Dissertation

Biomolecule Conformational Dynamics Studies using Two-Beam Fluorescence Fluctuation Spectroscopy: Dynamics of DNA Hairpin Structure

Submitted By

Jaemyeong Jung

Department of Chemistry

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall, 2005

COLORADO STATE UNIVERSITY

NOVEMBER 08, 2005

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY JAEMYEONG JUNG ENTITLED "BIOMOLECULE CONFORMATIONAL DYNAMICS STUDIES USING TWO-BEAM FLUORESCENCE FLUCTUATION SPECTROSCOPY: DYNAMICS OF DNA HAIRPIN STRUCTURE" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Adviser

Co-Adviser

Department Head/Director

ABSTRACT OF DISSERTATION

Biomolecule Conformational Dynamics Studies using Two-Beam Fluorescence Fluctuation Spectroscopy: Dynamics of DNA Hairpin Structure

This dissertation describes a new fluorescence-based technique for understanding biomolecule conformational dynamics, referred to as two-beam fluorescence fluctuation spectroscopy (2bFFS). In this approach, the fluorescence fluctuations associated with biomolecule conformational changes were observed as the molecules flowed sequentially between two spatially offset, microscopic detection volumes, and analyzed simultaneously from three different vantage points—cross-correlation analysis of the two detection channels relative to each other, autocorrelation analysis of the two detection channels independently, and photon counting histogram (PCH) analysis. The cross-correlation function characterizes the hydrodynamic properties (flow rate and diffusion rate) of the molecules under study, independent of conformational fluctuations. These parameters can then be used to constrain the measured autocorrelation function, from which precise measurements of the conformational kinetics parameters are obtained. The combined PCH analysis can resolve the equilibrium distribution between different conformations of the molecules. To demonstrate these principles, dynamic equilibrium between the folded and unfolded conformations of single stranded DNA hairpin molecules was investigated. The DNA hairpins contained a polythymine loop and a five base-pair stem sequence that was end-labeled with a fluorescent dye and a quencher. Folding and unfolding of the DNA hairpin structure caused the dye fluorescence to fluctuate on the same characteristic time scale as the conformational change. In addition to unambiguously characterizing the relaxation times of the process, the analysis revealed

non-exponential relaxation kinetics and DNA size-dependent folding times characteristic of dynamic heterogeneity in the DNA hairpin-forming mechanism. Analysis of the equilibrium distribution suggests a three-state mechanism for the reaction, involving a rapid equilibrium between open and a stable intermediate form of the DNA hairpin. The final, closed form of the DNA hairpin is suggested to be stable on a much longer time scale than that of the present FFS measurements.

Jaemyeong Jung
Department of Chemistry
Colorado State University
Fort Collins, CO. 80523
Fall, 2005

To my wife, Jee

and

my beloved ones, Claire and Daniel

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 BIOMOLECULE CONFORMATIONAL CHANGES

"Time never stops." It is one of the most common truths that everyone understands from simple experience. In some aspect, it can be rephrased as, "Everything is changing." Modern science shows that the more we view the world from a microscopic perspective, the more phenomena become dynamic. This dynamics is what controls the world surrounding us and sustains our biological systems. Understanding this dynamics on the microscopic level has been the main target of studies in all areas of physics, chemistry, and biology. This dissertation discusses an essential aspect of the dynamics in biological systems, namely conformational fluctuations of biomolecules. These fluctuations originate from the spatial rearrangement of atoms in the molecule, and form

many unique and different over all shapes, or conformations. The reconfigurations are known to be associated with particular properties that affect biological function, reactivity, transportation, assembly, tendency to aggregation, and potential cytotoxicity. These conformational changes are found in virtually every functional loci of biological processes. Protein folding is the most typical example of the dynamics. When DNA codes direct the generation of a protein during the translation process in a cell, the long sequence of the protein is somewhat like a random polymer chain. The chain folds quickly by itself, and then a unique conformation results from the self-assembly. This unique protein structure gives the essence of its biological function. Failure of the folding process may cause significant outcomes in the biological system. Recent studies show evidence relating mis-folding events to many kinds of genetic diseases such as Alzheimer's disease, Bovine Spongiform Encephalopathy (BSE) known as Mad cow disease, Cystic fibrosis, etc. In order to find ways to treat such diseases, intensive efforts are focused on the study of protein folding problem. The story of protein folding has not yet led to treatments for the diseases involved, but this could happen in the near future with better understanding of how proteins fold. We can also find similar stories in the dynamics of ribozyme docking and undocking, and DNA/RNA hairpin opening and closing. Figure 1.1 shows time scales of some biological process involving fast conformational dynamics (from nano-seconds to milli-seconds).

Despite this long journey of research, we are still in the early stages of unraveling these fundamental phenomena. Atomistic molecular dynamics simulations currently can access only the early part of the processes (~10 ns). A recent theoretical technique called

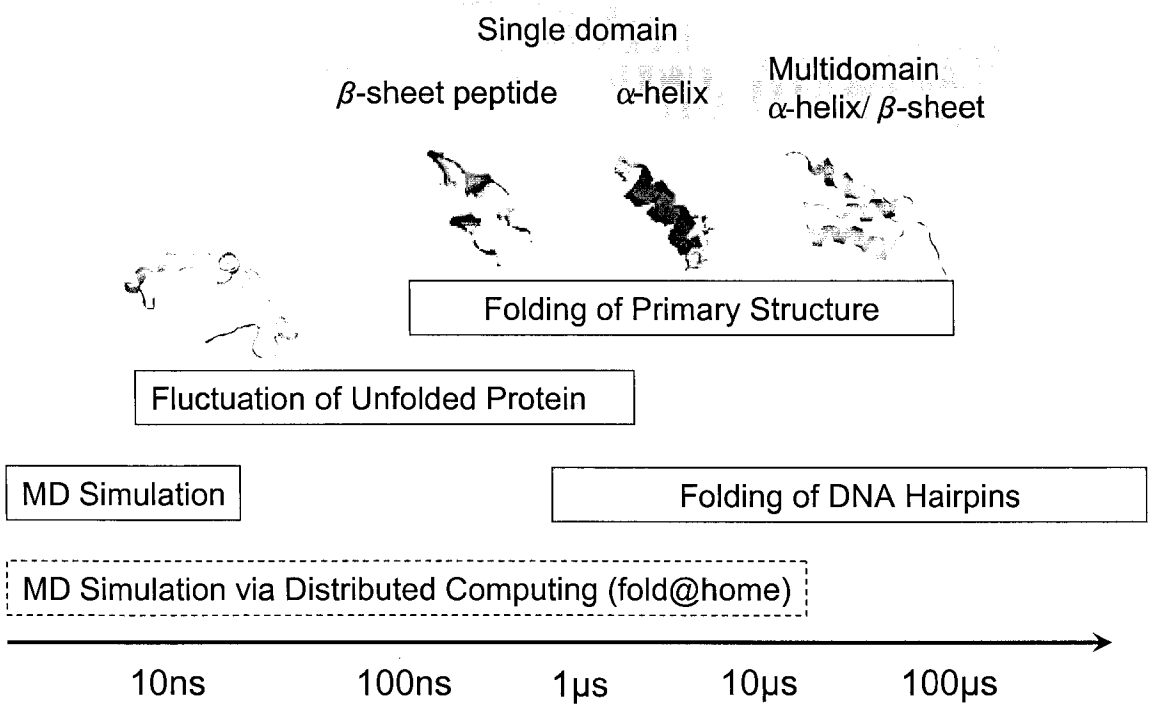

Figure 1.1 Examples of fast conformational dynamics in biomolecules (i.e. protein folding) and their corresponding time scales
DNA hairpin folding and unfolding reaction shows micro-second time scale. Current molecular dynamics simulation approach can access to the early part of these dynamics (~10 ns) with atomistic details.

distributed computing was developed for the study of much longer time process. Several promising results predicted the dynamic behaviors up to a few tens of microseconds based on the statistical interpretation of sampled data from large scale, parallel molecular dynamics simulations.[1-5] Experimentally, measuring the fast dynamics of complex biomolecules has not been an easy task with conventional techniques, even in the view of the latest cutting edge technologies. In fact, various controversies on the thermodynamic and kinetic description for most biomolecule conformational changes still remain. However, knowledge of conformational dynamics might be the first step to be achieved for the complete understanding of the biomolecules' functioning and the purpose assigned to them by Mother Nature. This point of view originally motivated the research described in this dissertation. We designed a unique analytical tool for the approach toward biomolecule conformational dynamics and applied it for the study of a model system—the DNA hairpin folding and unfolding reaction.

1.2 CONFORMATIONAL DYNAMICS OF DNA HAIRPINS

1.2.1 DNA Hairpin Structure

In addition to the double helix structure deciphered by Crick and Watson in 1953, DNA has alternative structures containing certain order in local sequences. Some examples are slipped-strand DNA, left-handed Z-DNA, cruciform DNA (four-way junctions) and triple-stranded DNA. Localized unwinding of the DNA structures are commonly found in the functional center of many biological processes in the cell, such as replication, recombination[6], mutagenic events[7], DNA repair and transcription. This

unwinding may occur as a result of local opening and closing of individual base-pairs,[8] or can be driven by DNA binding proteins, leading to the formation of single-stranded regions in the DNA structures.[9, 10] Hairpins are regions of unwound DNA forming irregular structures by base-pairing between neighboring self-complementary sequences. In general, DNA strands with self-complementary base sequences can form hairpin structures consisting of single stranded subsequences bounded by a stem containing contiguous complimentary base pairs.(Figure 1.2) Hairpin containing domains may play a significant role in control mechanism for gene expression. It has been shown that DNA hairpins can serve as intermediates in genetic recombination, [11-13] and as protein recognition sites, and can regulate transcription *in vivo*.[14, 15] The formation of hairpins and their involvement in biological processes is now well known in both prokaryotic and eukaryotic system.[14, 16-18] Recent biomedical researches shows that the formation of DNA hairpins by GC-rich triplet repeat and inordinate expansion of these triplets during DNA replication, leads to several genetic diseases such as myotonic dystrophy, a progressive neuromuscular disorder.[19, 20] Hairpins were also proposed as antisense drugs with high resistance to degradation by exonucleases.[21] Short hairpin structures are useful drug targets because their overall shape and geometry differ significantly from regular double-stranded DNA, and are used in hybridization studies to investigate the effect of secondary structure on probe hybridization.[22-25] Understanding the structural and dynamical aspects of how DNA hairpins fold and behave can be a starting point for detailed understanding the mechanism and control of various biological processes. This involves characterizing the stability and dynamics of such structures, and determining the

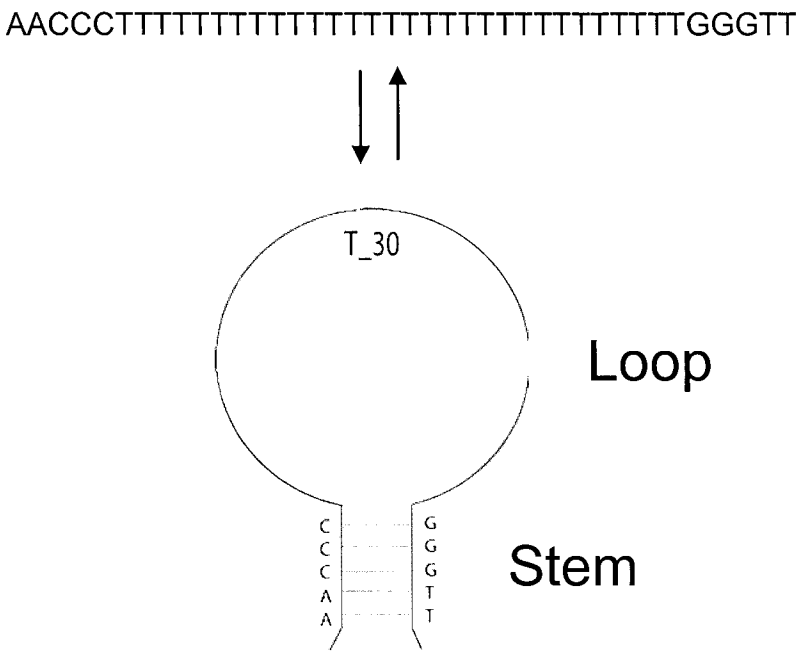

Figure 1.2 A schematic description of DNA hairpin formation
Simple DNA hairpin structure consists of single strands with self-complementary base sequences. Loop part is the stranded subsequences bounded by Stem, the contiguous base pairs. T_30 represent the thirty thymines in the loop. This structure has opening and closing dynamics with conformational changes.

factors that affect their stability. Recent applications of DNA hairpins in biosensor research using the concept of "molecular beacons" has shown promising results.[26] Furthermore, the knowledge of the basic kinetics will answer many fundamental questions regarding the folding of larger RNA which is more complex and involves parallel pathways and kinetic traps.[27-32]

1.2.2. Fundamental Questions in Biological Processes

The folding and unfolding dynamics of DNA hairpin molecules has been investigated for more than thirty years. Along the way, comparisons have been made to the protein folding problem. One of the most fundamental questions that has been raised is: Why do proteins fold so fast?[33] This question is frequently referred to as Levinthal's paradox.[34] In 1969, it was argued that it should take forever for proteins to fold if they randomly sampled all conformational space. For a simple example, let's consider a very short protein chain consisting of 100 monomers. If each monomer has only two conformational states, then, the total number of possible chain conformations would be 2^{100}. We can assume that the protein spends about 10 pico-seconds in the search of each chain conformation until it finds the native conformational state. In this scenario, the protein needs a folding time of $2^{100}\times 10$ pico-seconds, which is much longer than the age of universe! In a general explanation to this paradox, scientists recognize that the conformational search for the native state of the protein is not random, but biased on a "funnel-like" energy landscape with a pronounced global minimum corresponding to the native state.[35-39] The folding process is accelerated on the free energy surface through cooperative interactions between monomers. Since larger molecules like proteins have

many more vibrational motions available to them, they have various ways to scramble the energy, and do so on a faster time scale than small molecules. Zwanzig et al. showed that if an energy penalty of the order of a few k_BT is imposed to break favorable native interactions, then the conformational search time can be reduced to biologically relevant times.[40]

This question is opposite to the one posed here in connection with the folding of DNA or RNA hairpins. Again, let's consider a DNA hairpin consisting of about ten nucleotides as a semiflexible polymer chain. Theoretical models describing the characteristic time for two ends of the polymer chain to come into contact predict ~40 nano-seconds for the diffusion limited contact time.[41-43] However, hairpin closing times of ~10-μs were reported experimentally, which is ~250 times slower. It is well known that cyclization times for λ DNA molecule with cohesive ends are also much longer than the end-to-end contact times for a semiflexible polymer chain.[44, 45] Thus, the question underlying the field of hairpin structure in single stranded polynucleotieds folding becomes: *why do hairpins form so slowly*? Although this problem has been studied for decades, many of basic issues remain unresolved. In addition to the puzzles in the relaxation rate itself, the dependence on many basic physical parameters (i.e. loop size, loop sequence, temperature, solvent viscosity, molecular size) has not yet been addressed.

