.07 ± 0.01 s-1. These results show that LASR can be used to generate and characterize transiently stable RNA species that cannot be observed otherwise, which may be useful to study short RNA duplexes such as the 7 nt seed regions of micro RNAs (miRNAs) that are important for miRNA target recognition[90].
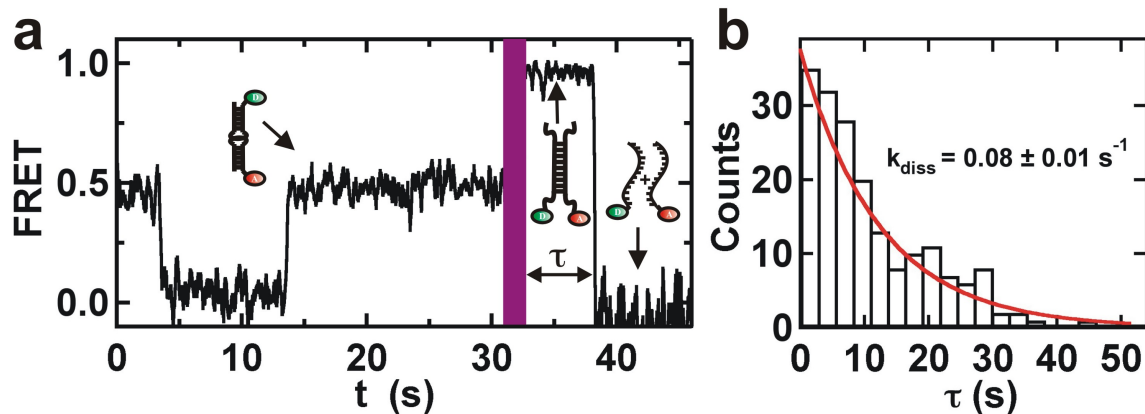
Figure 5.4 Formation of transient metastable strucutre. **a.** Single molecule FRET time course of a transient stable duplex formation and dissociation. τ is the dwell time in the transient duplex state. **b.** Histogram of HPIV mutant and Dwell time analysis to calculate k_{diss}. Reprinted from ref [73]

5.3.2 Transition state analysis using LASR

To obtain transition state parameters for the kissing complex dissociation and refolding reactions, we derived an expression based on Transition State Theory and the

Gibbs-Helmotz equation (see methods). This expression was used to fit the LASR dissociation and refolding curves (Figs. 5.2ab, red curves) to obtain the activation energy barrier ($\Delta H^{\ddagger}$). For the dissociation reaction, we obtain $\Delta H^{\ddagger}_d$ = 27 ± 3 kcal/mol, which is in excellent agreement with previous bulk studies. To further validate this analysis, we determined $\Delta H^{\ddagger}_d$, by single-molecule Eyring analysis. We measured k_{on} and k_{off} at temperatures ranging 15-23°C using a temperature-controlled microscope stage (Fig. 5.5). At higher temperatures k_{on} and k_{off} could not be determined because the number of kissing complexes observed decreased with temperature. A linear fit to an Eyring plot of k_{off} as a function of temperature yields a dissociation activation energy barrier $\Delta H^{\ddagger}_d$ = 29 ± 7 kcal/mol, which is in excellent agreement with the value obtained with LASR.

For the refolding reaction, a fit to the LASR data results in $\Delta H^{\ddagger}_r$ = 38 ± 5 kcal/mol (Fig. 5.2b), higher than $\Delta H^{\ddagger}_d$, as expected. However, because the kissing complex is a deep kinetic trap most molecules dissociate before forming the extended duplex, $\Delta H^{\ddagger}_r$ cannot be readily determined by other methods, including single-molecule Eyring analysis. Therefore, LASR presents a unique opportunity to determine transition state parameters that could not be otherwise measured.
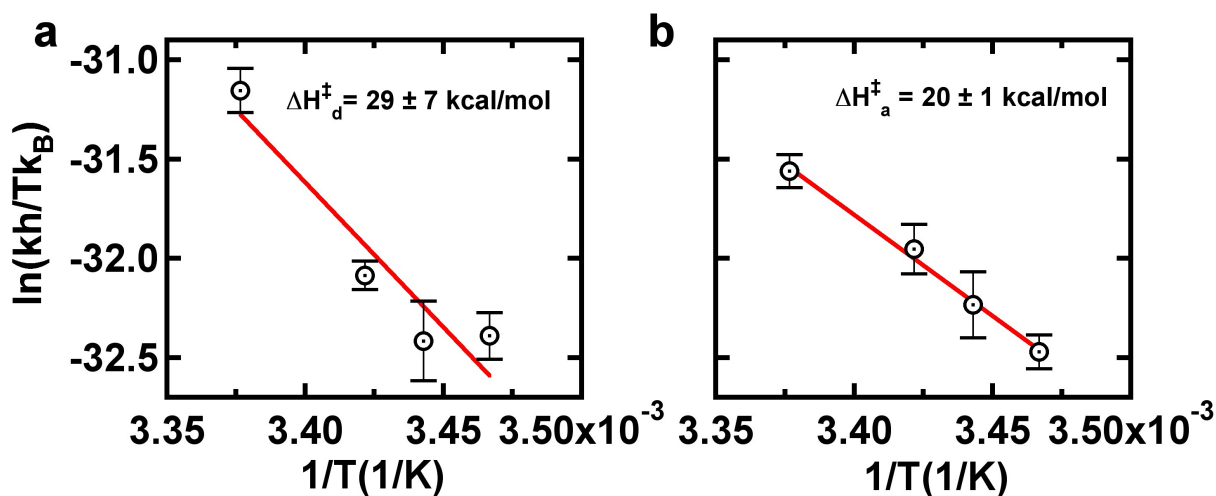
Figure 5.5 Enthalpy of barrier from Eyring analysis. a. Eyring's analysis gives enthalpy of the HP1+ HP2 kissing complex formation and dissociation. **b.** Eyring's analysis gives enthalpy of the HP1+ HP2 kissing complex dissociation Reprinted from ref [73]

5.4 Conclusions

The ability to characterize folding kinetic traps is critical to understanding the refolding mechanisms of biological macromolecules. However, this can be challenging because of the high-energy barriers involved. Here, we have demonstrated how a new single-molecule technique, Laser Assisted Single-molecule Refolding (LASR), can be used for this purpose. By combining temperature jump kinetics with single molecule spectroscopy, we are able to drive refolding reactions otherwise hindered by large kinetic barriers. We first showed how LASR could be used to study DNA melting at the single-molecule level for three dsDNAs with melting temperatures ranging 48–70 °C. A comparison with an ensemble averaged DNA melting assay confirms the accuracy of the LASR measurement. We then used LASR to study the refolding reaction of a model kinetic trap, an RNA kissing complex, into an extended duplex. From the resulting LASR temperature curves we obtain quantitative energy barriers for the kissing complex

reactions that we validate against two other measurements at the single molecule level and in bulk. The primary advantage of LASR is that it enables us to study reaction that would otherwise be very challenging with other methods. We also demonstrate how LASR combined with mutagenesis can be used to estimate the relative yields of competing reaction pathways, as well as to generate and characterize transiently stable RNAs that cannot be otherwise formed.

(Chapter 5 is adopted from paper *Zhao, R. et al. Laser-assisted single-molecule refolding (LASR). Biophys J* **99**, *1925-1931, doi:S0006-3495(10)00889-1*[73])

CHAPTER 6

Origin of memory effect in Hairpin Ribozyme docking kinetics

6.1 Introduction

RNA molecules are biopolymers with various important biological functions such as protein synthesis, gene regulation and splicing[4,5,7]. The catalytic function of RNA requires the precise formation of the secondary and tertiary structure, which is essential to the assembly of the catalytic core. RNA folds in a hierarchic way. Secondary structure folds first with the assistance of monovalent cation. With the help of divalent cation, RNAs can search through the potential energy surface to find the global minimum that corresponds to the native structure[11]. Due to the flexibility of RNA, the folding potential energy surface can be rugged and full of kinetic traps[13], which will keep the RNA in intermediate structures that have distinct dynamics from the native structure.

RNA structural dynamics have profound influence on the catalysis. A RNA, as simple and small as the hairpin ribozyme, has subpopulations with distinct structural fluctuations that are linked to distinct catalytic activity[14]. We hypothesize that this heterogeneity in structural dynamics originates from different intermediate structures that are kinetically entrapped. To test this hypothesis, the minimal active form of the hairpin ribozyme is adopted in this study as a model system[22]. The hairpin ribozyme RNA was found in tobacco ringspot virus satellite RNA rolling cycle replication process. It catalyzes a self-splicing reaction without the help of protein enzyme in a compact folded structure[27]. It has two independently folding helix-loop-helix domains, A and B[22]. The RNA is in equilibrium between two conformations, where two domains are coaxially

stacked (undocked) or closely interacting in loops A and B (docked) (Fig. 6.2b). Detailed reaction pathway has been described previously[23]. The catalytic core is only formed in the docked structure. Probably the most interesting characteristic of the hairpin ribozyme structural dynamics is the memory effect, in which molecules transit from docked structure to undocked structure with a distinct (three orders' of magnitude difference) kinetic rate constant and molecules tends to undock with the same rate constant for a long time that is greater than 3 hours[14,26]. Yet, hairpin ribozymes with distinct undocking rates are all catalytically active. Walter's lab has shown that this memory effect may last through native PAGE gel separation, denaturation and recombination[31]. Footprinting assay and mass spectrometry data shows almost no differences in the primary and secondary structure between subpopulations of the hairpin ribozyme[31]. Nevertheless, the possibility that the memory effect is an artifact of covalent difference from RNA synthesis is not completely ruled out. In this study, we combined a single-molecule FRET study and laser induced temperature jump to clarify that the memory effect is an intrinsic property of hairpin ribozyme folding. This study also allowed measurement of the energetic barrier height of inter-conversions.