1.2.3 Experimental Observations

The biological functioning of DNA or RNA hairpin structure results from its conformational change. Recent investigation reveals the rapid reaction between folded

and unfolded conformations due to the fluctuation of base pairing in the stem. Kinetic measurements on hairpin formation in single stranded polynucleotides provide insights into the conformational flexibility of these chains, and the nature and strength of the intrachain interactions. In this aspect, oligo-nucleotides that form hairpin structures in solution represent an ideal model system to investigate the nature of hairpin structure in polynucleotides, and to test our understanding of these structures. Experimentally, temperature jump (T-jump) is the traditional approach for the measurement.[46-55] The electrical pulses generated by discharging a capacitor were used to raise the temperature of sample on microsecond time scales.[56] From the relaxation data, e.g. rapid change of absorbance due to base stacking after the temperature jump, thermodynamic and kinetic information such as activation enthalpies and/or zipping rates were analyzed.[57] A modified T-jump technique with a coaxial cable capacitor has been used for some sub-microsecond measurement.[51, 58] In recent years, a variety of new experimental techniques, such as fluorescence correlation spectroscopy (FCS)[59-62], laser induced T-jump[63-67], and single-molecule techniques[68-70] have been used in the investigation of the hairpin kinetics. However, large variations in the measured parameters have been reported, and there are still many fundamental questions unresolved and still in debate. For example, the temperature dependence of the reaction rate provides information on the activation enthalpy of the reaction via Arrhenius analysis. The enthalpy of the transition state relative to the unfolded and folded states is essential for unraveling the mechanism and pathways for the transition. Activation enthalpies ranging from negative to positive values have been reported. For instance, an activation enthalpy of from +25 kcal/mol to +11 kcal/mol for $A_6C_NU_6$ hairpin was reported in separate thermodynamics studies in

1973[52, 71], while other group found ~0 kcal/mol value for $A_4GC_5U_4$ hairpin.[50] Recent FCS studies found the value to be +5~15 kcal/mol for hairpins with loop size of 8~30 bases.[59, 60] Laser induced T-jump measurements yield negative values of -10~13 kcal/mol for similar hairpins.[66, 67] So far, it is not clear whether the differences in the various sets of measurement are because of inherent differences in the data acquisition and analysis. Furthermore, recent observations of non-Arrhenius temperature dependence for the opening and closing rates,[61, 62] or evident deviations from an Arrehnius behavior [60] have added to the complexity of the data interpretation.

In contrast, microsecond DNA hairpin folding and unfolding reaction times have been reported consistently in recent laser T-jump and FCS experiments. However, interestingly, data from single molecule spectroscopy based measurements showed much longer reaction times for similar sized hairpin structures.[69, 72, 73] About 140ms of closing time was reported in the single pair fluorescence resonance energy transfer (FRET) experiment and ~1s of folding time was observed from the single molecule extension measurement under mechanical force of ~14 pN in the presence of 10 mM Mg^{2+}. Although the authors attributed the long time scale to artificial effects under the experimental conditions, such as interaction of hairpin with the derivatized glass surface or presence of the applied force, an alternative interpretation for these observations will be reported here.(See Section 1.2.5)

1.2.4 Theoretical Descriptions

Small DNA hairpins are useful test systems for energy-based structure prediction algorithms, due to their small size and well-defined conformations.[74-78] These

algorithms are widely used in predicting stable secondary structures for DNA and RNA molecules. The ability to predict the kinetics of these hairpins provides an additional constraint on the parameters that describe the nature of their free energies, along with corresponding energy surface and its roughness.[72, 79, 80] Hairpin structures also serve as testing grounds for the growing effort in molecular dynamics simulations of nucleic acids.[4, 5, 81-89] These computational studies suggest: (i) the hairpin folding process has a rugged energy landscape due to intrastrand interactions, such as non-native base-pairing, hydrogen bonding and base-stacking, (ii) most folding pathways lead to dead-ends or ensembles of collapsed states with misfolded traps, and (iii) a distinctly non-Arrhenius temperature dependence for the closing rates. To date, these predictions are not fully consistent with experimental results. However, the observed deviation from the Arrehenius behavior in many experimental and theoretical studies indicate the complexities of the dynamics involved.[62] One way such deviations are observed is through stretched exponential relaxation of the folding reaction (See figure 1.3).[61, 62] Rather than being described with one typical relaxation rate constant, complex reactions can be analyzed as an ensemble of reactions involving with multiple reaction pathways. The energy surface along the reaction coordinate has a different degree of ruggedness determined by experimental conditions. In the analysis of biomolecule folding and unfolding reactions, possible interpretations of the stretched parameter, β, is that $0 \leq \beta \leq 1$, with $\beta \approx 0$ corresponding to a completely random folding mechanism which is a continuous process on smooth energy surface, and $\beta \approx 1$ corresponding to normal two-state Arrhenius kinetics, with single discrete energy barrier.[61, 62] Intermediate β values can indicate varying degrees of heterogeneity in the folding mechanism, which is the case

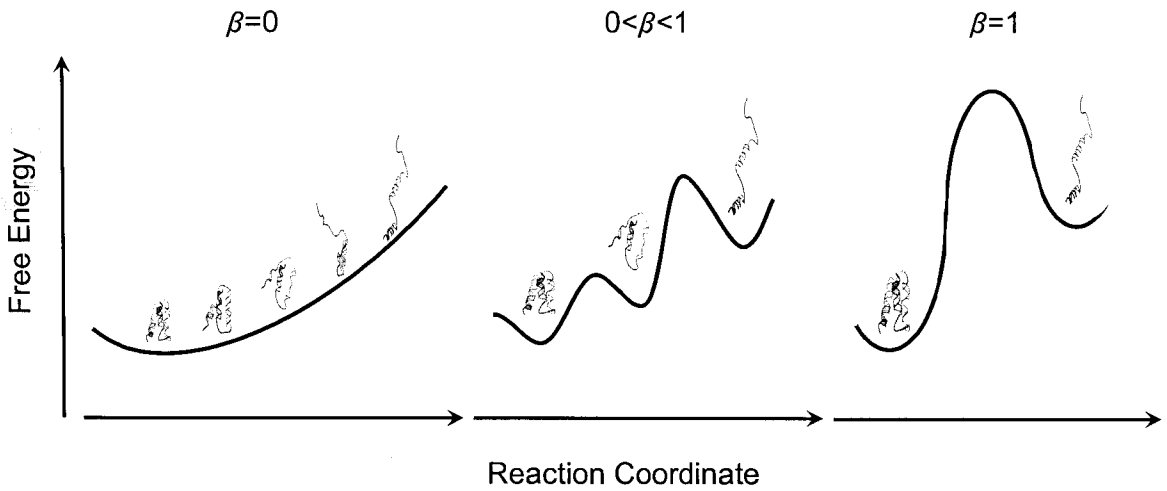

Figure 1.3 Schematic description of free energy landscape profile along with biomolecule folding and unfolding reaction coordinates
Different energy landscapes represent various degree of activation enthalpy, therefore different reaction mechanism. The values of stretched exponential parameter, β, are interpreted with corresponding ruggedness of the energy surface.

for most biomolecules. In this dissertation, we report the intermediate values of β (~0.6) for our five base-paired hairpin systems, consistent with dynamic heterogeneity in the DNA hairpin-forming mechanism.

As described previously, the rate-determining step that defines the folding times for short DNA or RNA hairpins in solution is not completely understood. Various models have been proposed to explain the kinetics of DNA hairpin formation, starting from basic polymer chain theory for describing how two ends of a polymer come into contact.[41-43, 90-92] In the case of hairpin formation in single stranded oligo-nucleotides, if intrachain interactions are ignored and the chain is treated like an ideal semiflexible polymer, end to end contact times expected from purely entropic conformational search is only tens of nanoseconds. Yet, the observed reaction times for DNA hairpin formation occur on the microsecond to tens of microsecond time scale. Therefore, more complicate mechanisms including intrachain interaction have been considered. The so-called "zipper model" of DNA hairpin folding has recently been proposed,[57, 93] which accounts for the nucleation and zipping steps in the folding process. Each nucleotide in the stem exists in one of two possible states, base-paired or open, and all of the base pairs occur contiguously in a single region. Initiation of nucleation can occur at any point along the stem where one or more base-pairs form to stabilize the looped conformations. The formation of a nucleus with one or more base pairs results in an ensemble of partially folded conformation. The addition of another base-pair leads to rapid zipping of the stem. In an extension of the original scheme, transient trapping of the single stranded chain gives rise to non-native loops with mismatched stems, which do not lead to the complete zipping of the stem. Thus, DNA hairpin folding can be described as a slower process

(micro-second time scale) than predicted from simple polymer chain dynamics. Several studies also proposed that the transient trapping in misfolded states can lead to non-Arrhenius behavior for the closing step.[66, 80, 94] However, force-extension measurements on polynucleotides have demonstrated that, under conditions of high ionic strength, intrachain interactions such as random base-pairing and intrastrand stacking interactions are non-negligible, and lead to significant deviations from the behavior expected for a semiflexible polymer.[69, 72] In comparison, the description of double-stranded DNA as a semiflexible polymer chain has been demonstrated in many experimental approaches consistently.[95-97] Thus, theoretical explanation of this slow dynamics of the single stranded DNA hairpin is still not well established. The main goal of this dissertation is to contribute some insight to the problems of hairpin dynamics. For this, our efforts were focused on the development of a new fluorescence based analysis strategy for analyzing biomolecule folding and unfolding kinetics.

1.2.5 Alternative Scenario in DNA Hairpin Folding Process

In the description of the DNA hairpin folding reaction, from the unfolded conformation to a closed hairpin, the hairpin formation would require several attempts in which non-native interactions have to be broken until the correct nucleation occurs.[66, 98] In many aspects, the reaction mechanism proposed in this dissertation is consistent with the model previously known as "zipper model with misfolded states". However, rather than the misfolded state as transient trap, our analysis provides strong evidence for the existence of stable intermediate state in the DNA hairpin folding process. These stable intermediates may correspond to partially folded conformations with mismatched

base pair(s) or collapsed states recently observed in molecular dynamics simulations of the hairpin folding process.[4, 5] We describe the hairpin formation process as rapid fluctuation between the unfolded, open chain conformation and the partially folded, intermediate state(s), followed by the slower folding reaction toward completely closed hairpin structure. This three-state mechanism supports our arguments that the micro-second time scale reaction rates reported in previous experiments, such as laser T-jump spectroscopy and FCS, do not corresponding to the entire folding reaction. Rather, these experiments probe the rapid fluctuations in the early stages of the process. One interesting observation regarding this point of view can be found in the direct measurement of the zipping and unzipping rates, carried out using a sub-microsecond T-jump experiment.[51] Kinetic measurements carried out below the melting transition temperature on 14 to 18 base-pair poly(A):poly(U) dimers revealed two distinct processes, one occurring with a time constant of about 0.2 μs and a much slower relaxation occurring with a time constant of a few seconds. In separate experiments, the same authors compared the estimated reaction enthalpy for forming the first base-pair of a $A_6C_6U_6$ hairpin with the activation enthalpy obtained from kinetics measurements.[52] They argued that a stable nucleus that leads to zipping occurs only after the formation of the fourth A-U base pair. Thus, more base pairs are required in first zipping step than suggested by the equilibrium zipper model. This lends support to the *three-state mechanism* proposed in this dissertation. Additionally, in early thermodynamic studies on λ-DNA cyclization, the authors explained the much slower cyclization times by arguing that the rate-determining step in the joining of the two ends is not the diffusion limited time for contact formation, but the very slow chemical step of base-pair formation.[45]

However, these ideas have not been supported widely in the study of hairpin folding dynamics. We envision that our observations can explain the experimental reports of slower reaction rates for the hairpin formation and resolve the discrepancies among various kinetics parameters that, in some cases, have existed in the literature for decades. We also hope the proposed *three-state mechanism* for the DNA hairpin folding process can provide a new insight toward development of a theoretical model describing this seemingly simple, but ever-complicate polymer chain behavior.

1.3 DISSERTATION SUMMARY

This dissertation is structured as follows. Chapter 2 introduces our new experimental approach, referred to as two-beam fluorescence fluctuation spectroscopy (2bFFS) for the dynamics study of biomolecules. This method was developed from the concept of fluorescence correlation spectroscopy (FCS). The background theory and its experimental realization are described. We also point out the motivation of 2bFFS approach especially for the study of biomolecule conformational dynamics under thermodynamic equilibrium. Chapter 3 addresses the kinetics measurements of DNA hairpin folding and unfolding reaction rates from the analysis of 2bFFS data. This analysis was based on the simultaneous measurements of auto- and cross-correlation of the fluorescence data. The equilibrium distribution in the reaction is discussed in Chapter 4. The combined approach of photon counting histogram with correlation analyses is described. Based on our observations, a *three-state mechanism* is proposed for the DNA hairpin formation. The mechanism provides an alternative interpretation for some

controversial experimental reports and plausible explanation for the fundamental questions still in debate. Summary and future prospective of our approach is presented in Chapter 5.

1.4 REFERENCE

1. Chong, L.T., et al., *Dimerization of the p53 Oligomerization Domain: Identification of a Folding Nucleus by Molecular Dynamics Simulations.* J. Mol. Biol., 2005. **345**: p. 869-878.
2. Snow, C.D., H. Ngyen, and V.S. Pande, *Absolute Comparison of Simulated and Experimental Protein Folding Dynamics.* Nature, 2002. **420**: p. 102-106.
3. Zagrovic, B., et al., *Unusual compactness of Polyproline Type ll Structure.* Proc. Natl. Acad. Sci. USA, 2005. **102**: p. 11698-11703.
4. Sorin, E.J., et al., *RNA Simulations: Probing Hairpin Unfolding and the Dynamic of a GNRA Tetraloop.* J. Mol. Biol., 2002. **317**: p. 493.
5. Sorin, E.J., et al., *Insights into Nucleic Acid Conformational Dynamics from Massively Parallel Stochastic Simulations.* Biophys. J., 2003. **85**: p. 790-803.
6. Froclich-Ammon, S.J., K.C. Galc, and N. Oshcroff, *Site-specific Cleavage of a DNA Hairpin by Topoisomerase II. DNA Secondary Structure as a Determinant of Enzyme Recognition/Cleavage.* J. Biol. Chem., 1944. **269**: p. 7719-7725.
7. Trinh, T.Q. and S.R. R., *The Influence of Primary and Secondary DNA Structure in Deletion and Duplication between Direct Repeats in Escherichia coli.* Genetics, 1993. **134**: p. 409-422.
8. Wartell, R.M. and A.S. Benight, *Thermal Denaturation of DNA Molecules: a Comparison of Theory with Experiment.* Phys. Rep., 1985. **126**: p. 67.
9. Lohman, T.M. and K.P. Bjorason, *Mechanisms of Helicase-Catalyzed DNA Unwiding.* Annu. Rev. Biochem., 1996. **65**: p. 169.
10. Bianco, P.R., et al., *Processive Translocation and DNA Unwinding by Individual RecBCD Enzyme Molecules.* Nature, 2001. **409**: p. 374.
11. Lilley, D.M., *Hairpin-Loop Formation by Inverted Repeats in Supercoiled DNA is a Local and Transmissible Property.* Nucleic Acids Res., 1981. **9**: p. 1271.
12. Romer, R., et al., *Escherichia Coli Single-Stranded DNA Binding Protein is Mobile on DNA: 1H NMR Study of its Interaction with Oligo- and Polynucleotides.* Biochemistry, 1984. **23**: p. 6132.
13. Roth, D.B., et al., *V(D)J Recombination: Broken DNA Molecules with Covalently Sealed (Hairpin) Coding Ends in Scid Mouse Thymocytes.* Cell, 1992. **70**: p. 983.
14. Dai, X., et al., *Supercoil-Induced Extrusion of a Regulatory DNA Hairpin.* Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 2174.
15. Dai, X., M. Kloster, and L.B. Rothman-Denes, *Sequence Dependent Extrusion of a Small DNA Hairpin at the N4 Virion RNA Polymerase Promoters.* J. Mol. Biol., 1998. **283**: p. 43.

16. Crews, S., et al., *Nucleotide Sequence of a Region of Human Mitochondrial DNA Containing the Precisely Identified Origin of Replication.* Proc. Natl. Acad. Sci. USA, 1979. **24**: p. 2174.

17. Wilson, K.S. and P.H. von Hippel, *Transcription Termination at Intrinsic Terminators: the Role of the RNA Hairpin.* Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 8793.

18. Glucksmann-Kuis, M.A., et al., *E. coli SSB Activates N4 Virion RNA Polymerase Promoters by Stabilizing a DNA Hairpin Required for Promoter Recognition.* Cell, 1996. **84**: p. 147.

19. Caskey, C.T., et al., *Triplet Repeat Mutations in Human Disease.* Science, 1992. **256**: p. 784.

20. Gacy, A.M., et al., *Trinucleotide Repeats that Expand in Human Disease Form Hairpin Structures in Vitro.* Cell, 1995. **81**: p. 533.

21. Tang, J., J. Temsamani, and S. Agrawal, *Self-stabilized Antisense Oigodeoxynucleotide Phosphorothioates: Properties and Anti-HIV Activity.* Nucleic Acids Res., 1993(21): p. 2729-2735.

22. Vesnaver, G. and K.J. Breslauer, *The Contribution of DNA Single-Stranded Order to the Thermodynamics of Duplex Formation.* Biopolymers, 1991. **50**: p. 425.

23. Gregorian, R.S. and D.M. Crothers, *Determinants of RNA Hairpin Loop-Loop Complex Stability.* J. Mol. Biol., 1995. **73**: p. 497.

24. Armitage, B., et al., *Hairpin-Forming Peptide Nucleic Acid Oligomers.* Biochemistry, 1998. **37**: p. 9417.

25. Bonnet, G., et al., *Thermodynamic Basis of the Enhanced Specificity of Structured DNA Probes.* Proc. Natl. Acad. Sci. USA, 1999. **96**(11): p. 6171-6176.

26. Tyagi, S. and F.R. Kramer, *Molecular Beacons: Probes that Flouresce upon Hybridization.* Nature Biotechnology, 1996. **14**: p. 303-308.

27. Zarrinkar, P.P. and J.R. Willianmson, *Kinetic Intermediates in RNA Folding.* Science, 1994. **265**: p. 918.

28. Zarrinkar, P.P., J. Wang, and J.R. Willianmson, *Slow Folding Kinetics of RNase P RNA.* RNA, 1996. **2**: p. 564.

29. Pan, T., D. Thirumalai, and S.A. Woodson, *Folding of RNA Involves Parallel Pathways.* J. Mol. Biol., 1997. **273**: p. 7.

30. Pan, T. and T.R. Sosnick, *Intermediates and Kinetic Traps in the Folding of a Large Ribozyme Revealed by Circular Dichroism and UV Absorbance Spectroscopies and Catalytic Acitvity.* Nat. Struct. Biol., 1997. **4**: p. 931.

31. Sclavi, B., et al., *RNA Folding at Millsecond Intervals by Synchrotron Hydroxyl Radical Footprinting.* Science, 1998. **279**: p. 1940.

32. Zhuang, X.W., et al., *A Single Molecule study of RNA catalysis and Folding.* Science, 2000. **288**: p. 2048.

33. Baldwin, R.L., *Why is protein folding so fast?* Proc. Natl. Acad. Sci. USA, 1996. **93**: p. 2627.

34. Levinthal, C. *How to Fold Graciously?* in *Mossbauer Spectroscopy in Biological Systems*. 1969. Allerton House, Monticello, Illinois: Univ. of Illinois Bulletin, Urbana, IL.

35. Sali, A., E. Shakhnovich, and M. Karplus, *How does a protein fold?* Nature, 1994. **369**: p. 248.

36. Sali, A., E. Shakhnovich, and M. Karplus, *Kinetics of protein folding. A lattice model study of the requirements for folding to the native state.* J. Mol. Biol., 1994. **235**: p. 1614.

37. Wolynes, P.G., J.N. Onuchic, and D. Thirumalai, *Navigating the folding routes.* Science, 1995. **267**: p. 1619.

38. Dill, K.A. and H.S. Chan, *From Levinthal to pathways to funnels.* Nat. Struct. Biol., 1997. **4**: p. 10.

39. Onuchic, J.N., Z. Luthey-Schulten, and P.G. Wolynes, *Theory of protein folding: the energy landscape perspective.* Annu. Rev. Phys. Chem., 1997. **48**: p. 545.

40. Zwanzig, R., A. Szabo, and B. Bagchi, *Levinthal's Paradox.* Proc. Natl. Acad. Sci. USA, 1992. **89**: p. 20.

41. Szabo, A., K. Schulten, and Z. Schulten, *First Passage Time Approach to Diffusion Controlled Reactions.* J. Chem. Phys., 1980. **72**: p. 4350.

42. Podtelezhnikov, A. and A. Vologodskii, *Simulation of Polymer Cyclization by Brownian Dynamics.* Macromolecules, 1997. **30**: p. 6668.

43. Friedman, B. and B. O'Shaughnessy, *Theory of Polymer Cyclization.* Phys. Rev. A., 1989. **40**: p. 5950.

44. Wang, J.C. and N. Davidson, *On the Probability of Ring Closure of Lamda DNA.* J. Mol. Biol., 1966. **19**: p. 496.

45. Wang, J.C. and N. Davidson, *Thermodynamic and Kinetic Studies on the Interconversion between the Linear and Circular Forms of Phage Lambda DNA.* J. Mol. Biol., 1966. **15**: p. 111.

46. Cohen, R.J. and D.M. Crothers, *Rate of Unwiding Small DNA.* J. Mol. Biol., 1971. **61**: p. 525.

47. Coutts, S.M., *Thermodynamics and Kinetics of G-C Base Pairing in the Isolated Extra Arm of Serine-specific Transfer RNA from Yeast.* Biochim, Biophys. Acta, 1971. **232**: p. 94.

48. Craig, M.E., D.M. Crothers, and P. Doty, *Relaxation Kinetics of Dimer Formation by Self complementary Oligonucleotides.* J. Mol. Biol., 1971. **62**: p. 383.

49. Porschike, D. and M. Eigen, *Co-operative Non-enzymic Base Recognition. 3. Kinetics of the Helix-coil Transition of the Oligoribouridylic-Okigoriboadenylic Acid System and of Oligoriboadenylic Acid Alone at Aidic pH.* J. Mol. Biol., 1971. **62**: p. 361.

50. Gralla, J. and D.M. Crothers, *Free Energy of Imperfect Nucleic Acid Helices. II. Small Hairpin Loops.* J. Mol. Biol., 1973. **73**: p. 497-511.

51. Porschike, D., *A Direct Measure of the Unzippering Rate of a Nucleic Acid Double Helix.* Biophys. Chem., 1974. **2**: p. 97.

52. Porschike, D., *Thermodynamics and Kinetics Parameters of an Oligonucleotide Hairpin Helix.* Biophys. Chem., 1974. **1**: p. 381.

53. Porschike, D., O.C. Uhlenbeck, and F.H. Martin, *Thermodynamics and Kinetics of the Helix-Coil Transition of the Oligomers Containing GC Base Pairs.* Biopolymers, 1977. **62**: p. 361.

54. Chu, Y.G. and I. Tinoco, *Temperature-Jump Kinetics of the dC-G-C-A-G-A-A-T-T-C-G-C-G Double Helix Containing a G. T Base-pair and the dG-C-A-G-A-A-T-T-C-G-C-G Double Helix Containing an Extra Adenine.* Biopolymers, 1983. **22**: p. 1235.

55. Xodo, L.E., et al., *The Duplex-Hairpin Conformational Transition of d(CGCGCGATCGCGCG) and d(CGCGCGTACGCGCG): A Thermodynamic and Kinetic Study.* J. Biomol. Struct. Dyn., 1988. **6**: p. 139.

56. Eigen, M. and L. de Maeyer, *Relaxation Methods*. 1963, New York: Interscience. 895.

57. Cantor, C.R. and P.R. Schimmel, *Biophysical Chemistry: the Behavior of Biological Macromolecules*. 1980, Freeman: New York.

58. Hoffman, G.W., *A Nanosecond Temperature-Jump Apparatus.* Rev. Sci. Instruments, 1971. **42**: p. 1643.

59. Bonnet, G., O. Krichevsky, and A. Libchaber, *Kinetics of Conformational Fluctuation in DNA Hairpin Loops.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 8602-8606.

60. Goddard, N.L., et al., *Sequence Dependent Rigidity of Single Stranded DNA.* Phys. Rev. Lett., 2000. **85**: p. 2400-2403.

61. Wallace, M.I., et al., *FRET Fluctuation Spectroscopy: Exploring the Conformational Dynamics of a DNA Hairpin Loop.* J. Phys. Chem. B, 2000. **104**: p. 11551-11555.

62. Wallace, M.I., et al., *Non-Arrhenius Kinetics for the Loop Closure of a DNA Hairpin.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 5584-5589.

63. Munoz, V., et al., *Folding Dynamics and Mechanism of beta-Hairpin Formation.* Nature, 1997. **390**: p. 196.

64. Dyer, R.B., et al., *Infrared Studies of Fast Events in Protein Folding.* Acc. Chem. Res., 1998. **31**: p. 709.

65. Shen, Y., S.V. kuznetsov, and A. Ansari, *Loop Dependence of the Dynamics of DNA Hairpins.* J. Phys. Chem. B, 2001. **105**: p. 12202-12211.

66. Ansari, A., S.V. kuznetsov, and Y. Shen, *Configurational Diffusion Down a Folding Funnel Describes the Dynamics of DNA Hairpins.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 7771-7776.

67. Kuznetsov, S.V., et al., *A Semiflexible Polymer Model Applied to Loop Formation in DNA Hairpins.* Biophys. J., 2001. **81**: p. 2864-2875.

68. Grunwell, J.R., et al., *Monitoring the Conformational Fluctuations of DNA Hairpins Using Single-Pair Fluorescence Resonance Energy Transfer.* J. Am.Chem. Soc., 2001. **123**: p. 4295-4303.

69. Liphardt, J., et al., *Reversible Unfolding of Single RNA Molecules b Mechanical Force.* Science, 2001. **292**: p. 733.

70. Deniz, A.A., et al., *Single-pair Fluorescence Resonance Energy Transfer on Freely Diffusing Molecules: Observation of Forster Distance Dependence and Subpopulations.* Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 3670.

71. Uhlenbeck, O.C., et al., *Stability of RNA Hairpin Loops: A6-Cm-U6.* J. Mol. Biol., 1973. **73**: p. 483.

72. Cocco, S., J.F. Marko, and R. Monasson, *Slow Nucleic Acid Unzipping Kinetics from Sequence-Defined Barriers.* Eur. Phys. J. E., 2003. **10**: p. 153.

73. Cocco, S., et al., *Force-Extension Behavior of Folding Polymers.* Eur. Phys. J. E., 2003. **10**: p. 249.

74. Turner, D.H. and N. Sugimoto, *RNA Structure Prediction.* Annu. Rev. Biophys. Biophys. Chem., 1988. **17**: p. 167.

75. Zuker, M., *On Finding All Suboptimal Foldings of an RNA Molecule.* Science, 1989. **244**: p. 48.

76. Jaeger, J.A., D.H. Turner, and M. Zuker, *Predicting Optimal and Subptimal Secondary Structure for RNA.* Methods Enzymol., 1990. **183**: p. 281.

77. Jaeger, J.A., J. SantaLucia, and I. Tinoco, *Determination of RNA Structure and Thermodynamics.* Annu. Rev. Biochem., 1993. **62**: p. 255.

78. Mathews, D.H., et al., *Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure.* J. Mol. Biol., 1999. **288**: p. 911.

79. Flamm, C., et al., *RNA Folding at Elementary Step Resolution.* RNA, 2000. **6**: p. 325.

80. Zhang, W. and S.J. Chen, *RNA Hairpin Folding Kinetics.* Proc. Natl. Acad. Sci. USA, 2002. **99**: p. 1931.

81. Zichi, D.A., *Molecular Dynamics of RNA with the OPLS force field. Aqueous Simulations of a Hairpin Containing a Tetranucleotide Loop.* J. Am.Chem. Soc., 1995. **117**: p. 2957.