A two-way junction construct of the hairpin ribozyme with a mutation to prevent cleavage was created to focus solely on the structural dynamics. Cy3 and Cy5 were added to the 3' and 5' end of strand A, respectively, to allow monitoring of distance by FRET. 5' end biotin was added to strand B for surface immobilization. Single molecule fluorescent emission was measured using prism based total internal reflection fluorescent microscopy as described in chapter two. IR laser illumination of the

observation volume was used to heat up the solution around the RNA sample as desired as described in chapter two.

6.2 Methods

Table 6.1 Hairpin ribozyme used in this study:

Name	Description	Sequence[1,2]
RZA	hairpin strand 1	5' Cy5-AAA UAG AGA AGC GAA CCA GAG AAA CAC ACG CCA AA U-3' Cy3
RZB	hairpin strand 2	5' B-AUA UAU UUG GCG UGG UAC AUC ACC UGG UAC CCC CUC GCA GUC CUA UUU-3'
RZA'	One piece Construct strand1	X AAA UAG AGA AGC GAA CCA CAG AAA CAC ACG CCA A Q AUU GGC GU
RZB'	One piece Construct strand 2	PGGU ACA UUA CCU GGU ACC CCC UCG C Q G UCC UAU UUA UAU AX
Splint	DNA splint	5'- CAG GTA ATG TAC CAC GCC AAT ATT GGC- 3'

[1] Q=dT with C6 amino liner for Cy3-NHS labeling, P=5' phosphate for ligation. Am= 2' methyl A, and B=Biotin

Ligation[91] was carried out to ligate RZA' and RZB' to give a one-piece construct, so the dye labeled strand does not dissociate at high temperatures. T4 RNA ligase 2 (M0239L, New England BioLabs Inc.) was used as per the manufacturer's protocol. After ligation, the RNA was purified using PAGE and then eluted out, filtered, ethanol precipitated, dried and dissolved.

6.2.1 Laser-assisted single molecule refolding (LASR)
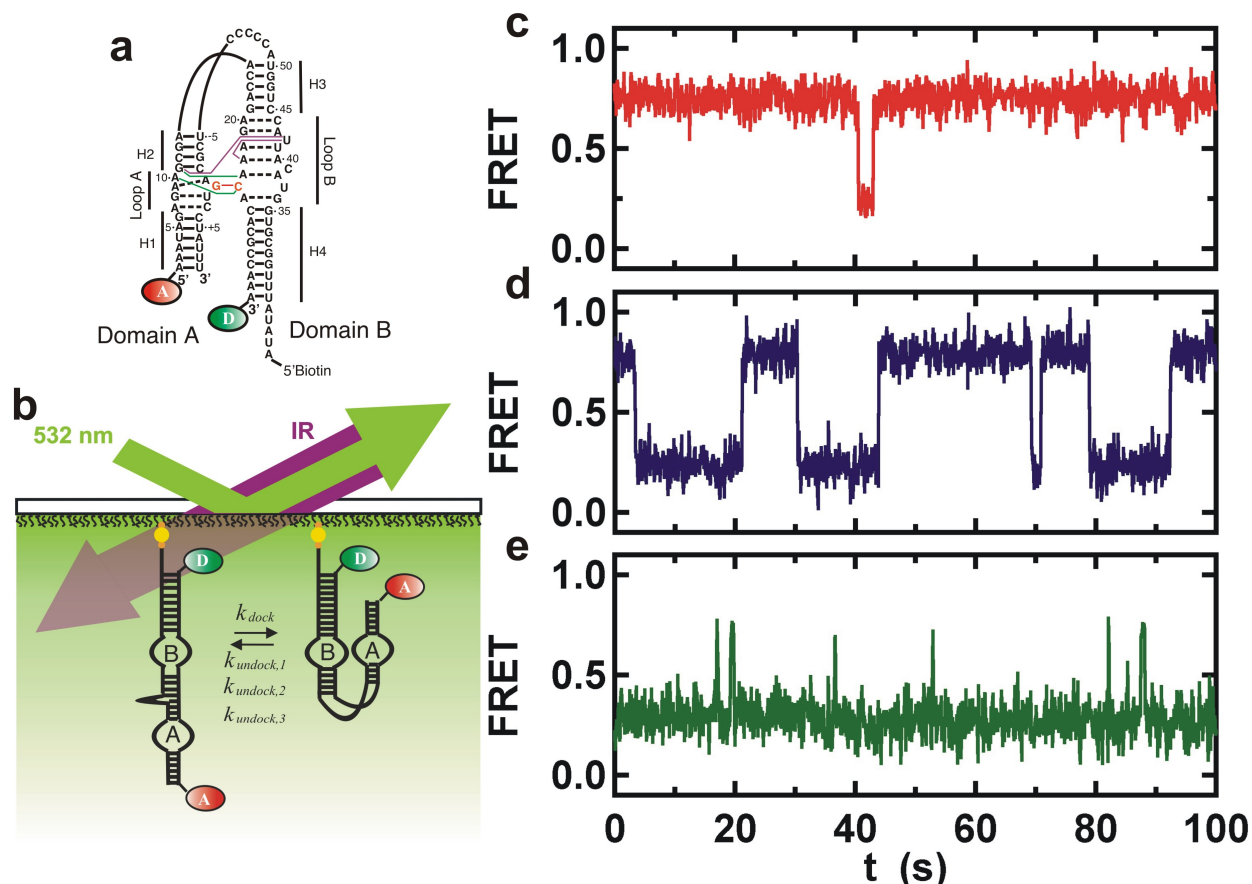
Figure 6.1 Laser-assisted single molecule refolding (LASR) of hairpin ribozyme. a, Secondary and tertiary structure of hairpin ribozyme. Black lines denote Watson-Crick base pairs. Dashed lines indicate non-canonical base pairs. Purple lines indicate U42 binding pocket. Green lines indicate ribose zipper interaction. Orange line indicates tertiary GC base pair. **b.** Experimental design of LASR of hairpin ribozyme. Two strand construct hairpin ribozyme was surface immobilized to a quartz slides via biotin-streptavidin-BSA biotin bridge (orange and yellow circles). FRET donor (green oval) and acceptor (red oval) were attached to the end of one strand via a C6 linker. Green arrow: 532 nm diode pumped green laser. Purple arrow: IR laser. **c,d,e,** Three distinct single molecule FRET time courses shows subpopulations of hairpin ribozyme of distinct undocking rates. The high FRET~0.8 corresponds to the docked structure. The low FRET~0.2 corresponds to the undocked structure. Red: Subpopulation 1. Blue: subpopulation 2. Green: subpopulation 3.

LASR is a new technique developed in the Rueda Lab to study kinetically trapped biopolymers[73]. It enables the characterization of the activation barriers of deep kinetic traps by combining traditional single-molecule FRET detection with infrared (IR)

laser-induced temperature jumps (T-jumps) [69,80,81]. An IR laser is used to excite the second overtone of the water OH stretch (~1400 nm)[71] and thus rapidly increase the solvent temperature. As a result, kinetically trapped biopolymers in solution receive sufficient thermal energy to overcome the refolding activation barrier (Fig. 6.2b). The T-jumps are transient (but reversible) temperature increases, and thus differ from the traditional T-jump[69]. The accuracy of LASR has been demonstrated in chapter 3.

In LASR, the IR laser introduced to the excitation spot through the prism at more acute angle than the critical angle (Fig. 6.2b). The IR laser irradiation was gated using a computer driven shutter. In this study, heating duration was fixed as 1 s. The IR laser power was adjusted using combinations of neutral density filters to give varied temperature jump magnitude. Multiple experiments were carried out at varied temperatures to obtain a LASR melting curve. The final effective IR laser power delivered to the water solution is estimated to be ~100 mW, taking into account the energy loss from reflection and the water absorption efficiency.

A micro-sized gold wire microsensor was adopted to calibrate the magnitude of the temperature jump as previously described in chapter 2.

6.2.2 Native gel separation

Sub-populations were separated on 12% polyacrylamide native gels running at 4 ˚C and 120 volts in 50 mM TRIS pH 7.5, 100 mM NaAc, 20mM Mg(Ac)2, 20 mM Ca(Ac)2. The running buffer solution was changed every two hours to maintain pH. The finished gel was scanned on a 'Typhoon' phosphorimager (GE healthcare) to locate the separated subpopulations. The separated hairpin ribozyme subpopulations were

excised from the gel and eluted. Gel fragments were removed by a poly-prep column and store at 4 °C for future T-jump study.

Sub-population enriched hairpin ribozyme was diluted to the proper concentration and immobilized on the quartz slide surface via biotin streptavidin interaction as described previously[7]. Then the LASR experiment was carried out on these hairpin ribozyme as described previously[2].
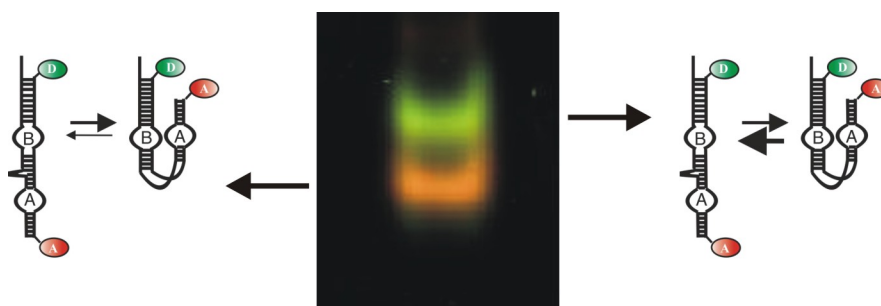
Figure 6.2 Native FRET gel shows at least two subpopulations; top band favors more undocked structure and the bottom band favors more the docked structure**.**

6.2.3 Data analysis

Sub-populations of the hairpin ribozyme with different undocking rate constants were categorized as slow (>50 s), medium (~14 s) and fast (< 4s) according to the average time spent in the docked structure. Molecules that change sub-population before and after the T-jump were denoted as 'inter-converting'. (About 25% of molecules showed FRET at 0.2 and never dock; these molecules were not included in statistical analysis.) The fraction (Φ) of inter-converting molecules was calculated as: $N_{inter-conversion}/N_{total}$, where N_{total} is the number of total molecules.

Single molecule Erying analysis
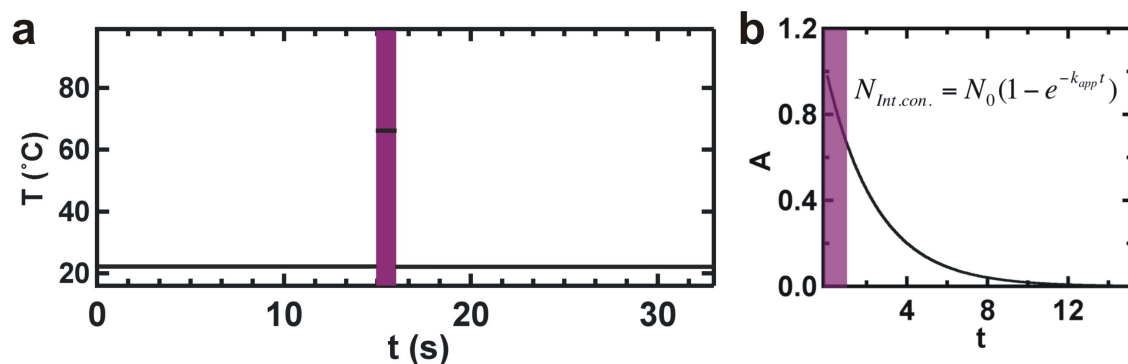
Figure 6.3 Calculation of apparent rate constant. a. Assumption of step function behavior of T-jump profile. **b.** Calculation of apparent rate constant of reaction.

To do Eyring analysis, we must determine the kinetic rate constant of the reaction. The apparent kinetic rate constant k_{app} was obtained from equation 13

$$k_{app} = -\ln(1-\phi)/t \qquad \text{(Eq. 6.1)}$$

where ϕ is the fraction of the reacted molecule , t is the heating duration. This equation is true if we assume the temperature jump profile is a step function, which is approximately true given that t is long (>1s).

Then, the rate constants can be used for Eyring analysis:

$$\ln \frac{k_{app}h}{k_B T} = -\frac{\Delta H^{\pm}}{R}\frac{1}{T} + \frac{\Delta S^{\pm}}{R} \qquad \text{(Eq. 6.2)}$$

where h is Plank's constant, k_B is the Boltzman constant, R is the gas constant, T is the final temperature in Kelvin, $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are the enthalpy and entropy changes between the ground state and transition state. The slope and interception from linear fit of the experimental data can be used to calculate $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$, respectively.

The error bar in Eyring's analysis is calculated as:

$$err = \frac{-\sqrt{n_{Int.con}}}{\left(n_{total} - n_{Int.con.}\right)\ln\left(1 - \dfrac{n_{Int.con.}}{n_{total}}\right)} \qquad \text{(Eq. 6.2)}$$

6.3 Results and discussion

First, we measure the heterogeneity in the hairpin ribozyme folding. Efficiency of FRET is calculated experimentally as FRET= $I_A/(I_A+I_D)$, where I_A and I_D is the emission intensity of the Cy3 and Cy5 from a single hairpin ribozyme. In docked structure, Cy3 and Cy5 are in proximity, thus is characterized by a high FRET value at ~0.8. Whereas in undocked structure, Cy3 and Cy5 are far apart, thus the FRET is around ~0.25. The rate constant of undocking is calculated from multiple exponential fitting to the histogram of time spent on docked structure. Four subpopulations of hairpin ribozyme with distinct undocking rate constants were observed, which is in good agreement with the previous study[32]. Salt concentration was increased to accelerate the folding dynamic, so that more frequent undocking events is observed to improve the accuracy of analysis. For simplicity of analysis, molecules are categorized as SubPop 1 (slow, $k_{undocking}4$, 58%) SubPop 2 (medium, $k_{undocking}3$, 15%) and SubPop 3 (fast, $k_{undocking}1$ and 2, 27%) according to the undocking rate constant. Fractions of each sub-population were calculated (0.58, 0.08 and 0.35) by counting molecules in different categories.
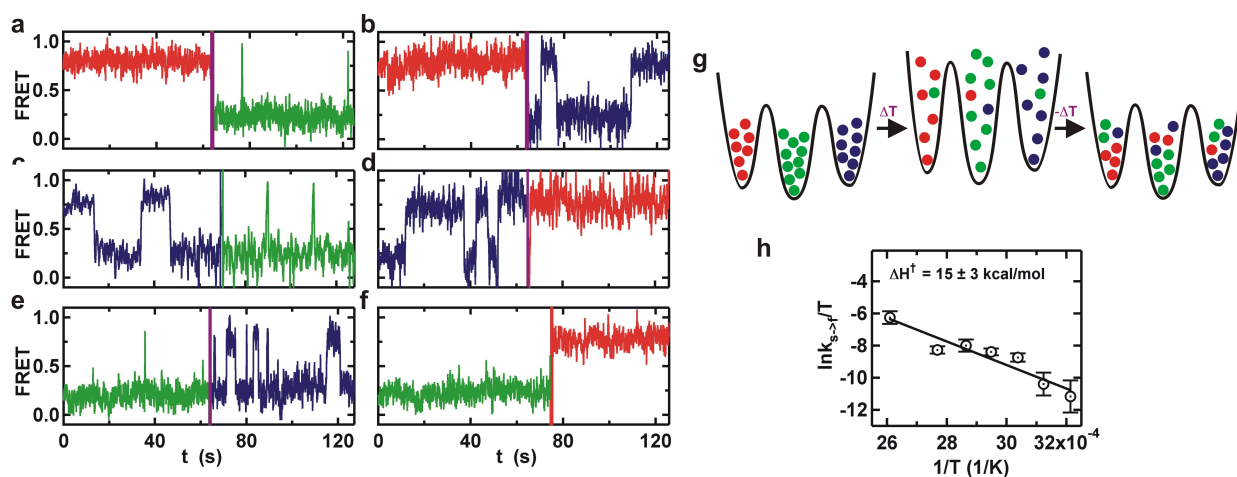
Figure 6.4 Temperature introduced inter-conversion between docking dynamics subpopulations. Red: SubPop1, blue: SubPop2. Green: SubPop3. **a-f,** Representative FRET time courses showing inter-conversions between subpopulations. Purple bar indicates the duration of heating. **g.** Schematic cartoon shows that an elevation in temperature increases the barrier on the potential energy surface for inter-conversion between subpopulations (labeled with red, green and purple). Molecules redistribute to other states after temperature relaxation. **h.** Single molecule Eyring's analysis gives an enthalpy change of $DH^{\ddagger}$= 15 ± 3 kcal/mol between the slow undocking state and the transition state towards fast undocking state.

We have found direct evidence showing that the heterogeneity of hairpin ribozyme folding is intrinsic. The application of LASR for 1s to a final temperature of 76 ˚C induced many more inter-conversions between populations compared to room temperature. The fraction of molecules that inter-convert increases slowly with temperature and remains small even at high temperatures (<15% at 78˚C). The FRET was monitored for the same hairpin ribozyme before and after the T-jump. As shown in Figure 6.5a, this molecule appears as SubPop 1 initially, and then it converts to be SubPop 3 after the T-jump. This is the most commonly observed inter-conversion. All subpopulations were able to interconvert between each other as shown in Fig 6.5b-f If the heterogeneity of the folding dynamics were due to covalent differences from synthesis, then at least some of the populations would be unable to interconvert. The

results support the hypothesis that folding heterogeneity originates from the rugged folding free energy surface, the barrier of which can be overcome at high temperatures. As indicated in table 6.2, the fraction of inter-conversion from SubPop 2 to other subpopulations is much higher than would be statistically expected based on the abundance of each populations at the beginning of the experiment, which indicates that subpopulation 2 has the lowest barrier height to overcome for inter-conversion. Similarly, SubPop3 has the highest barrier for inter-conversion.