82. Miller, J.L. and P.A. Kollman, *Observation of an A-DNA to B-DNA Transition in a Non-helical nucleic Acid hairpin Molecule using Molecular Dynamics.* Biophys. J., 1997. **73**: p. 2702.

83. Miller, J.L. and P.A. Kollman, *Theoretical Studies of an Exceptionally Stable RNA Tetraloop: Observation of Convergence from an Incorrect NMR Structure to the Correct One using Unrestrained Molecular dynamics.* J. Mol. Biol., 1997. **270**: p. 436.

84. Young, M.A., G. Ravishanker, and D.L. Berveridge, *A 5-nonosecond Molecular Dynamics Trajectory for B-DNA: Analysis of Structure, Motions, and Solvation.* Biophys. J., 1997. **73**: p. 2313.

85. Srinivasan, J., et al., *Continuum Solvent Studies of the Stability of RNA Hairpin Loops and Helices.* J. Biomol. Struct. Dyn., 1998. **16**: p. 671.

86. Williams, D.J. and K.B. Hall, *Unrestrained Stochastic Dynamics Simulations of the UUCG Tetraloop using an Implicit Solvation Model.* Biophys. J., 1999. **76**: p. 3192.

87. Sarzynska, J., T. Kulinski, and L. Nilsson, *Conformational Dynamics of a 5S rRNA Hairpin Domain Containing Loop D and a Single Nucleotide Bulge.* Biophys. J., 2000. **79**: p. 1213.

88. Zacharias, M., *Conformational Analysis of DNA-Trinucleotide-Hairpin-Loop Structure using a Continuum Solvent Model.* Biophys. J., 2001. **80**: p. 2350.

89. Sorin, E.J., Y.M. Rhee, and V.S. Pande, *Does Water Play a Structural Role in the Folding of Small Nucleic Acids?* Biophys. J., 2005. **88**: p. 2516-2524.

90. Wilemski, G. and M. Fixman, *Diffusion-controlled Intrachain Reactions of Polymers. 1. Theory.* J. Chem. Phys., 1974. **60**: p. 866.

91. Doi, M., *Diffusion-controlled Reaction of Polymers.* Chem. Phys., 1975. **9**: p. 455.

92. Guo, Z. and D. Thirumalai, *Kinetics of Protein Folding: Nucleation Mechanism, Time Scales, and Pathways.* Biopolymers, 1995. **36**: p. 83.

93. Poland, D. and H.A. Scheraga, *Theory of Helix-coil Transitions in Biopolymers; Statistical Mechanical Theory of Order-disorder Transitions in Biological Macromolecules*. Vol. xvii. 1970, New York: Academic Press.

94. Socci, N.D., J.N. Onuchic, and P.G. Wolynes, *Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels.* J. Chem. Phys., 1996. **104**: p. 5860.

95. Bustamante, C., et al., *Entropic Elasticity of Lamda-Phage DNA.* Science, 1994. **265**: p. 1599.

96. Marko, J.F. and E.D. Siggia, *Fluctuations and Supercoiling of DNA.* Science, 1994. **265**: p. 506.

97. Hagerman, P.J., *Flexibility of DNA.* Annu. Rev. Biophys. Biophys. Chem., 1988. **17**: p. 265.

98. Ansari, A. and S.V. Kuznetsov, *Is Hairpin Formation in Single-Stranded Polynucleotide Diffusion-Controlled?* J. Phys. Chem. B, 2005. **109**: p. 12982-12989.

Chapter 2

Fluorescence Fluctuation Spectroscopy

2.1 BACKGROUND AND THEORY

2.1.1 Fluorescence Correlation Spectroscopy

2.1.1.1 Basic Concepts

In present work, we use the ultra-sensitive fluorescence based chemical analysis techniques known collectively as fluorescence fluctuation spectroscopy (FFS).[1] The most prominent example of FFS is fluorescence correlation spectroscopy (FCS), which was introduced as a way to non-invasively monitor the dynamics of chemically reacting system.[2-5] Beginning in 1972, Elson, Magde, and Webb published a series of papers describing a fluorescence-based analogue of Dynamic Light Scattering (DSL)[6, 7], a

commonly used technique for measuring solute diffusivities. The authors realized that using fluorescence measurements to analyze the particle motion imparts several key advantages over conventional DLS. Fluorescence is more chemically selective than light scattering and allows greater flexibility in studying the motion of specific analytes. Also, fluorescence can be detected from molecules that are much too small to be detected by DLS, which allows one to characterize the motion of small molecules as well as large macromolecules. DSL is limited to measuring the diffusivities of macromolecules greater than ~1nm in size. Finally, fluorescence analysis opens the way to characterizing chemically reacting systems. Chemical reactions generally do not create a large enough index of refraction change to be studied by DLS. However, spontaneous chemical reactions can create fluctuations in the molecular diffusion properties and/or other fluorescence characteristics that could conceivably be analyzed via fluorescence detection with much greater sensitivity. FCS was accomplished by focusing an excitation laser beam into the analyte solution and then monitoring the fluorescence generated from the laser beam's focal region.(See Figure 2.1) Diffusion of fluorophors into and out of the focal volume altered the local concentration of the fluorophors, giving rise to spontaneous fluorescence intensity fluctuations that could be analyzed using a statistical analysis technique referred to as *autocorrelation* analysis to characterize the time dependence of these fluctuations. In general, autocorrelation analysis measures the time dependence of a given set of measurement values, which in the case of FCS refers to the

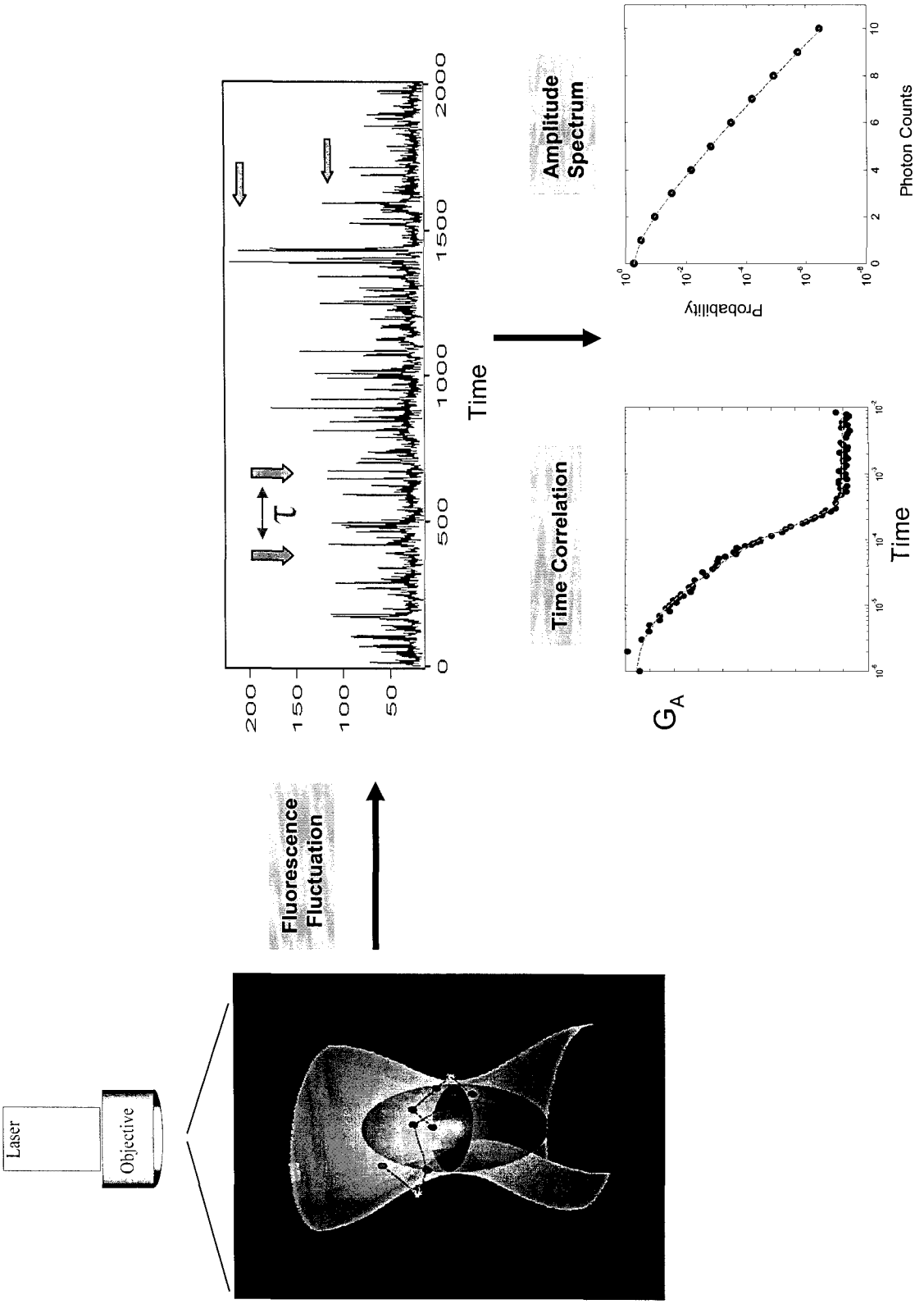

Figure 2.1 Schematic description for the fluorescence fluctuation spectroscopy concept (Fluorescence auto correlation and photon counting histogram analyses are shown.) The laser beam is focused in the sample solution by confocal microscope objective. The size is less than 1 micrometers and the occupying volume is ~1 femto liter scale. The oval shape in this cartoon shows the laser focus. And a fluorescent molecule moving around it by diffusion is shown. The pathway of the molecule movement was generated by Monte Carlo simulation. Whenever the molecule is inside the laser focus which is called detection volume, gives off fluorescent light. Because, the molecules in solution moves in and out of the focus continuously, the fluctuation in fluorescence intensity is observed and recorded. The fluctuations at time t and time $t + \tau$ are now correlated using Equation (2.25) and produce the auto-correlation curve. The curve shows higher correlation at shorter time and decrease as time goes on. In other approach, the amplitude spectrum of fluorescence intensity is histogrammed. Fitting the photon counting histogram by a theoretical model such as Equation (2.32) provides information on the molecular brightness of the sample molecule as well as sample concentration.

photocurrent (or photoelectron counts) from a photo-detector as function of time. FCS monitors and compares the fluorescence intensity, $\langle I(t)\rangle$, measured at time t, with the intensity measured at a later time $t + \tau$, averaged over all values of t. this information is contained in an auto-correlation function, $G(\tau)$.

$$G(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \frac{I(t)I(t+\tau)}{\langle I(t)\rangle^2}dt = \frac{\langle(\langle I\rangle+\delta I(t))(\langle I\rangle+\delta I(t+\tau))\rangle}{\langle I\rangle^2\langle I\rangle^2}$$
$$G(\tau)-1 = \frac{\langle\delta I(t+\tau)\delta I(t)\rangle}{\langle I\rangle^2} \qquad (2.1)$$

where, T is the total measurement time and τ is the lagtime. The autocorrelation function compares the fluorescence light intensity measured at time t, with the intensity measured at a later time, $t + \tau$. when τ is small compared to t, there is little time for the particle diffuse, giving rise to comparatively large values of $G(\tau)$. As τ increases relative to t, $G(\tau)$ begins to decay, finally reaching a constant value at infinitely long lagtimes. The fluorescence intensity, $I(t)$ is given by,

$$I(t) = \kappa\int_V I_{ex}(\vec{r})\cdot S(\vec{r})\cdot\sigma\cdot\Phi(\vec{r})\cdot C(\vec{r},t)\ dV \qquad (2.2)$$

where, κ is overall detection efficiency; I_{ex} is excitation energy profile with the maximum amplitude I_0; $S(r)$ is collection efficiency; σ is molecular absorption cross-section; Φ is quantum yield; and $C(r,t)$ is the local particle concentration at time t. In the our FCS optical set-up, the Equation (2.2) can be simplified by,

$$I(t) = \int_V W(\vec{r})\cdot\varepsilon\cdot C(\vec{r},t)\ dt \qquad (2.3)$$

Here, ϵ is the photon count rate per detected molecule per second, or called molecular brightness given by $\epsilon = I_0 \cdot \kappa \cdot \sigma \cdot \Phi$. $W(r)$ is the spatial distribution of the laser focus given by the convolution of $I_{ex}(r) / I_0$ and $S(r)$, and can be approximated by a 3D-Gaussian function $W_G(r)$:

$$W_G(\vec{r}) = \exp\left(-2\frac{x^2+y^2}{\omega_0^{\,2}} - 2\frac{z^2}{z_0^{\,2}}\right) \tag{2.4}$$

with the e^{-2} focal radius in the radial dimension, ω_0, of the focal volume and the axial e^{-2} radius, z_0. In that case, the focal volume, V_G, is defined by

$$V_G = \frac{\left(\int W(\vec{r})dV\right)^2}{\int W(\vec{r})^2 dV} = \pi^{\frac{3}{2}}\omega_0^{\,2} z_0 \tag{2.5}$$

Using these terms, the Equation (2.1) can be expressed as

$$G(\tau) - 1 = \frac{\iint W(\vec{r})W(\vec{r}\,')\left\langle \delta[\varepsilon \cdot C(\vec{r},t)]\delta[\varepsilon \cdot C(\vec{r}\,',t+\tau)]\right\rangle dVdV'}{\left(\int W(\vec{r})\left\langle \delta[\varepsilon \cdot C(\vec{r},t)]\right\rangle dV\right)^2} \tag{2.6}$$

If ϵ is constant for a given system ($\delta\epsilon = 0$), so that the changes in local concentration only come from free diffusion of molecules in the solution, the concentration fluctuation $\delta C(r,t)$ can be derived by solving following diffusion equation.

$$\frac{\partial \delta C_j(\vec{r},t)}{\partial t} = D_j \nabla^2 \delta C_j(\vec{r},t) \tag{2.7}$$

where, D_j is the diffusion coefficient of species j. Equation (2.7) can be evaluated using the Fourier transform method. A derivation of corresponding correlation function, which assumes a single component solution and only considers diffusion within the laser beam, yields the equation:

$$G(\tau)-1=g_{Diffusion}(\tau)=\frac{1}{N}\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2}=\frac{1}{N}g_D(\tau) \tag{2.8}$$

where, κ_0 is the ratio of the radial and axial e^{-2} radii ($\kappa_0 = \omega_0/z_0$). We defined the diffusion term as $g_D(\tau)$:

$$g_D(\tau)=\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2} \tag{2.9}$$

Note that according to Equation (2.8), $g_{Diffusion}(\tau)$ decays with τ from a maximum at $\tau = 0$. The amplitude of $g_{Diffusion}(\tau)$ depends on the average number of molecules, N, occupying the observation volume (i.e., the occupancy), and its width depends on τ_D, the average time it takes for a molecule to diffuse through the observation volume. The rate of decay of decay of $G(\tau)$ vs. τ is characteristic of the diffusion rate of the particles, which , in turn, yields the diffusion coefficient, D. The relationship of τ_D to the molecular diffusion coefficient is given by:

$$\tau_D=\frac{\omega_0{}^2}{4D} \tag{2.10}$$

Knowledge of the diffusion coefficient gives rise to the effective hydrodynamic radius, R_H, of the particle, according to the Stokes-Einstein equation:

$$D=\frac{k_BT}{6\pi\eta R_H} \tag{2.11}$$

where k_B is the Boltzmann constant, T is the temperature, and η is the solvent viscosity. Hence, measuring the decay rate of $G(\tau)$ vs. τ gives the diffusion coefficient of the analyte, and from thence, its molecular size. The case that will interest us here is where a

small ligand molecule binds to a much larger receptor, thus reducing the diffusion coefficient of the ligand. If the ligand-receptor complex is stable on the time scale of its transit time through the detection region, then the auto-correlation function represents a linear combination of two single-component auto-correlation functions, corresponding to the free and bound ligand:

$$G(\tau)-1=\frac{\sum_{i=1}^{2} N_i \varepsilon_i^{\,2} g_{D,i}(\tau)}{< I >^2} \quad . \tag{2.12}$$

Here, i represents the free or the bound ligand, therefore, ϵ_i is the molecular brightness of species i (the average emission rate of a single molecule of species i), and $\langle I \rangle$ is the time-averaged total signal intensity due to the total fluorescence plus the background signal. If the diffusion rates of the different species are sufficiently resolved, then the auto-correlation function can be analyzed to determine the relative concentrations of bound and unbound ligand, providing a measure of the equilibrium binding constant.