Initial	Final	N 56°C	N% 56°C	Con.%	Initial%
Pop. 1	Pop. 2	6	17 ± 7	63 ± 13	58 ± 8
Pop. 1	Pop. 3	16	46 ± 12		
Pop. 2	Pop. 1	6	17 ± 7	34 ± 10	15 ± 3
Pop. 2	Pop. 3	6	17 ± 7		
Pop. 3	Pop. 1	0	0 ± N.A.	3 ± 3	27 ± 5
Pop. 3	Pop. 2	1	3 ± 3		

Table 6.2 Fraction of inter-conversion at 56 °C were calculated and categorized according to the initial population and compared to fraction of initial subpopulations. First column: Initial state of the hairpin ribozyme before T-jump. Second column: Final state of the hairpin ribozyme after T-jump. Third column: number of inter-conversions specified by first and second column at 56°C. Fourth column: percentage of the corresponding inter-conversions. Fifth column: the percentage of inter-conversions starting at the same initial state as specified by first column. Sixth column: percentage of hairpin ribozyme in the state specified by first column before T-jump.
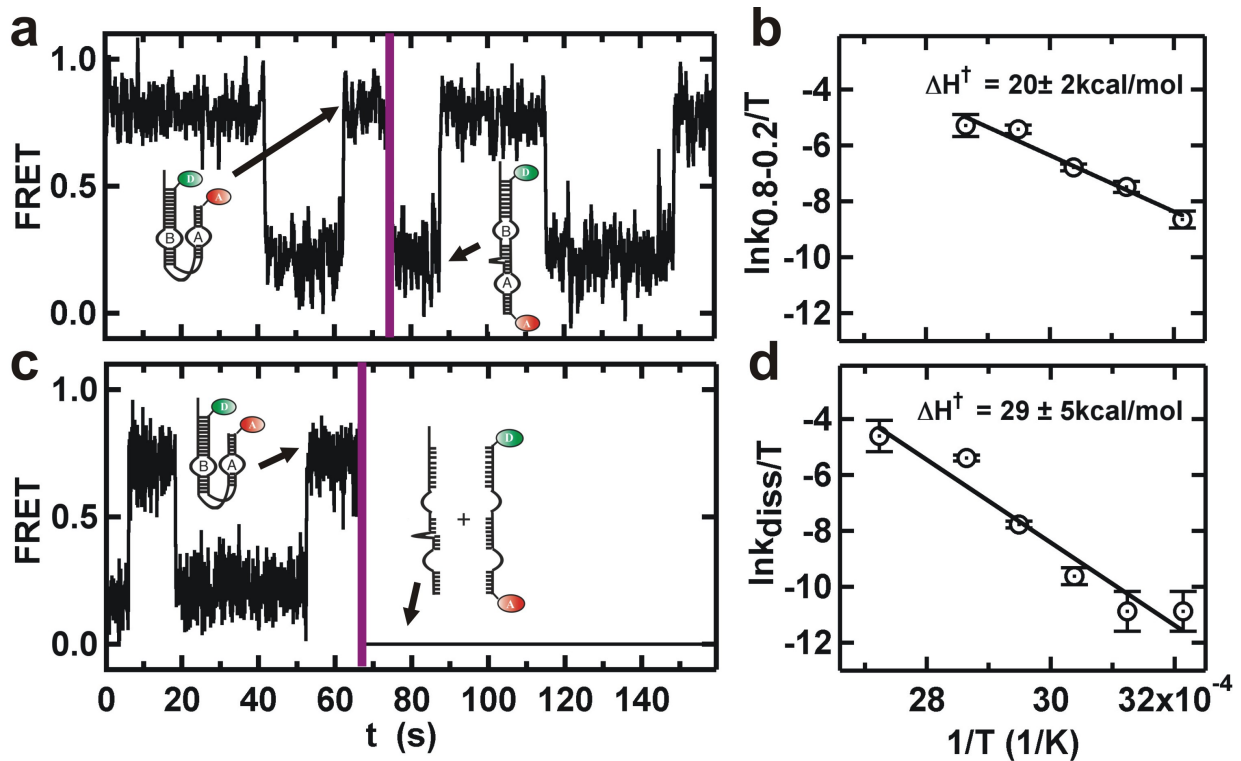
Figure 6.5 Single-molecule Eyring's analysis of structure disruptions. a. A typical FRET time course shows a transition from FRET=0.8 to 0.2. **b.** single-molecule Eyring analysis gives an enthalpy change $\Delta H^{\ddagger}$=20 ± 2 kcal/mol for overcoming the energy barrier to disrupt the tertiary interactions formed in the docked structure. No events were observed at temperature 110 °C. **c.** A typical FRET time course shows a transition from FRET=0.8 or 0.2 to the loss of both donor and acceptor signal. **d.** single-molecule Eyring analysis gives an enthalpy change $\Delta H^{\ddagger}$=29 ± 5 kcal/mol for overcoming the energy barrier to dissociate two strands of hairpin ribozyme.

To further determine the barrier heights of inter-conversions, we have adopted single molecule Eyring analysis. This allows estimation of the thermodynamic parameters of the barrier height between subpopulations. These parameters are a means to interrogate the nature of the interactions between the two domains in the hairpin ribozyme. From the temperature dependence of the apparent rate constants, we can fit for the enthalpic and entropic change of the barrier height for all the molecular reactions as shown in table 6.3. We assumed a step function for the T-jump profile and one-way reaction during T-jump to derive the apparent rate constant. These make the

barrier height we obtained in this study not readily comparable to other studies. Nevertheless, the numbers obtained are valuable for comparing barrier heights rather than as absolute values as all the assumptions and conditions are identical. We obtained an enthalpic change of the barrier height of 15 ± 3 kcal/mol for the slow to fast undocking inter-conversion (SubPop.1->SubPop.3), which is smaller than the enthalpy of 20 ± 2 kcal/mol calculated for the disruption of tertiary interaction in the docked state (0.8 to 0.2 FRET state). This indicated that there are fewer interactions broken during subpopulation inter-conversion than complete tertiary interaction disruption. Given that SubPop 1 resides mostly in the 0.8 FRET state and SubPop 3 resides mostly in the 0.2 FRET state. It is likely that most of the enthalpic term for conversion from SubPop 1 to 3 is due to abrogation of tertiary interactions. The enthalpy directly contributing to inter-conversion is very smaller and hence, the inter-conversion is apparently an entropy-controlled process. The reason for the fewer observed inter-conversion than the undocking (0.8->0.2) lays in the fact of very negative entropy (-43 cal/mol/k, table 5.1). Negative entropy means that the transition state has a more rigid structure than the ground state and the reaction will take place slowly. Entropy contributes to 30% of the barrier height of inter-conversion from SubPop 1 to SubPop 3 at room temperature. This entropic contribution is significantly larger compared to the undocking and dissociation of hairpin ribozyme.

Table 6.3 Barrier information from Eyring analysis

	Inter-conversion (1->3)	Undocking (0.8->0.2)	Dissociation
$\Delta H^{\ddagger}$ (kcal/mol)	15 ± 3	20 ± 2	29 ± 5
$\Delta S^{\ddagger}$ (cal/mol/K)	-22 ± 6	10 ± 4	24 ± 6
$\Delta G^{\ddagger}_{295}$ (kcal/mol)	22 ± 4	17 ± 3	22 ± 6

Another effect from very negative entropy is the increase of free energy barrier height at higher temperature as indicated by Van't Hoff's equation. Nevertheless, the Boltzmann distribution of the RNAs' energy becomes broader at higher temperature, which helps more molecules to overcome the barrier even though it is higher (Fig 6.5g). This cancellation between two effects explains why the percentage of the observed inter-conversion is low (<11%) even at a temperature of 88°C. From the difference of free energy barrier height between undocking and inter-conversion, we can estimate the lifetime of the inter-conversion using equation $k^{\pm} = k_0 e^{-\Delta G^{\pm}/RT}$ assuming the difference in the pre-factor k_0 for both reactions is negligible and the average lifetime of the docked structure is 50 s. We obtained a lifetime of 70 hours for the inter-conversion from SubPop 1 to SubPop 3, which is in good agreement with previous observation[14] (<5% after 3 hrs).

To calculate the barrier height for each inter-conversion, we separated subpopulations of the hairpin ribozyme by EMSA as described in the methods. Slow

undocking subpopulations were enriched in the bottom band and fast undocking subpopulations were enriched in the top band (Fig 1.6). The fraction of each subpopulation is summarized in table 6.4

Table 6.4 Factions of subpopulations in two bands separated by EMSA before and after T-jump of 56 ˚C.

Fractions		SubPop1	SubPop2	SubPop3
Unseparated	Before	0.58 ± 0.08	0.15 ± 0.03	0.27 ± 0.05
	After	0.4 ± 0.05	0.12 ± 0.03	0.48 ± 0.05
Top	Before	0.35 ± 0.04	0.11 ± 0.02	0.54 ± 0.05
	After	0.26 ± 0.04	0.16 ± 0.03	0.58 ± 0.05
Bottom	Before	0.71 ± 0.06	0.20 ± 0.03	0.09 ± 0.02
	After	0.45 ± 0.06	0.33 ± 0.05	0.22 ± 0.04

There are only two major bands shown in the native gel instead of three or four bands as indicated from single molecule study. We think there are actual three subpopulations with SubPop 1 and 3 showing on bottom and top, respectively. SubPop 2 is not visible on the gel because it is minor population and has broader distribution. SubPop 2 spends almost the same amount of time on docked and undocked structure, which travel very differently on the gel and hence give a broad band. So top band is a mixture of SubPop 2 and 3. Bottom band is a mixture of SubPop 1 and 2. The small percentage of SubPop 1 in top band and the SubPop 3 in bottom band were most likely due to slow inter-conversion of subpopulations with time.

LASR experiment for hairpin ribozyme eluted from both bands gives similar results as the unseparated one. The mainly observed inter-conversion at 80˚C is SubPop1 to

SubPop3. SubPop3 barely interconverts into other subpopulations (Table 6.4). The

barrier information is summarized in table 6.5

Table 6.5 Barrier information of the inter-conversions between subpopulations.