Another important aspect of FCS is its ability to characterize chemically reacting systems. Spontaneous fluctuations in the local concentration of an analyte can be caused by chemical reactions occurring under equilibrium conditions, as well as by molecular diffusion.(Figure 2.2) By characterizing these fluctuations, FCS can measure equilibrium constants and reaction rate parameters without the need to perturb the system equilibrium. Elson, Magde, and Webb derived auto-correlation functions for a number of different scenarios in which the local concentration of the analyte was fluctuating due to some type

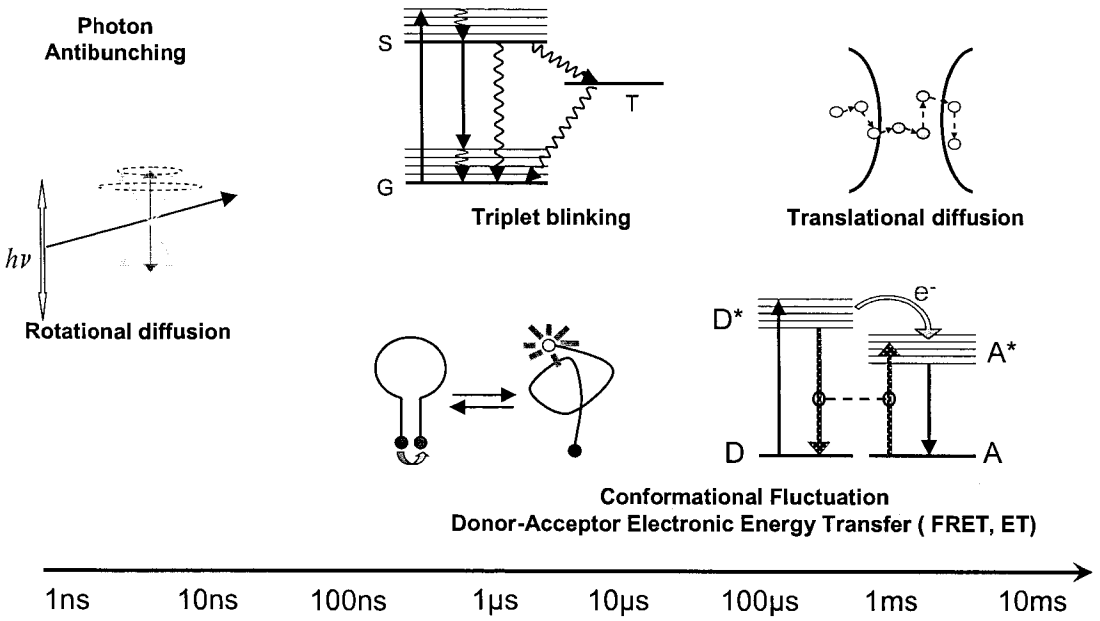

Figure 2.2 Schematic description for the examples of Dynamic processes which can be accessed by fluorescence correlation spectroscopy (FCS) and their corresponding time scales

The time scale in the conformational fluctuation of biomolecules is often overlapped with other process such as triplet blinking or translational diffusion of the molecules. Therefore, analysis of the coupled dynamics is difficult from one data set of auto correlation curve alone. In order to circumvent this problem, we developed a new approach presented in this dissertation.

of spontaneous chemical reaction. If ϵ is changes due to chemical reaction, $\delta C(r,t)$ is described by following differential equation.

$$\frac{\partial \delta C_j(\vec{r},t)}{\partial t} = \sum_{k=1}^{m} T_{jk} \delta C_j(\vec{r},t) \tag{2.13}$$

where, m is the total number of species participating in the chemical reaction and T_{jk} represents the matrix of the kinetic rate coefficients. For a simple reaction between two components, the corresponding correlation function can be expressed as following,

$$g_{Reaction}(\tau) = \frac{1-P+Pe^{(-\tau/\tau_P)}}{1-P} = 1+P_{eq}e^{(-\tau/\tau_P)} \tag{2.14}$$

where, P is an amplitude term which is a function of the relative fraction of each species and their molecular brightnesses and τ_P is the relaxation time of the chemical process. Examples of these reactions are triplet blinking and biomolecule conformational change which is the main subject of this dissertation. When both diffusion and reaction process are coupled and involved in the fluctuation, $\delta C(r,t)$ is given by the combined form of Equation (2.7) and (2.13). However, in most case, each process covers different time scales, so that it is possible to separate the dynamic contributions.

$$G(\tau)-1 = g_{Diffusion}(\tau)\cdot g_{Reaction}(\tau) \tag{2.15}$$

FCS experiments can be also performed under the condition of flowing solution such as electrophoretic flow or pressure-driven flow. The correlation function is derived by solving the governing hydrodynamic equations with appropriate boundary conditions.

$$\frac{\partial \delta C_j(\vec{r},t)}{\partial t} = -V(\vec{r}) \cdot \nabla \delta C_j(\vec{r},t) \tag{2.16}$$

where, $V(r)$ is the flow rate at the detection volume. In most situation, particle diffusion is coupled to the flow, so that $\delta C(r,t)$ is given by the combined form of Equation (2.7) and (2.16). The equations lead to a flow term, $g_{Fa}(\tau)$, in the autocorrelation function:

$$g_{Fa}(\tau) = \exp\left[-\frac{\left(V(\vec{r})\cdot\tau\right)^2}{\omega_0^{\,2}\left(1+\tau/\tau_D\right)}\right] = \exp\left[-\left(\frac{\tau}{\tau_{Fa}}\right)^2\left(\frac{1}{\left(1+\tau/\tau_D\right)}\right)\right] \tag{2.17}$$

where, τ_{Fa} is the transit time of molecules across the detection volume with radius of ω_0. Flowing the sample through the excitation beam suppresses photobleaching problem especially under high laser power and provides fresh fluorophors into the excitation region faster than can be accomplished by diffusion only. When combined with electrophoretic flow, this technique can be used to characterize sample properties associated with the molecular charge.

2.1.1.2 Single Molecule Confocal Spectroscopy

Problems in the early development of FCS (i.e. high background noise which is mainly the scattered light from solvent, and poor efficiency in light collection and detection), have been overcome in recent years via the combination of FCS with single molecule confocal microscopy. Rigler and co-workers[8-12] and Zare and co-workers[13] demonstrated that confocal microscopy could be used to directly detect the fluorescence emitted by individual molecules as they diffused through the microscopic

focal volume of the confocal microscope. This was an important extension of the original single molecule detection studies reported by Keller and co-workers.[14] The correlation function could be determined based on a relatively small number of single molecule fluorescence signals, detected over a period of a few seconds.(Figure 2.3) FCS is normally done by focusing an excitation laser beam to its diffraction limit using a high numerical aperture (NA) microscope objective, positioning the focal region into the analyte solution, and monitoring the fluorescence generated from within the focal volume over time. The same objective lens also serves to collect fluorescence from the sample, an arrangement referred to as epi-illumination a small pinhole, positioned at the image plane of he objective (the position where the image comes into focus behind the rear aperture of the objective), acts as a spatial filter to block fluorescence generated outside the focal region from reaching the detector, thus ensuring that only the fluorescence generated within the focal region can be detected. The spatial distribution of the light intensity within the laser beam focus serves as the detection volume. The size of the detection volume can be estimated by assuming a cylindrically shaped focal volume with radius, ω_0, and height, $2z_0$, where z_0 is the axial radius of the focal volume. ω_0 and z_0 are related to the NA of the objective, the wavelength, λ, of the excitation light, and the index of the refraction, n, of the sample medium according to the equations:

$$\omega_0 = \frac{1.22\lambda}{2 \cdot NA} \qquad z_0 = \frac{2n\lambda}{(NA)^2} \tag{2.18}$$

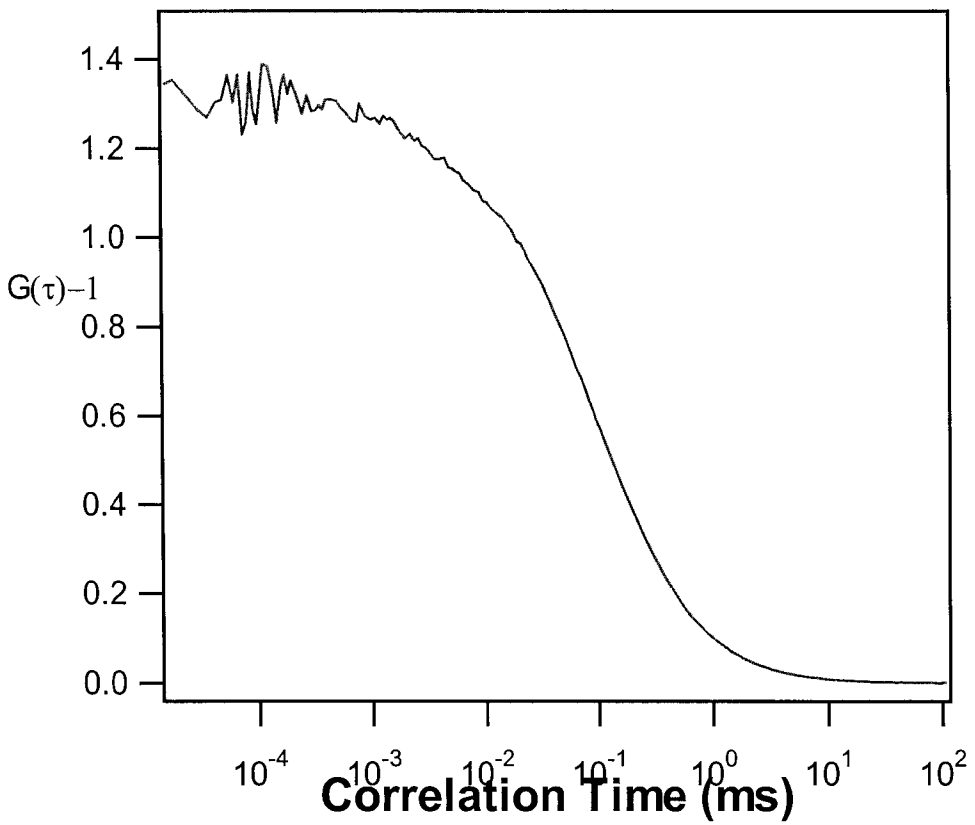

Figure 2.3 Auto-correlation data from FCS measurement of ~1nM Rhodamine 6G solution

Molecular process with time scale from 10 nano-second to 10 milli-second can be accessed by the correlation analysis.

In an experiment that utilizes a 1.3 NA objective, a 514.5-nm laser beam as the excitation source, and an aqueous medium (n=1.33), the resulting detection volume is ~0.3 femtoliters (fl). This extremely small detection volume is important for several reasons. It suppresses the background noise caused by backscattering of the excitation laser beam through Raleigh and Raman scattering processes; it enables optical excitation of the fluorophors to their saturation point using a modest average laser power (< 1 mW); it ensures that the number of fluorophors being probed at any given time is small; and it allows samples with extremely small volumes (microliters or less) to be analyzed. Other aspects of confocal microscopy that are important for single molecule detection include the high collection efficiency of the objective lens (~25% for a 1.3 NA oil-immersion objective), the high transmission efficiency of the optical components in the wavelength range of interest, and an efficient single photon counting detector. Modern single photon counting avalanche photodiode modules are able to detect visible photons with 30 to 70% quantum efficiency. All in all, collection/detection efficiencies of 5 to 10% are attainable with modern confocal microscope setups. Considering that many fluorophors can emit up to 10^6 to 10^8 photons per second (prior to photobleaching) when pumped near their optical saturation point, this can lead to photodetection rates that exceed 10^5 photons per second per molecule, albeit over brief time periods.

Modern FFS takes advantage of the fact that dilute solutions (sub-nanomolar to sub-micromolar) of fluorophors exhibit large amplitude fluorescence intensity fluctuations

when probed by single molecule confocal microscopy. This allows the fluctuations to be characterized in a matter of seconds, rather than the tens of minutes to hours needed in the earlier days. Large amplitude fluctuations arise because the average number of fluorophors occupying the detection volume (i.e. the occupancy) is small compared to the deviation from the average at any given time. Random diffusion of fluorophors into and out of the detection volume ensures that the number of flurophors being probed is never the same from one moment to the next. Consider, for example, an analyte concentration of 1 nM. At this concentration, the average number of fluorophors within a 1-femtolitter detection volume is ~0.6 molecules. This means that, on average, the occupancy fluctuates between zero and one, corresponding to deviations from the mean occupancy of 0.6 and 0.4, respectively. If the microscope is properly configured for single molecule detection, then the fluorescence signal will be characterized by "quiet" periods, during which only background noise is observed, punctuated by brief, intense "bursts" of signal, due to the passage of a single molecule through the detection volume (See Figure 2.4A). The durations of the bursts are characteristic of the diffusion rate of the molecules, with average burst durations typically ranging from a few tens of microseconds to a few milliseconds, depending on the molecule's diffusion rate.