Inter-conversion		1-2	1-3	2-1	2-3	3-1	3-2
Top	ΔH(kcal/mol)	11 ± 1	13 ± 3	NA	9 ± 1	NA	-1 ± 2
	ΔS(cal/mol/k)	-34 ± 2	-24 ± 10	NA	-41 ± 4	NA	-70 ± 5
	ΔG(kcal/mol)	21 ± 1	20 ± 4	NA	21 ± 2	NA	20 ± 3
Bottom	ΔH(kcal/mol)	12 ± 4	14 ± 3	1 ± 3	15 ± 8	NA	1 ± 5
	ΔS(cal/mol/k)	-27 ± 12	-24 ± 10	-69 ± 10	-25 ± 20	NA	-71 ± 15
	ΔG(kcal/mol)	20 ± 5	21 ± 4	22 ± 4	22 ± 10	NA	22 ± 7
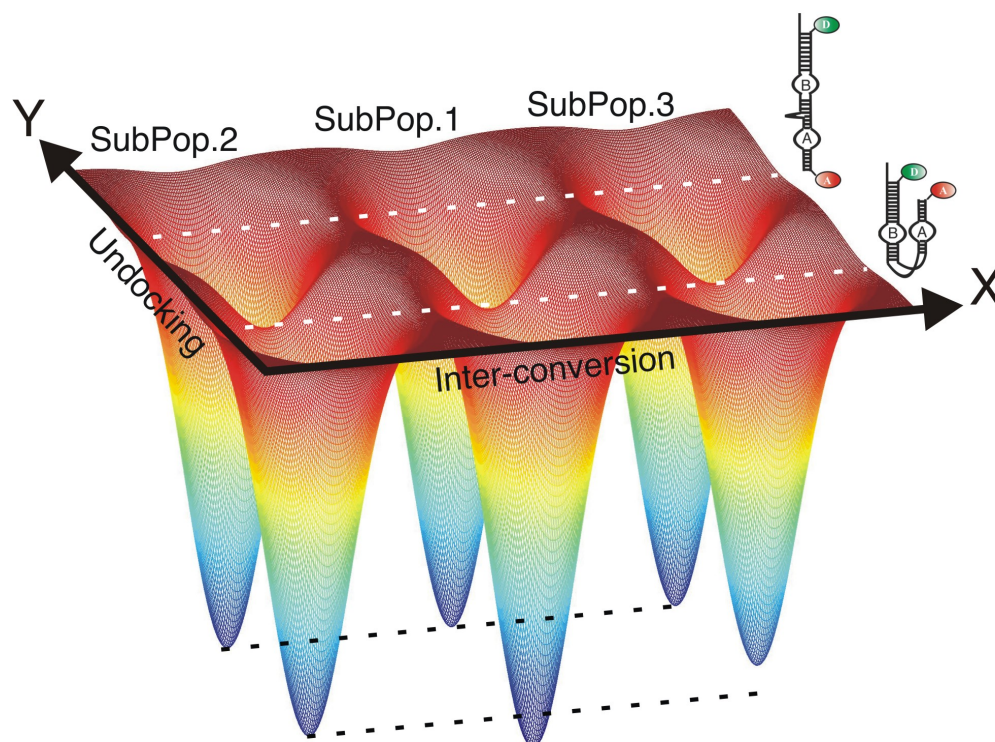
Figure 6.6 Free energy diagram of hairpin ribozyme undocking and inter-conversion. Drawing is not to scale. Front and back wells correspond to docked and undocked hairpin ribozyme, respectively, as indicated by white dashed line. Y-axis denotes the undocking (or docking in the reverse direction) coordinates. X-axis denotes the inter-conversion coordinates. Sub-populations with distinct undocking rate constant were indicated with different depth of the wells and denoted as SubPop1-3.

To better visualize the hairpin ribozyme docking/undocking dynamics and the inter-conversions between subpopulations, we have proposed a model as shown in figure 6.7 The well depths of the undocked structure on the free energy surface are the same, whereas the well depths for the docked structure are very different, which give rise to the heterogeneity in undocking. RNAs in different subpopulations are in equilibrium between the docked and undocked structure at room temperature as the barriers for docking and undocking are easily overcome by thermal fluctuation. The RNA docks and undocks with a rate constant associated with the barrier height. However, there are higher barriers between the subpopulations preventing the RNA from inter-converting between subpopulations. This high barrier height is the origin of the memory effect in the hairpin ribozyme folding.
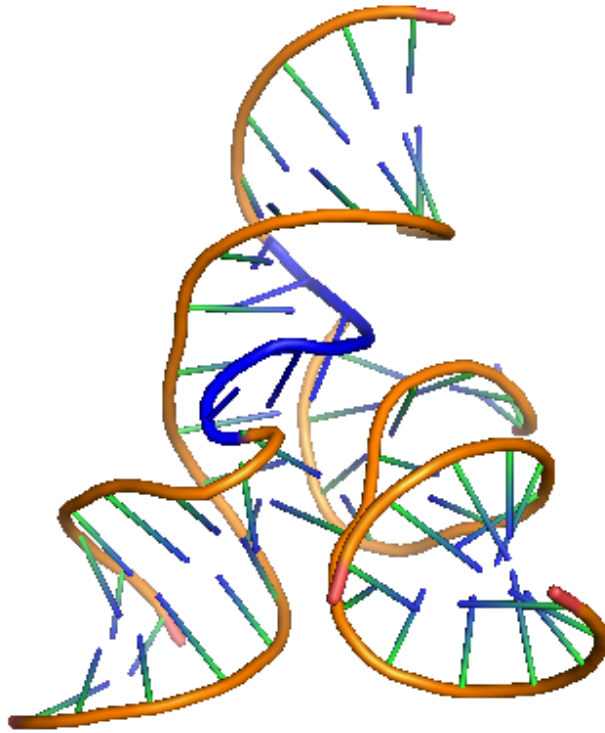
Figure 6.7 S-turn in loop B with sequence: CAXUAC. Similar long lasting memory effect of folding heterogeneity was observed in the study of Sarcin-Ricin loop that shares structure homology and S-turn motif. The s-turn motif is highlighted in blue in the crystal structure[22].0

The Structural origin of the memory effect is probably the flexible loop regions. NMR and crystal structures show differences in the loop[3,22,92]. Previous study of the Sarcin-Ricin loop, which includes sequence CAXUAC, (S-turn[93]) like that found in loop of the hairpin ribozyme, shows similar structural heterogeneity. In the context of this study, S-turn structure heterogeneity is a plausible explanation for hairpin ribozyme folding heterogeneity. As we can see from the crystal structure, the S-turn is completely buried inside the interface of tertiary contacting. Orientation of the base flipping in the S-turn or the conformation of the S-turn itself could introduce steric effect in the docked

structure, hence resulting heterogeneity of the stability of docked structure, which shows up as memory effect in folding. If this hypothesis is true, the inter-conversion between subpopulations is likely involving structural rearrangement of S-turn, which should involve little enthalpic change and large entropic change. And our result provides support this hypothesis. Further mutational studies may allow identification of the interactions that contribute to the barrier height of inter-conversions.

6.4 Conclusions

We have demonstrated that the heterogeneity in the hairpin ribozyme undocking is an intrinsic property of the RNA by forcing inter-conversion between subpopulations using LASR. Also, we calculated enthalpy of the barrier for inter-conversion using Eyring analysis. The small enthalpic contribution of the barrier suggests that only a subset of tertiary interactions need to be broken for inter-conversion. The entropy change contributes 30% of the ruggedness on the hairpin ribozyme folding free energy surface and results in an extremely high temperature for efficient inter-conversion. The negative entropy of the transition state indicates a rigid transition state for the inter-conversions and makes the inter-conversion reaction slow. This work highlights the power of LASR in the study of the transition states in biopolymer folding and demonstrates that memory effect are naturally occurring rather than artifactual.

CHAPTER 7

Conclusions and future directions

To perform their biological functions, biopolymers such as RNA, protein and DNA must fold into the correct three-dimensional structure. This can be challenging given the extremely large structural sampling space determined by combination of all possible dihedral angles of the chemical bonds and base pairing. However, it is clear that biopolymers fold in a hierarchical way[10,94,95], in which local interactions form first followed by global interactions. This significantly simplifies the folding problem of biopolymers, but still gives a complex folding free energy surface with many kinetic traps[13] and global minima corresponding to inactive intermediate states and active native states, respectively. Due to the rugged folding free energy, biopolymers tend to fold into different states with distinct catalytic activities or structural fluctuations[14,15,26,51,96,97]. The inter-conversion between them can be slow and hence result in a memory effect in biopolymer folding. Previous studies on protein folding have shown the presence of multiple native conformations with varied activities[96,97] that can interconvert by refolding from denatured structures[98]. This indicates that folding heterogeneity is an intrinsic property of biopolymers rather than an artifact such as covalent difference from synthesis.

Similar folding heterogeneity has been observed for non-coding RNAs[14,15,26] (ncRNAs). Compared to protein, ncRNAs have even more complex and rugged folding free energy surfaces due to larger structural sampling spaces and the stableness of the local structures formed at the early stage of RNA folding[94,99]. Once an alternative

structure is formed, it normally requires an enzyme, such as chaperone[17], helicase[19] or cofactor[20], to interconvert it to other conformations. Kinetically trapped RNA in the intermediate state with non-native interaction may cause the RNA to fold in a distinct fashion that is different from the native structure[14,26]. When the kinetic trap is deep enough, RNAs will stay in that intermediate state for a long time without converting to native structure[31] and hence showing a distinct structural dynamic for a long time, which was referred to as the memory effect. Previous studies of RNA folding have also shown this folding heterogeneity and memory effect are common and intrinsic to RNA. This memory effects were observed in Ribosome[100], hairpin riboyzme[14], tetrahymena group 1 intron (TG1I) ribozyme[15], sarcin-rincin loop[93] and other engineered RNA aptamers[101]. It was also speculated that different folding structure and dynamics may respond to cellular and environmental cues differently and hence make the host cell of these RNAs more adaptive to the evolutionary pressure[15]. Previous studies on hairpin ribozymes[14] and TG1I ribozyme[15] have shown both ribozymes have to dock the substrate domain into the catalytic core to perform catalysis. Memory effect of the heterogeneous docking and undocking rate constants were observed. Molecules with distinct folding dynamics are all biologically active. This is contradictory to traditional belief that there is only one active native stricture. Instead, it indicates there are multiple folding pathways leading to multiple native structures. The inter-conversions between multiple native states were achieved by removing and restoring magnesium ions[15], which are required for TG1I folding. This indicates the folding heterogeneity is an intrinsic property of TG1I ribozyme rather than an artifact such as covalent difference from synthesis or UV-induced crosslinks. Otherwise, the inter-conversions between all subpopulations cannot be

observed. In all the previous studies, however, no information about the free energy barrier separating the multiple native structures was determined, leaving the phenomenon of long-lived memory effects unsolved.

In this study, we have developed a new technique called laser-assisted single-molecule refolding, in which we combined single-molecule FRET with laser-induced T-jump to study kinetically trapped RNAs. We seek to elucidate whether or not the folding heterogeneity and memory effect in intrinsic property of hairpin ribozyme and how high the barrier is to give such a long lasting memory effect. To validate our technique, we have applied this technique to study the melting of DNA duplex and compare the result from the UV melting experiment. We have observed the dissociation of DNA duplexes and the formation of DNA hairpins. For the first time we have driven a molecular reaction using heat and monitored the change at the single molecule level. The close agreement between the two melting curve indicates the accuracy of the temperature calibration of our setup. The minor discrepancy at low temperature region of short DNA duplex is probably from the different equilibrium time of two methods, which has negligible effect on activation barrier height measurement.

We have also applied LASR to study the dimerziation reaction of HIV DIS RNA analogs. In the HIV virus replication cycle, two identical HIV RNAs will form dimer before the viral RNA encapsulation process[76]The DIS RNAs will form a kissing complex first, and then, with the help of enzymes (such as NC protein[102].), two DIS RNAs will form a duplex. This kissing complex intermediate state is very stable indicating a deep kinetic trap. Previous studies have shown that the kissing complex is so stable that it can be crystallized[77]. The dimerization reaction pathway may or may not through the

kissing complex structure[103]. The HIV DIS RNA kissing hairpin analogs are ideal model systems to study kinetically trapped RNA. The LASR experiment on the RNA kissing hairpins shows that we can drive the dimerization reaction with a T-jump. The T-jump made the majority of kissing complex dissociate, as the barrier height is low for dissociation. The melting curve from LASR studies is more cooperative for dimerization than for dissociation indicating there are more hydrogen bonds and other interaction to be broken during the dimerization. The kinetic barrier heights were calculated for the dimerization and dissociation reaction. The barrier of dissociation agrees well with the surface Plasmon resonance bulk measurement. While the barrier of dissociation has no comparison as it is impossible to study a reaction starting from an intermediate state in bulk. In the LASR study of the non-kissing mutant, we see no kissing at all; but there is still a small portion of RNA that will dimerize after the T-jump. Assuming the mutation in the loop does not affect the dimerizaiton through non-kissing interactions, the ratio of two competing reaction pathways (kissing *vs.* non-kissing) is calculated to be 4 at all temperatures. This suggests that dimerization reaction pathway through the kissing complex is predominant. The surprising result from the 'non-dimerization' mutant shows the ability of LASR to introduce and monitor transient meta-stable species, which is not possible in ensemble bulk measurement. The lifetime of the meta-stable specie is calculated to be 7 seconds, which is a reasonable number for 7 effective base pairs under the salt conditions in this study. These again proved the advantage of LASR in generating and capturing the transient meta-stable species and then study a reaction starting from it.

LASR study applied to the hairpin ribozyme has introduced inter-conversion

between all subpopulations that have distinct memory effects in their folding dynamics. This memory effect was extremely long lived at room temperature according to previous studies[14]. This result clearly shows that the memory effect is an intrinsic property of heterogeneous RNA folding. Otherwise, some of the subpopulation should not be accessible if it is due to an artifact such as covalent difference. It supports the hypothesis of multiple native structures separated by high-energy barriers forming kinetic traps. For the first time, we have also quantitatively determined the enthalpy and entropy of the barrier heights of the transition between wells of the RNA folding free energy diagram. This quantitative information shed a light into the mechanism of the long lasting memory effect. The enthalpy change for the inter-conversion barrier is lower than the tertiary and secondary structure disruption. This indicates there are not many interactions to be broken for the inter-conversion between subpopulations. Previous studies have shown the loss of memory effect when loop A and B were eliminated, suggesting the memory effect originates from the loop region[14]. Also the mutants with mutation in the tertiary interaction showed no effect in eliminating the folding heterogeneity[21], which indicates the tertiary interactions stabilizing the docked structure were not the origin of the memory effect. Instead, all evidence indicates subtle conformational or topological difference may have played important role in the memory effect[31].

The infrequent inter-conversion lays in the fact that the entropy is very negative and contributes 30% of the barrier height. This is huge contribution considering the small difference in the barrier heights of inter-conversion and undocking. Another effect from the negative entropy is that the free energy barrier height increases as the

temperature goes up. So even at the high temperature of 88 °C, where most of the hairpin ribozyme dissociates, inter-conversion is still not prevailing (~11%). The small increase in the inter-conversion at high temperature is due to the broader Boltzmann distribution of the molecules into the higher energy level.

The physical meaning of a negative entropy change of the reaction is a rigid transition state. For the inter-conversion of a subpopulation to happen, the hairpin ribozyme has to form a rigid transition state. In the context of the hairpin ribozyme secondary structure, one possible mechanism of inter-conversion could involve the stretch of the S-turn, which also causes a memory effect in the Sarcin-Ricin loop[93]. S-turn was locked inside the loop B. The steric effect or the tension may have prevented the S-turn from freely changing conformation. In this mechanism, not many interactions need to be broken, which agrees with the experiment observation. However, this process is entropically unfavorable. The S-turn is fully buried in the contact between loop A and B. If there is heterogeneity in S-turn conformation, it is very likely it will cause heterogeneity in the stability of the docked structure and hence heterogeneous undocking rate constant. The free energy barrier height for inter-conversion is as high as it for dissociation, which results in a barely observed inter-conversion. However, alternative explanations are possible, considering S-turn alone may not be enough to create 4 distinct subpopulations.

For future study, Bishnu Paudel will make mutations to the S-turn to further identify the interactions responsible for the memory effect. An extra nucleotide can be added before, after or in the middle of the S-turn to change the length of S-turn. If the memory effect in the hairpin ribozyme folding was due to stretching of the S-turn, the

length of it should play an important role. To prevent the added nucleotide from interacting with other nucleotides, a modified nucleotide can be used. Also, mutation can be made to alter the base paring or tertiary interaction inside of the S-turn. This will help to understand the interactions needed to be broken for the inter-conversion.

For further broad the application of LASR, Hansini Mundigala will apply LASR to study the dimerization reaction of DIS RNA with the presence of antibody or NC proteins. This will help understand the dimerization mechanism of DIS RNA and provide useful insight into preventing the dimerization as a way of stopping the HIV replication cycle.

REFERENCES:

1 Mattick, J. S. The hidden genetic program of complex organisms. *Sci Am* **291**, 60-67 (2004).

2 Kruger, K. *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**, 147-157, doi:0092-8674(82)90414-7 [pii] (1982).

3 Cai, Z. & Tinoco, I., Jr. Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry* **35**, 6026-6036, doi:10.1021/bi952985g

bi952985g [pii] (1996).

4 Valadkhan, S. & Manley, J. L. Splicing-related catalysis by protein-free snRNAs. *Nature* **413**, 701-707 (2001).

5 Benne, R. & Sloof, P. Evolution of the mitochondrial protein synthetic machinery. *Biosystems* **21**, 51-68 (1987).

6 Guo, S., Tschammer, N., Mohammed, S. & Guo, P. Specific delivery of therapeutic RNAs to cancer cells via the dimerization mechanism of phi29 motor pRNA. *Hum Gene Ther* **16**, 1097-1109, doi:10.1089/hum.2005.16.1097 (2005).

7 Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498 (2001).

8 Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E. & Breaker, R. R. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10**, 701-707 (2003).

9 Gilber, W. Origin of life : The RNA world. *Nature* **319**, 618 (1986).

10   Pyle, A. M. Metal ions in the structure and function of RNA. *J Biol Inorg Chem* **7**, 679-690 (2002).

11   Woodson, S. A. Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr Opin Chem Biol* **9**, 104-109 (2005).

12   Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-19 (1997).

13   Treiber, D. K. & Williamson, J. R. Beyond kinetic traps in RNA folding. *Curr Opin Struct Biol* **11**, 309-314 (2001).

14   Zhuang, X. *et al.* Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**, 1473-1476 (2002).

15   Solomatin, S. V., Greenfeld, M., Chu, S. & Herschlag, D. Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature* **463**, 681-684, doi:nature08717 [pii]

10.1038/nature08717 (2010).

16   Herschlag, D., Khosla, M., Tsuchihashi, Z. & Karpel, R. L. An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J* **13**, 2913-2924 (1994).

17   Herschlag, D. RNA chaperones and the RNA folding problem. *J Biol Chem* **270**, 20871-20874 (1995).

18   Ha, T. *et al.* Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **419**, 638-641, doi:10.1038/nature01083

nature01083 [pii] (2002).

19   Pyle, A. M. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* **37**, 317-336, doi:10.1146/annurev.biophys.37.032807.125908 (2008).

20   Mikulecky, P. J. *et al.* Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol* **11**, 1206-1214, doi:nsmb858 [pii]

10.1038/nsmb858 (2004).

21   Rueda, D. *et al.* Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance. *Proc Natl Acad Sci U S A* **101**, 10066-10071, doi:10.1073/pnas.0403575101

0403575101 [pii] (2004).