At fluorophor concentrations between ~10 and 100nM, the number of molecules occupying the detection volume, and hence, the fluorescence signal, varies about a certain mean value (See Figure 2.4 B). The fluorescence data collected under these

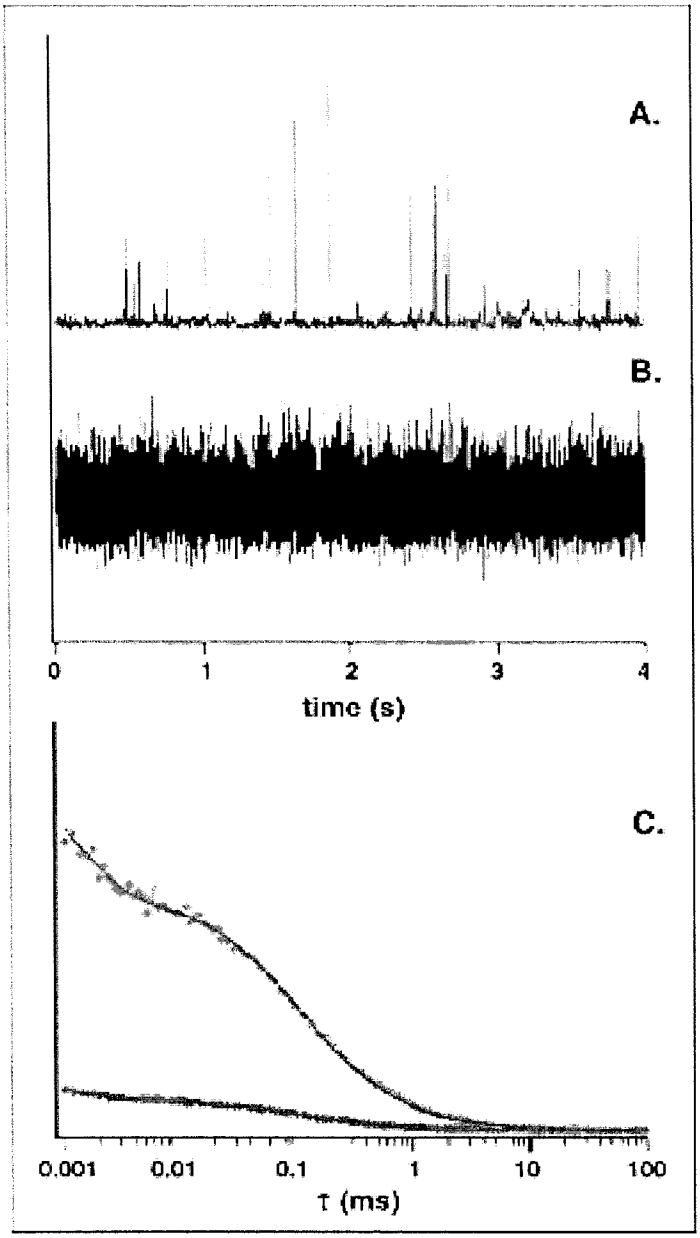

Figure 2.4 Time dependent fluorescence photo count data and auto-correlation functions obtained from static solutions of Rodamine 6G molecules being probed single molecule confocal fluorescence detection experiment. (a) Fluorescence data obtained from a dilute (sub-nM) solutions of fluorophors. Photon bursts from individual molecules are clearly resolved. (b) Fluorescence data obtained from a more concentrated (>10nM) solution of fluorophors. The solution is too concentrated for single molecule bursts to be clearly differentiated from the overall fluorescence. (c) Auto-correlation curves typical of a dilute solution (upper) and a more concentrated solution (lower). The solid diamonds are experimental data points, and the solid curves represent fits to a model equation.

conditions is still representative of individual molecule transits even though more than one molecule is being probed at a time. The amplitude of the autocorrelation function taken under these conditions will be reduced due to the inverse relationship with the occupancy number (Figure 2.4C). As the concentration is increased above ~100 nM, the deviation in the occupancy becomes small compared to the average fluorescence signal, and the detector starts to reach its saturation point. This requires lowering the laser power, thus reducing the molecular brightness of the fluorophors. These two effects place an upper limit on the fluorophor concentration in FCS analysis. At the lower end of the concentration scale, the lengthy time interval between detected molecules becomes a limiting factor, as does background radiation coming from Raman scattering by the solvent. In general, FFS is useful for analyte concentrations in the range of ~0.1 nM to ~100 nM. It is sometimes possible to attain lower detection limits by rapidly scanning the focal volume of the laser beam relative to the sample (or vice versa).[15] This enlarges the effective detection volume, and, hence, the average molecular occupancy, without introducing unwanted background radiation.

2.1.1.3 Fluorescence Cross Correlation Spectroscopy

Diffusional FCS only works for assays that involve a larger change in molecular size. The bound complex needs to be on the order of ~8 times more massive than the free ligand. Otherwise, the different species are difficult to distinguish based on their diffusion

times alone. Hence, there has been a strong motivation to develop alternative FFS strategies that are sensitive to other properties of the system besides diffusion. One of such strategies is a FCS based technique using the concept of cross-correlation. Two-color fluorescence cross-correlation spectroscopy (2cFCCS) and two-beam fluorescence cross-correlation spectroscopy (2bFCCS) are two modes for the approach. In 2cFCCS the detection volume is formed by spatially overlapping two excitation laser beams, each operating at a different wavelength.[16-22] Two different dye molecules that absorb light in different spectral regions can both be excited within the same detection volume. Fluorescence generated in the detection volume is split into two different detection channels, each sensitive to the emission spectrum of one of the dyes. The signals from the two detectors are then subjected to cross-correlation analysis. Instead of comparing signals from the same detector at two different times, as in autocorrelation analysis, the comparison is made between the signals from detector 1 at time t and detector 2 at time $t + \tau$. The cross-correlation function is then obtained by integrating over all values of t. Mathematically, this is expressed as:

$$G(\tau) = \lim_{T\to\infty} \int_0^T \frac{I_1(t)I_2(t+\tau)}{\langle I_1\rangle\langle I_2\rangle} dt = \frac{\langle I_1(t)I_2(t+\tau)\rangle}{\langle I_1\rangle\langle I_2\rangle} \tag{2.19}$$

where I_1 and I_2 are the fluorescence signals from detectors 1 and 2, respectively. The key aspect of 2cFCCS is that contributions to the cross-correlation function only occur when both fluorophors are simultaneously present in the detection volume. This means that

binding assays can be constructed in which each binding partner is labeled with a different fluorophor. The binding reaction creates a doubly labeled complex that can be detected via 2cFCCS, whereas the singly labeled unbound species make no contribution. By analogy, assays involving the decomposition of a doubly labeled molecule to form two singly labeled products can also be studied in this way. The assays do not depend on changes in molecular size, but only on the coincident detection of both fluorophors. Another advantage over diffusional FCS is that the amplitude of the correlation function, $Gc(0)$, with the autocorrelation function amplitudes from each detection volume, $G_1(0)$ and $G_2(0)$, determined from the same data set, one can directly measure the concentration of the doubly labeled complex using:

$$\overline{C} = \frac{G_C(0)}{V_{eff} G_1(0) G_2(0)} \tag{2.20}$$

where, C is the average concentration of the complex and V_{eff} is the confocal detection volume.

The 2bFCCS technique was devised by using spatially separate laser beams to determine the transit times of individual molecules between the two beams.[23-26] (See Figure 2.5) In this approach, the focal regions are offset from each other by only a few micrometers distance, R, which enables the fluorescence from both focal regions to be collected by the same high efficiency objective. As in single-beam FCS, 2bFCCS has been used in the past to characterize the transport properties of single

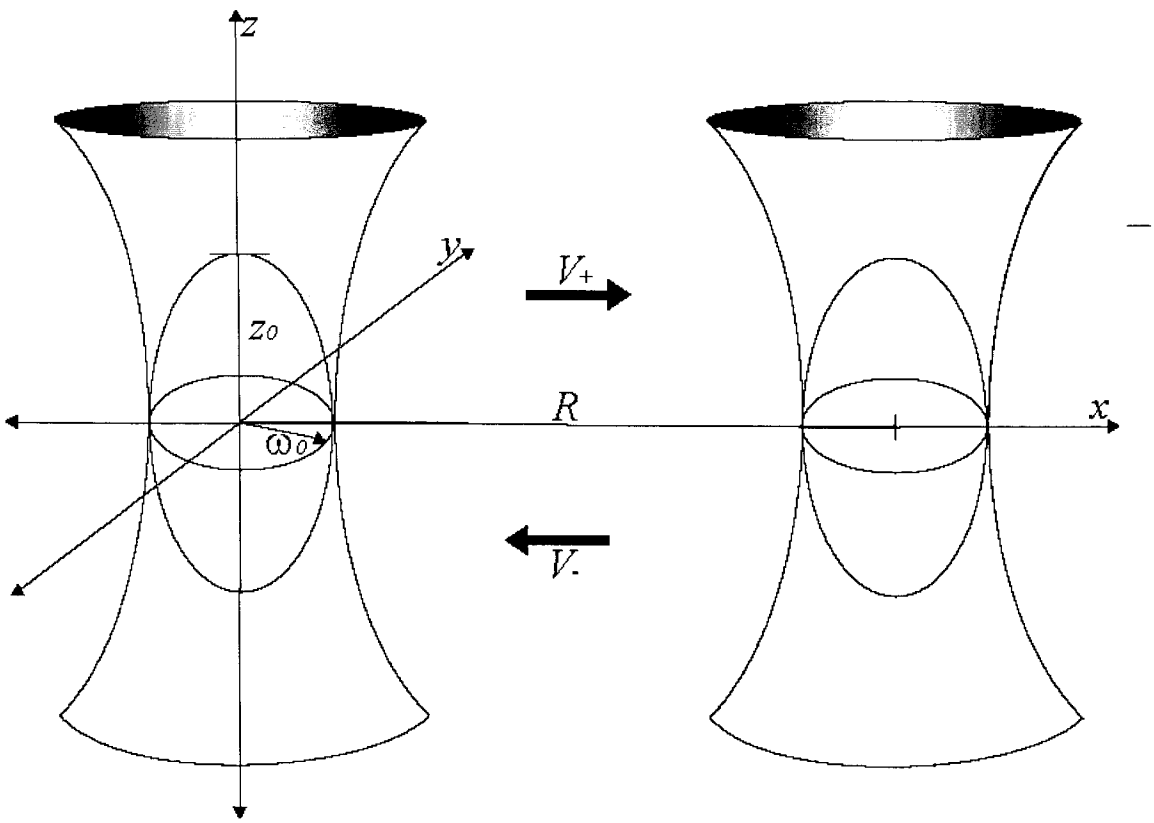

Figure 2.5 Schematic description for the geometry of detection volumes in the 2bFCCS experiment

Two identical laser beams are tightly focused in the solution and form Gaussian-like foci with lateral radius w_0 along coordinates of the molecule path, x. z_0 represents the longer axial radius. The two detection volumes are separated by R. In our experiments, w_0 was ~0.3 micrometer and R was ~5 micrometer.

molecules in microstructured flow channels. Previously demonstrated advantages of this technique relative to single-beam FCS include the ability to measure slower flow velocities at higher precision, and to resolve the different components of a multi-component mixture based on their molecular transit time and direction.

The cross correlation function can be derived as same way as autocorrelation case.:

$$G_C(\tau)-1=\frac{1}{N}\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\tau/\tau_D}\right)^{1/2}\exp\left(-\frac{\left(\vec{V}_\rho\cdot\tau-\vec{R}\right)^2}{\omega_0^{\,2}\left(1+\tau/\tau_D\right)}-\frac{V_z^{\,2}\tau^2}{z_0^{\,2}\left(1+\kappa_o^2\tau/\tau_D\right)}\right) \tag{2.21}$$

where, V_ρ is the x- and y- depending term of the flow velocity, which is separated from V_z, the z-depending term. τ_{Fc} is the average time a molecule needs to travel from one detection volume to the other and thus, defined as $\tau_{Fc}=R/V(\vec{r})$. For the case of directional flow which is parallel to the axis between the two laser beams, the Equation (2.21) can be reduced to:

$$G_C(\tau)-1=g_{Diffusion}\cdot\exp\left(-\frac{R^2\left(1-\tau/\tau_{Fc}\right)^2}{\omega_0^{\,2}\left(1+\tau/\tau_D\right)}\right) \quad . \tag{2.22}$$

Here, we defined the flow term as $g_{Fc}(\tau)$:

$$g_{Fc}(\tau)=\exp\left(-\frac{R^2\left(1-\tau/\tau_{Fc}\right)^2}{\omega_0^{\,2}\left(1+\tau/\tau_D\right)}\right) \tag{2.23}$$

This function rises to a maximum at $\tau=\tau_{\max}$ which is given by

$$\tau_{\max}=-\left(\tau_D+\frac{\tau_{Fc}^{\,2}}{2\tau_D}\left(\frac{\omega_0}{R}\right)^2\right)+\left[\left(\tau_D+\frac{\tau_{Fc}^{\,2}}{2\tau_D}\left(\frac{\omega_0}{R}\right)^2\right)^2+\tau_{Fc}^{\,2}\left(1+\left(\frac{\omega_0}{R}\right)^2\right)+2\tau_D\tau_{Fc}\right]^{1/2} \tag{2.24}$$

and decreases again to zero for very high values of τ. We can find that $\tau_{\max} \approx \tau_{Fc}$ and $g_{Fc}(\tau) \approx 1$ in our experimental condition where $R >> \omega_0$. The cross correlation analysis gives essentially the same information about the molecular processes, such as diffusion, flow, and chemical reaction as autocorrelation analysis does. However, it should be noted that the cross correlation covers a much narrower range of time scale compared to the autocorrelation analysis.[27] (See Figure 2.6) This fact gives the possibility to exclude undesired process(es) that occur outside the time range under investigation. In Chapter 3, we demonstrate this advantage in the study of biomolecule conformational fluctuations, in which the time range is usually coupled with other processes such as triplet blinking and molecular diffusion.