22   Rupert, P. B. & Ferre-D'Amare, A. R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **410**, 780-786 (2001).

23   Walter, N. G., Hampel, K. J., Brown, K. M. & Burke, J. M. Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer. *EMBO J* **17**, 2378-2391, doi:10.1093/emboj/17.8.2378 (1998).

24   Walter, N. G., Burke, J. M. & Millar, D. P. Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction. *Nat Struct Biol* **6**, 544-549, doi:10.1038/9316 (1999).

25   Nahas, M. K. *et al.* Observation of internal cleavage and ligation reactions of a ribozyme. *Nat Struct Mol Biol* **11**, 1107-1113, doi:nsmb842 [pii]

10.1038/nsmb842 (2004).

26   Tan, E. *et al.* A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc Natl Acad Sci U S A* **100**, 9308-9313, doi:10.1073/pnas.1233536100

1233536100 [pii] (2003).

27   Walter, N. G. & Burke, J. M. The hairpin ribozyme: structure, assembly and catalysis. *Curr Opin Chem Biol* **2**, 303, doi:S1367-5931(98)80073-2 [pii] (1998).

28   Zhao, R. & Rueda, D. RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **49**, 112-117, doi:S1046-2023(09)00103-0 [pii]

10.1016/j.ymeth.2009.04.017 (2009).

29   Bartley, L. E., Zhuang, X., Das, R., Chu, S. & Herschlag, D. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J Mol Biol* **328**, 1011-1026, doi:S0022283603002729 [pii] (2003).

30   Okumus, B., Wilson, T. J., Lilley, D. M. & Ha, T. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys J* **87**, 2798-2806, doi:10.1529/biophysj.104.045971

S0006-3495(04)73750-9 [pii] (2004).

31   Ditzler, M. A., Rueda, D., Mo, J., Hakansson, K. & Walter, N. G. A rugged free energy landscape separates multiple functional RNA folds throughout denaturation. *Nucleic Acids Res* **36**, 7088-7099, doi:gkn871 [pii]

10.1093/nar/gkn871 (2008).

32   Bokinsky, G. *et al.* Single-molecule transition-state analysis of RNA folding. *Proc Natl Acad Sci U S A* **100**, 9302-9307 (2003).

33   Rueda, D. & Walter, N. G. Fluorescent energy transfer readout of an aptazyme-based biosensor. *Methods Mol Biol* **335**, 289-310, doi:1-59745-069-3:289 [pii]

10.1385/1-59745-069-3:289 (2006).

34   Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78-86, doi:10.1006/meth.2001.1217

S1046-2023(01)91217-4 [pii] (2001).

35   Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. & Ha, T. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys J* **86**, 2530-2537, doi:S0006-3495(04)74308-8 [pii]

10.1016/S0006-3495(04)74308-8 (2004).

36   Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K., Jr. Direct observation of the rotation of F1-ATPase. *Nature* **386**, 299-302 (1997).

37   Christian, T. D., Romano, L. J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc Natl Acad Sci U S A* **106**, 21109-21114, doi:0908640106 [pii]

10.1073/pnas.0908640106 (2009).

38   Lakowicz, J. R. Principles of Fluorescence Spectroscopy. *Plenum Publishing Corporation* **2nd edition** (1999).

39   Lee, C. Y. Fluorescence Spectroscopy. *Current Protocols Essential Laboratory Techniques* **2.4** (2009).

40   Rueda, D. & Walter, N. G. Single molecule fluorescence control for nanotechnology. *J Nanosci Nanotechnol* **5**, 1990-2000 (2005).

41   Axelrod, D. Total internal reflection fluorescence microscopy in cell biology. *Traffic* **2**, 764-774, doi:021104 [pii] (2001).

42   Selvin, P. R. & Ha, T. *Single Molecule Techniques: A Laboratory Manual*. 1st edition edn, (Cold Spring Harbor Laboratory Press, 2007).

43   Yasuda, R., Noji, H., Yoshida, M., Kinosita, K., Jr. & Itoh, H. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. *Nature* **410**, 898-904 (2001).

44   Diez, M. *et al.* Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. *Nat Struct Mol Biol* **11**, 135-141, doi:10.1038/nsmb718 nsmb718 [pii] (2004).

45   Yildiz, A. *et al.* Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061-2065 (2003).

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

47   Smith, D. E. *et al.* The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-752 (2001).

48   Chemla, Y. R. *et al.* Mechanism of force generation of a viral DNA packaging motor. *Cell* **122**, 683-692 (2005).

49   Blanchard, S. C., Gonzalez, R. L., Kim, H. D., Chu, S. & Puglisi, J. D. tRNA selection and kinetic proofreading in translation. *Nat Struct Mol Biol* **11**, 1008-1014 (2004).

50   Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Jr., Puglisi, J. D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-12898 (2004).

51   Xie, Z., Srividya, N., Sosnick, T. R., Pan, T. & Scherer, N. F. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc Natl Acad Sci U S A* **101**, 534-539, doi:10.1073/pnas.2636333100

2636333100 [pii] (2004).

52   Min, W. *et al.* Fluctuating enzymes: lessons from single-molecule studies. *Acc Chem Res* **38**, 923-931, doi:10.1021/ar040133f (2005).

53   English, B. P. *et al.* Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat Chem Biol* **2**, 87-94 (2006).

54   Digman, M. A. *et al.* Fluctuation correlation spectroscopy with a laser-scanning microscope: exploiting the hidden time structure. *Biophys J* **88**, L33-36, doi:S0006-3495(05)73358-0 [pii]

10.1529/biophysj.105.061788 (2005).

55   Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys J* **72**, 1335-1346, doi:S0006-3495(97)78780-0 [pii]

10.1016/S0006-3495(97)78780-0 (1997).

56   Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* **5**, 491-505, doi:nmeth.1218 [pii]

10.1038/nmeth.1218 (2008).

57   Trache, A. & Meininger, G. A. Atomic force microscopy (AFM). *Curr Protoc Microbiol* **Chapter 2**, Unit 2C 2, doi:10.1002/9780471729259.mc02c02s8 (2008).

58   Jaiswal, J. K. & Simon, S. M. Imaging single events at the cell membrane. *Nat Chem Biol* **3**, 92-98, doi:nchembio855 [pii]

10.1038/nchembio855 (2007).

59   Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat Methods* **5**, 507-516, doi:nmeth.1208 [pii]

10.1038/nmeth.1208 (2008).

60   Pereira, M. J. *et al.* Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis. *J Mol Biol* **382**, 496-509, doi:S0022-2836(08)00862-0 [pii]

10.1016/j.jmb.2008.07.020 (2008).

61   Posson, D. J. & Selvin, P. R. Extent of voltage sensor movement during gating of shaker K+ channels. *Neuron* **59**, 98-109, doi:S0896-6273(08)00415-7 [pii]

10.1016/j.neuron.2008.05.006 (2008).

62   Tahmassebi, D. C. & Millar, D. P. Fluorophore-quencher pair for monitoring protein motion. *Biochem Biophys Res Commun* **380**, 277-280, doi:S0006-291X(09)00093-X [pii]

10.1016/j.bbrc.2009.01.055 (2009).

63   Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J* **94**, 1826-1835, doi:S0006-3495(08)70621-0 [pii]

10.1529/biophysj.107.117689 (2008).

64   Crawford, D. J., Hoskins, A. A., Friedman, L. J., Gelles, J. & Moore, M. J. Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA* **14**, 170-179, doi:rna.794808 [pii]

10.1261/rna.794808 (2008).

65   Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* **3**, 891-893, doi:nmeth934 [pii]

10.1038/nmeth934 (2006).

66   Hodak, J. H., Downey, C. D., Fiore, J. L., Pardi, A. & Nesbitt, D. J. Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc Natl Acad Sci U S A* **102**, 10505-10510 (2005).

67   McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **91**, 1941-1951, doi:S0006-3495(06)71903-8 [pii]

10.1529/biophysj.106.082487 (2006).

68   Ballew, R. M., Sabelko, J. & Gruebele, M. Direct observation of fast protein folding: the initial collapse of apomyoglobin. *Proc Natl Acad Sci U S A* **93**, 5759-5764 (1996).

69   Ballew, R. M., Sabelko, J. & Gruebele, M. Observation of distinct nanosecond and microsecond protein folding events. *Nat Struct Biol* **3**, 923-926 (1996).

70   Ma, H. *et al.* Exploring the energy landscape of a small RNA hairpin. *J Am Chem Soc* **128**, 1523-1530 (2006).

71   HITRAN. *http://savi.weber.edu/hi_plot/* (Accessed July 2007).

72   database, H. *http://savi.weber.edu/hi_plot/*.

73   Zhao, R. *et al.* Laser-assisted single-molecule refolding (LASR). *Biophys J* **99**, 1925-1931, doi:S0006-3495(10)00889-1 [pii]

10.1016/j.bpj.2010.07.019 (2010).

74   Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-3415 (2003).

75   Andersen, A. A. & Collins, R. A. Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme. *Proc Natl Acad Sci U S A* **98**, 7730-7735, doi:10.1073/pnas.141039198

141039198 [pii] (2001).

76   Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc Natl Acad Sci U S A* **93**, 5572-5577 (1996).

77   Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C. & Dumas, P. Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat Struct Biol* **8**, 1064-1068, doi:10.1038/nsb727

nsb727 [pii] (2001).

78   Bernacchi, S., Henriet, S., Dumas, P., Paillart, J. C. & Marquet, R. RNA and DNA binding properties of HIV-1 Vif protein: a fluorescence study. *J Biol Chem* **282**, 26361-26368, doi:M703122200 [pii]

10.1074/jbc.M703122200 (2007).

79   Russell, R. RNA misfolding and the action of chaperones. *Front Biosci* **13**, 1-20, doi:2557 [pii] (2008).

80   Kirschner, K., Eigen, M., Bittman, R. & Voigt, B. The Binding of Nicotinamide-Adenine Dinucleotide to Yeast D-Glyceraldehyde-3-Phosphate Dehydrogenase: Temperature-Jump Relaxation Studies on the Mechanism of an Allosteric Enzyme. *Proc Natl Acad Sci U S A* **56**, 1661-1667 (1966).

81   Phillips, C. M., Mizutani, Y. & Hochstrasser, R. M. Ultrafast thermally induced unfolding of RNase A. *Proc Natl Acad Sci U S A* **92**, 7292-7296 (1995).

82   Proctor, D. J. *et al.* Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. *Biochemistry* **43**, 14004-14014 (2004).

83   Rueda, D., Hsieh, J., Day-Storms, J. J., Fierke, C. A. & Walter, N. G. The 5' leader of precursor tRNAAsp bound to the Bacillus subtilis RNase P holoenzyme has an extended conformation. *Biochemistry* **44**, 16130-16139, doi:10.1021/bi0519093 (2005).

84   Stancik, A. L. & Brauns, E. B. Rearrangement of partially ordered stacked conformations contributes to the rugged energy landscape of a small RNA hairpin. *Biochemistry* **47**, 10834-10840, doi:10.1021/bi801170c (2008).

85   Fiore, J. & Nesbitt, D. Personal communication. (2010).

86   Rastogi, T., Beattie, T. L., Olive, J. E. & Collins, R. A. A long-range pseudoknot is required for activity of the Neurospora VS ribozyme. *EMBO J* **15**, 2820-2825 (1996).

87   Guo, P., Zhang, C., Chen, C., Garver, K. & Trottier, M. Inter-RNA interaction of phage phi29 pRNA to form a hexameric complex for viral DNA transportation. *Mol Cell* **2**, 149-155, doi:S1097-2765(00)80124-0 [pii] (1998).

88   Moore, P. B. & Steitz, T. A. The structural basis of large ribosomal subunit function. *Annu Rev Biochem* **72**, 813-850, doi:10.1146/annurev.biochem.72.110601.135450 (2003).

89   Khvorova, A., Lescoute, A., Westhof, E. & Jayasena, S. D. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nat Struct Biol* **10**, 708-712, doi:10.1038/nsb959

nsb959 [pii] (2003).

90   Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787-798, doi:S0092867403010183 [pii] (2003).

91   Stone, M. D. *et al.* Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-461, doi:nature05600 [pii]

10.1038/nature05600 (2007).

92   Butcher, S. E., Allain, F. H. & Feigon, J. Solution structure of the loop B domain from the hairpin ribozyme. *Nat Struct Biol* **6**, 212-216, doi:10.1038/6651 (1999).

93   Korennykh, A. V., Plantinga, M. J., Correll, C. C. & Piccirilli, J. A. Linkage between substrate recognition and catalysis during cleavage of sarcin/ricin loop RNA by restrictocin. *Biochemistry* **46**, 12744-12756, doi:10.1021/bi700931y (2007).

94   Brion, P. & Westhof, E. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct* **26**, 113-137, doi:10.1146/annurev.biophys.26.1.113 (1997).

95   Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. The protein folding problem. *Annu Rev Biophys* **37**, 289-316, doi:10.1146/annurev.biophys.37.092707.153558 (2008).

96   van Oijen, A. M. *et al.* Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* **301**, 1235-1238, doi:10.1126/science.1084387

301/5637/1235 [pii] (2003).

97   Lu, H. P., Xun, L. & Xie, X. S. Single-molecule enzymatic dynamics. *Science* **282**, 1877-1882 (1998).

98   Baldini, G. *et al.* Evidence of discrete substates and unfolding pathways in green fluorescent protein. *Biophys J* **92**, 1724-1731, doi:S0006-3495(07)70980-3 [pii]

10.1529/biophysj.106.093567 (2007).

99   Sigler, P. B. An analysis of the structure of tRNA. *Annu Rev Biophys Bioeng* **4**, 477-527, doi:10.1146/annurev.bb.04.060175.002401 (1975).

100   Altuntop, M. E., Ly, C. T. & Wang, Y. Single-molecule study of ribosome hierarchic dynamics at the peptidyl transferase center. *Biophys J* **99**, 3002-3009, doi:S0006-3495(10)01032-5 [pii]

10.1016/j.bpj.2010.08.037 (2010).

101   Huang, Z., Pei, W., Jayaseelan, S., Shi, H. & Niu, L. RNA aptamers selected against the GluR2 glutamate receptor channel. *Biochemistry* **46**, 12648-12655, doi:10.1021/bi701036p (2007).

102   Cosa, G. *et al.* Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys J* **87**, 2759-2767 (2004).

103   Vo, M. N., Barany, G., Rouzina, I. & Musier-Forsyth, K. HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper". *J Mol Biol* **386**, 789-801, doi:S0022-2836(08)01607-0 [pii]

10.1016/j.jmb.2008.12.070 (2009).

ABSTRACT

LASER-ASSISTED SINGLE-MOLECULE REFOLDING

by

RUI ZHAO

May 2011

Advisor: Prof. David Rueda

Major: Chemistry (Physical Chemistry)

Degree: Doctor of Philosophy

Non-coding RNAs must fold into precise secondary and tertiary structures in order to perform the biological functions. Due to the flexibility of RNA, the RNA folding energy landscape can be rugged and full of local minimum (kinetic trap). To provide a means to study kinetically trapped RNAs, we have developed a new technique combining single-molecule FRET detection with laser induced temperature jump. We have calibrated the magnitude of the temperature jump with 1°C accuracy using gold micro-size sensor. The accuracy of temperature calibration was confirmed by close agreement between single-molecule and bulk DNA duplex melting experiments.

HIV 1 DIS RNAs form a kissing complex in the loop region and proceed to the extended duplex structure with the help of enzymes or other cofactors in the later stage of viral replication. The kissing complex itself is very stable, which makes it a unique optimal model system to study kinetically trapped RNAs. The application of LASR to kissing hairpins has allowed us, for the first time, to drive a molecular reaction and monitor the process at the single molecule level. The melting curve for the dissociation

and dimerization was used to estimate the thermodynamic properties of the reaction, such as melting temperature, cooperativity, and the enthalpy change. Mutational studies have allowed dissection of the contribution of base pairs to kissing complex stability. And the ratio of the competing reaction pathways was determined.

LASR experiments designed to the study of the origin of the memory effect in the hairpin ribozyme folding introduced inter-conversions between the subpopulations of hairpin ribozyme with distinct undocking kinetic rates. This provides strong evidence for the hypothesis that the memory effect is an intrinsic property of hairpin ribozyme folding. Eyring analysis was adopted to fit the enthalpic and entropic height of the inter-conversion barrier. Our results suggest that for inter-conversion to occur, surprisingly, a small number of interactions are broken. However inter-conversion is rare due to a large negative entropic term. Negative entropy indicates a rigid transition state for inter-conversion and an increased free energy barrier height with increased temperature. This explains why there are few inter-conversions even at high temperatures such as 78 ˚C. The entropic barrier may primarily arise from the stretching of the S-turn.

AUTOBIOGRAPHICAL STATEMENT

Name: Rui Zhao

Education:
- **Ph.D. in Chemistry**   *(2005- Present)*
  Wayne State University, Detroit, MI
- **B.S. in Chemical Physics**   *(2001- 2005)*
  University of Science & Technology of China, Hefei, Anhui, China

Publications:
- R. Zhao, and D. Rueda, *RNA folding dynamics by single-molecule fluorescence resonance energy transfer.* Methods, 2009.
- R. Zhao, M. Marshal, E. A. Alemán, R. Lamichhane and D. Rueda. *Laser-Assisted Single-Molecule Refolding,* Biophysics Journal, 2010
- E. A. Alemán, H. Yu, R. Zhao, S. L. Brock, and D. Rueda, *Cooperative Behavior in CdSe Quantum Dot Assemblies (Aerogels) as Probed by Single Particle Fluorescence Intermittency (submitted)*
- N. Salim, R. Zhao, R. Lamichhane, D. Rueda and A. L Feig, *Thermodynamic and Kinetic Analysis of an RNA Kissing Interaction and Its Resolution into an Extended Duplex (submitted)*
- R. Zhao, B. Paudel, and D. Rueda, *Exploring the memory effect in RNA ribozyme folding (in preparation)*

Presentations:

Oral

1. ***American Physical Society March Meeting 2010****, Portland, OR "Laser-Assisted Single-Molecule refolding*

2. **Michigan RNA Society tenth Meeting 2009**, Albion, MI
   "Laser-Assisted Single-Molecule refolding"

3. **Rustbelt RNA meeting 2008**, Mt. Sterling, OH
   "Laser-Assisted Single-Molecule refolding"

4. **10th Wayne State University Graduate Symposium 2008**, Detroit, MI
   "Laser-Assisted Single-Molecule refolding""

Poster

- **Michigan RNA Society Eleventh Meeting 2010**, Lansing, MI
  "Laser-Assisted Single-Molecule refolding"
- **Rustbelt RNA meeting 2009**, Mt. Sterling, OH
  "Laser-Assisted Single-Molecule refolding"
- **11th Wayne State University Graduate Symposium 2009**, Detroit, MI
  "Laser-Assisted Single-Molecule refolding"
- **Biophysical Society 52nd Annual Meeting 2008,** L.A., CA
  "RNA Folding Dynamics by Single-Molecule Temperature Jump kinetics"
- **Twelfth Annual Meeting of the RNA society 2007**, Madison, WI
  "RNA Folding Dynamics by Single-Molecule Temperature Jump kinetics"