2.1.1.4 Measuring Correlation Functions

In confocal microscopy based FCS, single photon counting methods are used to measure the autocorrelation function. Experimentally, this is accomplished by accumulating the detected photons into successive time bins of duration Δt. The fluorescence intensity, $I(t)$, at any given time is equivalent to the number of detected photons n_i, divided by the time interval, Δt, corresponding to $t = i\ \Delta t$. The autocorrelation function for a given lagtime is calculated from Equation (2.25), after an appropriately large number of time intervals have been accumulated[28]:

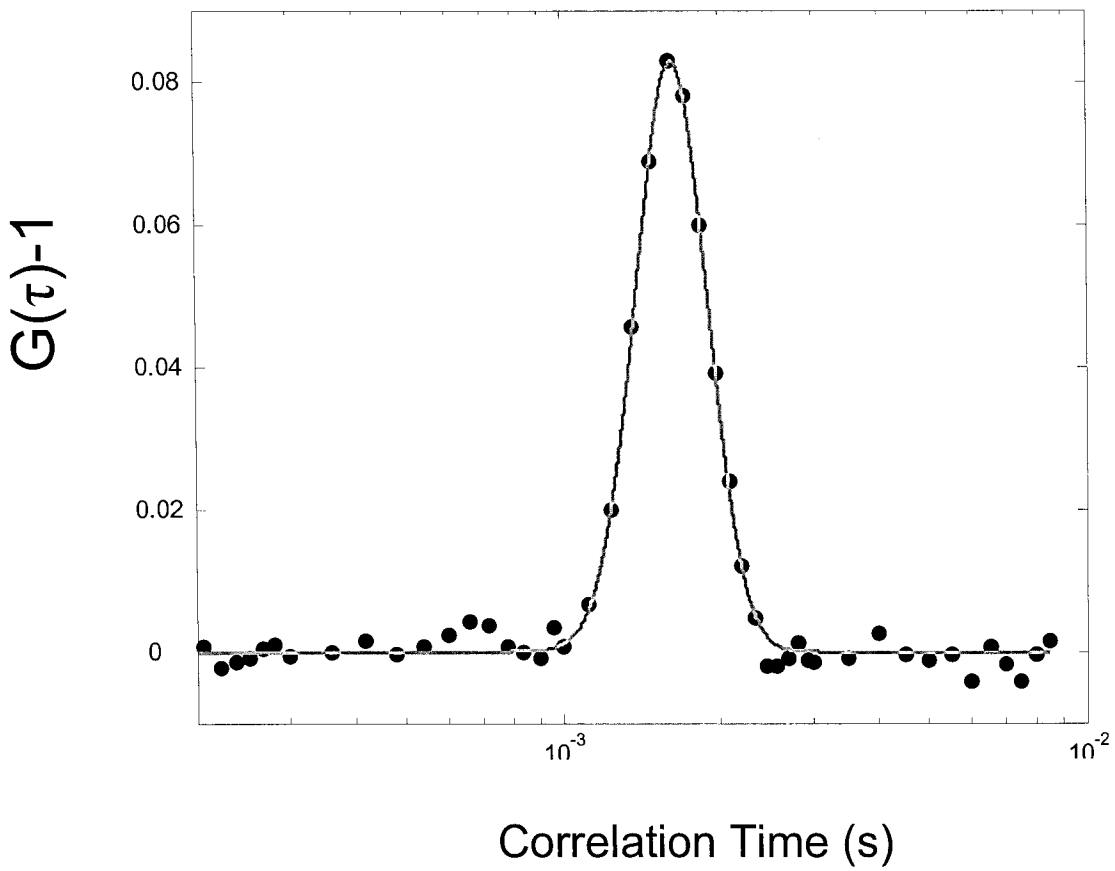

Figure 2.6 Cross-correlation data from 2bFCCS measurement of ~5nM Rhodamine 6G labeled polynucleotides (40mer of poly-Thymine)
Theoretical model (solid line) fits the experimental data (dots) according to Equation (2.22).

$$G(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I\rangle^2} = \frac{(M-k)\sum_{i=1}^{M-k} n_i n_{i+k}}{\left(\sum_{i=1}^{M-k} n_i\right)^2} \quad . \tag{2.25}$$

Here, M is the total number of time bins, and k indicates the time interval corresponding to lagtime $\tau = k\Delta t$. In practice, it is often convenient to allow the lengths of the successive time intervals to vary. Photons are initially collected into time bins of a few nanoseconds duration. Subsequent photons are then accumulated into time bins of increasingly longer durations ranging from tens of nanoseconds to seconds. This "multiple-τ" approach imparts sensitivity to fluctuations over a broad range of time scales (nanoseconds to tens of seconds) without requiring excessive data accumulations.

For cross correlation function[27],

$$G_{1,2}(\tau) = \frac{\langle I_1(t)I_2(t+\tau)\rangle}{\langle I_1\rangle\langle I_2\rangle} = \frac{\frac{1}{M-k}\sum_{t=1}^{M-k} n_1(t)n_2(t+\tau)}{\left(\frac{1}{M-k}\right)^2 \sum_{t=1}^{M-k} n_1(t)\sum_{t=1}^{M-k} n_2(t+\tau)} \tag{2.26}$$

where, 1 and 2 represent different detectors 1 and 2, such that $n_1(t)$ and $n_2(t+\tau)$ are the number of photocounts accumulated from detectors 1 and 2 at counting interval t and $t+\tau$, respectively. $G_{12}(\tau)$ and $G_{21}(\tau)$ are the "forward" and "reverse" cross-correlation functions of detector 1 relative to detector 2. In 2bFCCS, each function is only sensitive to molecules that flow in the one direction (i.e. forward direction, from focal region 1 to focal region 2 or opposite direction). In this way, 2bFCCS can resolve mixture of molecules flowing different direction.

2.1.2 Photon counting histogram

The photon counting histogram (PCH)[29], also referred to as the fluorescence intensity distribution analysis (FIDA)[30], is the latest development in FFS analysis, and perhaps the one that is currently experiencing the most widespread acceptance in many research areas. PCH was developed independently by Gratton and co-workers and Gall and co-workers in 1999. It is essentially a confocal microscopy based variation of a technique originally proposed by Qian and Elson in 1990 for analyzing the moments of the fluorescence intensity distribution in macroscopic sample volumes.[31, 32] PCH derives its chemical selectivity from differences in the molecular brightness, ϵ_i defined in Equation (1.12), of the analyte molecules. Fluorescence emitted from the confocal detection volume is monitored by accumulating the detected photons into successive time bins of equal sampling time per bin. If the duration of each bin is much shorter than the diffusion time of the molecules through the detection volume, then each bin represents a snapshot of the fluorescence emitted from the molecules occupying the detection volume at that particular moment in time. The fluorescence data is histogrammmed according to the number of photons detected per sampling time.(Figure 2.7) The shape of the histogram is a complicated function of the spatial distribution of the excitation/detection volume, the analyte concentration, and the molecular brightness of the analytes. The probability of

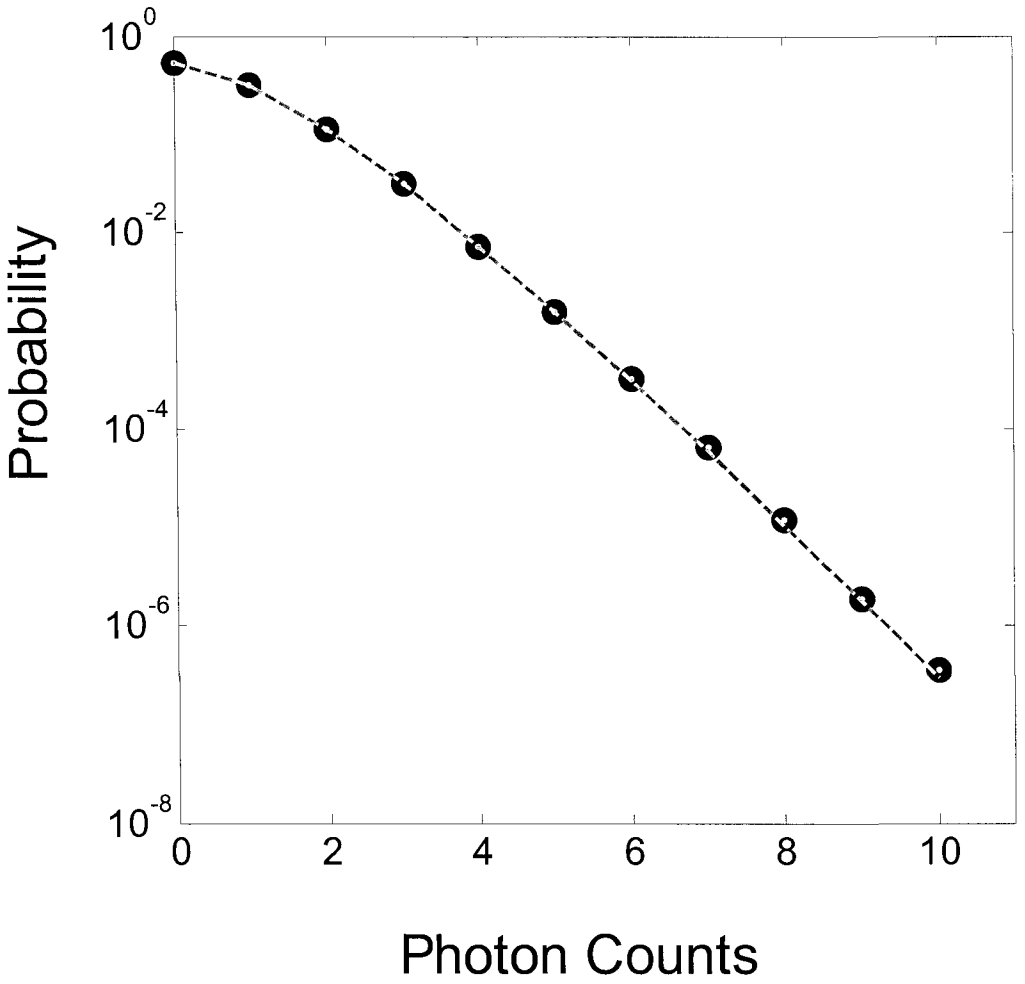

Figure 2.7 Normalized photon counting histogram data from PCH measurement of ~5nM Rhodamine 6G labeled polynucleotides (40mer of poly-Thymine) Theoretical model (solid line) fits the experimental data (dots) according to Equation (2.32). The photon counts represents the number of photons collected from sample within the time bin. In our experiments, 9-microsecond was used.

detecting k photons ($k > 0$) from one fluorescent molecule in a sufficiently large reference volume V_0, is given by:

$$p^{(1)}(k;V_o,\varepsilon) = \frac{1}{V_o}\int Poisson\left[k,\varepsilon\cdot W(\vec{r})\right]d\vec{r} \tag{2.27}$$

where, $Poisson(k, x)$ is the Poisson distribution that describes the probability of getting k photons when the average number of photons received is x:

$$Poisson(k,x) = \frac{x^k e^{-x}}{k!} \tag{2.28}$$

Chen et al. proposed the PCH integration using the 3D-Gaussian approximation for $W(r)$.

$$p_G^{(1)}(k;V_0,\varepsilon) = \frac{\pi\omega_0^{\,2} z_0}{V_0 k!}\int_0^\infty \gamma(k,\varepsilon e^{-2x^2})dx \tag{2.29}$$

where, $\gamma(k, x)$ is the incomplete gamma function given by [29]

$$\gamma(k,x) = \frac{\int_0^x t^{k-1}e^{-t}dt}{(k-1)!} \tag{2.30}$$

and, N particle probability function can be derived by convoluting the one particle probability function in N times:

$$p^{(N)}(k;V_0,\varepsilon) = (\underbrace{p^{(1)}\otimes p^{(1)}\otimes\cdots\otimes p^{(1)}}_{N}) \tag{2.31}$$

The photon counts probability describing a system consisting of one species is given by:

$$P(k;\bar{N},\varepsilon) = \sum_{N=0}^{\infty} p^{(N)}(k;\mathrm{q},\varepsilon)\cdot Poisson(N,\bar{N}) \tag{2.32}$$

and, that of a system consisting of multiple species is by the convolution approach:

$$P(k;\bar{N}_1,\varepsilon_1,\bar{N}_2,\varepsilon_2\cdots,\bar{N}_N,\varepsilon_N) = P(k;\bar{N}_1,\varepsilon_1)\otimes P(k;\bar{N}_2,\varepsilon_2)\otimes\cdots\otimes P(k;\bar{N}_N,\varepsilon_N) \tag{2.33}$$

Therefore, PCH can resolve multiple species in a mixture system based on their molecular brightnesses expressed in the distribution of fluorescence intensity amplitudes, while FCS analysis mainly depends on time-dependent fluctuation in the fluorescence intensity. Due to this aspect, FCS and PCH provide complementary information for the target system.

Several variations of PCH have been developed that enhance its chemical selectivity even further. For example, multiple distributions can be obtained by analyzing the photocount data using varying sampling times. Molecular diffusion causes the shape of the distribution to depend on the sampling time. This effect is ignored in conventional PCH by making the sampling time so small that the molecular motion is essentially frozen in time during each sampling interval. By characterizing the sampling time dependence over a large time scale (micro-seconds to milliseconds), one extracts the diffusion rates of the analytes in addition to their molecular brightness values. This is referred to as photon counting *multiple* histograms (PCMH)[33] and also as fluorescence intensity *multiple* distribution analysis (FIMDA)[34]. Another alternative is dual-color PCH[35]. The approach is based on same concept as two-dimensional FIDA (2d-FIDA)[36]. In these methods, the fluorescence is monitored on two detectors, each sensitive to different emission wavelengths or to orthogonal emission polarizations. A two-dimensional histogram is constructed according to the number of detected photons per bin for each detection channel. The shape of the histogram depends not only on the

analyte concentrations and molecular brightness values, but also on the emission wavelengths of the fluorophors or their rotational anisotropies, depending on whether the two detection channels are differentiated according to wavelength or polarization. Finally, fluorescence intensity and lifetime distribution analysis (FILDA) combines the molecular brightness information with the fluorescence lifetimes of the analytes[37]. The fluorescence is excited using a pulsed laser source, and each detected photon is recorded along with the elapsed time between the excitation pulse and the time of detection. The data is histogrammed according to the number of photons per bin and the sum of elapsed times for each bin. The resulting histogram reveals the concentrations, molecular brightness values, and fluorescence lifetimes of each analyte. FILDA is conceptually similar to a related technique developed by Seidel and co-workers, referred to as burst integrated fluorescence lifetime (BIFL) analysis[38-40]. In short, PCH based methods have been devised as powerful tool that can exploit differences in a variety of fluorescence characteristics. The PCH analysis has been applied mostly to the study of binding of ligands to their receptors because the binding event creates an enormous difference between the brightness of the bound complex and that of the free ligands, making it easy to discriminate the bound complex. In Chapter 4, we present for the first time, its application to the study of biomolecule conformational fluctuation between different species, which causes the changes of molecular brightness.

FFS is influencing many areas of research such as the characterization of photo-physical and photochemical processes, biomolecular conformational dynamics, adsorption/desorption and molecular diffusion at solid-liquid interfaces and biological membranes; molecular flow profiling in microfluidics devices; multicomponent elecrophoretic analysis; and intracellular molecular dynamics and imaging. In many of these examples, FFS is providing crucial new insight into the nature of the system that would be difficult or impossible to attain in any other way.

As you can see in the following Chapters, all the work presented in this dissertation are based on the our effort to develop an FFS-based analysis strategy for analyzing biomolecule conformational dynamics, such as folding and unfolding, which can extract all the relevant information from a single set of fluorescence data, obtained from the molecule of interest alone. This approach shows high information content in single molecule level, under the condition of thermodynamics equilibrium, high precision of the measured parameters, and high temporal resolution.

2.2 INSTRUMENTATION

2.2.1 Capillary Preparation

Our two beam FFS experiments were performed by focusing two laser beams at the center of capillary. By pressurizing solution along the capillary, target molecules can

flow from the first laser beam to the second laser beam. A square bore fused silica capillary (Polymicro, Phoenix, AZ) with 50μm-inner-diameter was chosen so that the two laser focal regions could be positioned at the center of capillary internal space with minimal inhomogeneous flow profile and beam distortion. We found that these problems are significant in circular capillary set up, especially at higher flow rates. A small section of the outer polyimide coating was dissolved in concentrated sulfuric acid at 75°C to create a transparent window for the laser focusing. The outer diameter of the capillary was around 400 μm. Therefore, the beam path through capillary glass is comparable to the conventional micro-glass slide which has ~170 μm thickness. To prevent sample adsorption to the capillary walls, the interior wall of the capillary was coated with poly(vinyl acohol) (PVA, Mw ~4×10^4g/mol, Sigma) according to the procedure of Belder *et al.*[41] A clean, 60 cm long capillary was filled with an acidified aqueous glutaraldehyde (Fisher, Houston, TX) in 800 μL of 0.5 M hydrochloric acid (Mallinckrodt, Hazelwood, MO). The filling can be confirmed by checking solution acidity with pH-indication paper. A plug of 2.5 wt % PVA in 0.6 M hydrochloric acid was injected into the glutaraldehyde-filled capillary using 0.5 MPa of nitrogen gas for 10 seconds. The immobilization of the polymer on the capillary surface can be achieved by fast cross-linking reaction of PVA with glutaraldehyde, the cross-linking agent. The capillary was then emptied and dried with continuous N_2 gas flow for ~4 hours, after which it was ready for use. Capillary cleaning was accomplished by rinsing with a mild

surfactant and a blank buffer solution prior to use. The PVA-coated capillary is placed onto flat glass slide using optical glue which needs UV light exposing. The angle between excitation laser light path and capillary surface is critical for the FFS experiment and should be perpendicular so that beam distortion can be minimized. It can be easily checked by monitoring change of beam shape along the capillary surface. Any change in the shape results from tilted surface relative to the beam path.

2.2.2 Optical Setup

The experimental setups used for present studies(See Figure 2.8) were modified from our two-beam fluorescence cross-correlation spectroscopy/ capillary electrophoresis experiment.[42] The micro-tube containing sample solution was placed inside a gas chamber connected to N_2 gas tank. The sample solutions were flowed continuously through the capillary by pressurizing the capillary inlet with N_2 gas. A pneumatic pressure regulator (Fairchild Model 81, Winston-Salem, NC) held the pressure of the N_2 gas constant. The sample flow can also be generated by applying high electric field in the same manner as the capillary electrophoresis experiments use. Fluorescence Rhodamine 6-Green (R6G) fluorescent dye. The laser beam was expanded and collimated by a 4× telescope, split, and then recombined into two nearly parallel beams by two 50/50 beam splitters (Newport, Irvine, CA.). The laser power of each beam was equalized using the appropriate absorptive neutral density filters (Newport). A laser power of ~ 100-μW

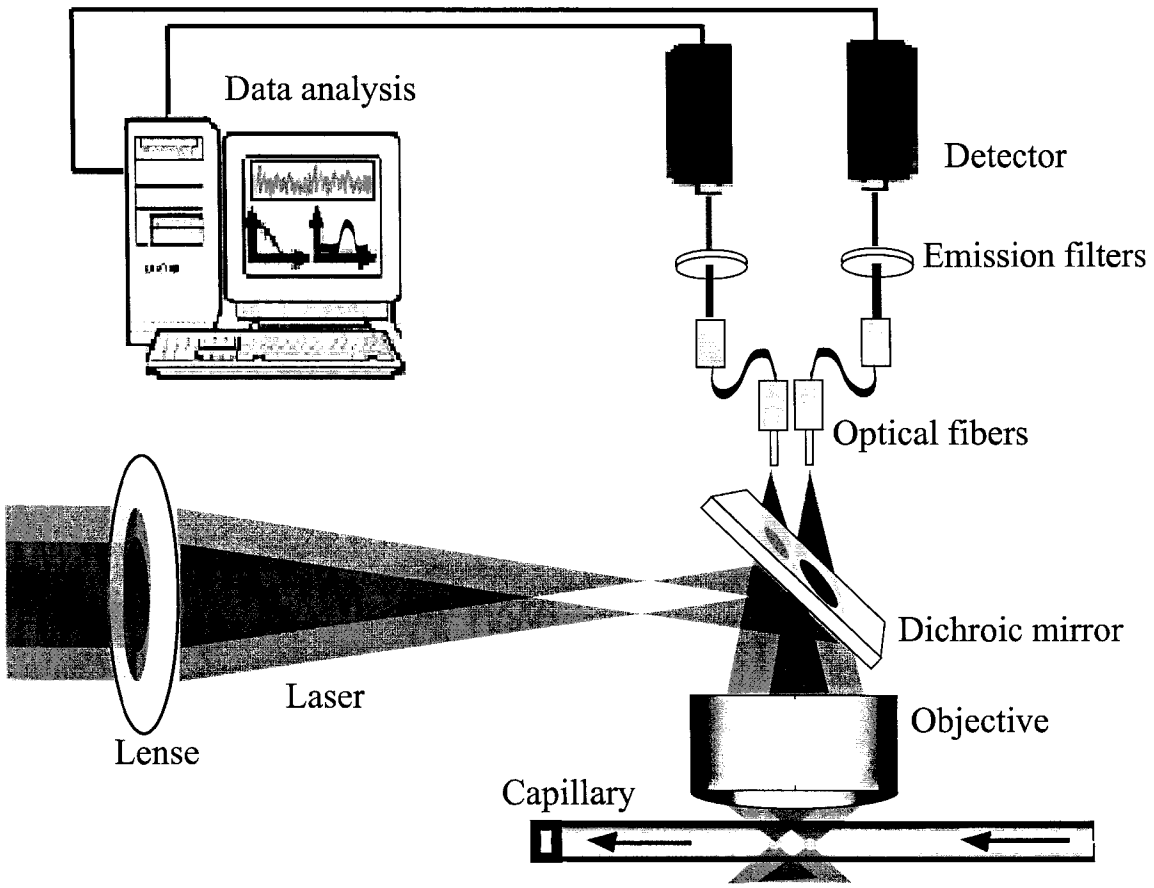

Figure 2.8 Schematic representation of the two- beam FCS experiment
The optical setup was designed to position the two laser beams in the center of the squared capillary, separated by a distance of *R*. The sample solution flows through the capillary under applied gas pressure. Fluorescence from each laser focal volume is collected and used for the correlation analyses.

measurements were performed using a home-built single molecule confocal fluorescence microscope. A 514.5 nm cw laser beam from an air-cooled, Ar^+ laser (Omnichrome/Melles-Griot, Carlsbad, CA) was used as the excitation source for the per beam was chosen to offer the best compromise between maximum signal-to-noise ratio, and minimum interference from photo-driven processes (*i.e.,* singlet to triplet intersystem crossing). The laser beams were prefocused by a 150-mm focal length spherical lens; reflected by a 530-nm long pass dichroic beam splitter (CVI, Albuquerque, NM); and focused into the capillary by a 100× 1.25 NA oil immersion microscope objective (Edmund Industrial Optics, Barrington, NJ). The distance between the focal lens and back aperture of microscope objective determine the size of laser focus at the sample solution. The longer distance provides a bigger detection volume. The back aperture should be underfilled by the laser beam in order to avoid severe deviation of laser focus from idea Gaussian beam profile. However, overfilling provides smaller detection volume which can achieve better signal to ratio in general. We adjusted the distance so that about 85% of the aperture area is filled with excitation beam. Fluorescence from each focal region was collected through the same objective and imaged onto the apertures of two 50-μm-core-diameter multimode optical fibers or pinholes positioned at the image plane of the objective. It provides axial resolution by rejecting out-of-focus light. The output of the fibers was band-pass filtered (550DF30, Chroma, Brattleboro, VT) and

focused onto the active areas of two single-photon-counting avalanche photodiode detectors (PerkinElmer Optoelectronics, model SPCM-AQR-14,Wellesley, MA).

The two laser beams formed nearly identical diffraction-limited focal regions, positioned near the center of the inner capillary space (~25-μm from the inner surfaces), separated along the axial dimension of the capillary by a distance, R. From autocorrelation analysis of a standard R6G solution, the e^{-2} focal radius in the radial dimension, ω_0, of the focal volume was determined to be ~ 0.3 μm. The ratio of the radial and axial e^{-2} radii ($\kappa_0=\omega_0/z_0$, where z_0 is the axial radius) was also determined, and found to be ~ 0.1. The separation distance, R, between the two foci was determined by cross-correlation analysis of the same sample and found to be ~5-μm. The interbeam distance can be adjusted by changing the angle of beam splitters used for splitting the original excitation laser beam into two. If the distance becomes very small (i.e. R < 1-μm), the overlap between the individual molecular detection efficiencies increases. This results in an increase of the probability that a photon from the detection volume one, V_1, arrives at detector 2 and vice versa. In the case of cross-correlation, this is equivalent to an autocorrelation, since the correlated photons originated from the same volume element. With smaller diameter of pinhole or fiber optics, the pseudo-autocorrelation effects can be minimized, but the overall collection efficiency would be poorer. If the distance becomes very large (i.e. R > 10-μm), the signal to noise ratio is too low due to the loss of correlation between fluorescence signals collected from individual molecule at two

detection volumes. In other words, most of the molecules passing the first detection volume diffuse out of the flow path, so that they can not reach into the second detection volume. Also, some parts of the laser beam are cropped by an optical path. The position of the laser beam foci relative to the inner surface of the capillary was reproducibly controlled using a sub-micrometer resolution differential micrometer (Newport) mounted on the z-axis of the sample stage. Additionally, a precision rotation stage (Newport) was mounted on the sample stage to allow adjustment of the flow axis relative to the axis defined by the position of the two laser beams. The optimum position of the two laser beam foci relative to the *z*-position of the capillary and the flow axis could be confirmed by cross-correlation analysis. The fastest correlation peak position and highest peak amplitude is the general indication for the optimum position. Also the symmetry in the peak shape is achieved from well aligned setup. The quality of laser alignment was monitored by checking the shape of fluorescence light reflected from an additional mirror positioned before the detectors. The entire experiment was placed inside an enclosure to block stray laser light. All experiments were carried out at an ambient temperature of ~19°C.

2.2.3 Data Acquisition

The photocounts from the two detectors (detectors 1 and 2) were recorded using the two channels of a 800-MHz gated photon counter (PMS-400, Becker & Hickl GmbH,

Berlin, Germany), interfaced to a Pentium computer. A sampling time of 1-μs per sampling interval was used for each channel. A LabView (National Instruments, Austin, TX) computer program, written in-house, was used to perform real-time correlation analysis on the detected photons as they were being accumulated. The program simultaneously calculated a set of four normalized correlation functions, as well as a photon counting histogram. The present study only reports our analysis of the correlation functions, which were obtained using the following set of equations:

$$G_{l,m}(\tau) = \frac{\langle I_l(t) I_m(t+\tau) \rangle}{\langle I_l \rangle \langle I_m \rangle} = \frac{\frac{1}{M-k} \sum_{t=1}^{M-k} n_l(t) n_m(t+\tau)}{\left(\frac{1}{M-k}\right)^2 \sum_{t=1}^{M-k} n_l(t) \sum_{t=1}^{M-k} n_m(t+\tau)} \tag{2.34}$$

where, l and m represent detectors 1 and 2. M is the total number of sampling intervals per scan for each detector, k is the number of sampling intervals corresponding to lagtime τ, and $n_l(t)$ and $n_m(t+\tau)$ are the number of photocounts accumulated from detectors l and m at counting interval t and $t+\tau$, respectively. $G_{11}(\tau)$ and $G_{22}(\tau)$ are the autocorrelation functions for each individual detector. With careful equalization of the laser power in each beam, and careful adjustment of the optical alignment, the two autocorrelation functions are identical within experimental error. $G_{12}(\tau)$ and $G_{21}(\tau)$ are the "forward" and "reverse" cross-correlation functions of detector 1 relative to detector 2. The cross-correlation functions are calculated after first re-binning the accumulated photocounts into 9-μs sampling intervals to achieve improved signal-to-noise ratio. $G_{12}(\tau)$ is only

sensitive to molecules that flow in the forward direction, from focal region 1 to focal region 2; whereas, $G_{21}(\tau)$ only analyzes molecules flowing in the opposite direction. For all the experiments reported here, the flow takes place in the forward direction. Hence, only $G_{12}(\tau)$ is analyzed. All correlation functions were calculated using the "multiple-τ" concept, in which the intervals between successive lagtimes increase as the lagtime increases. The lagtime can start from 50 ns and cover up to 1 second. The correlation functions were calculated for successive scans of 65,536 sampling intervals each. The rebinned photocounts of 9-μs sampling intervals were also used to construct photon counting histograms. The histograms obtained from each detection channel were identical within experimental error. The user interface of the LabView program continuously displays a running average of the correlation functions and the cumulative histograms. In most experiments, the final set of correlation functions represent averages of ~1500 scans, corresponding to a total data accumulation time of ~100-s. There is another panel displaying the real-time fluorescence signal intensity. By monitoring the stability of the intensity time trace during the experiment, any artificial effects (i.e. surface interaction of sample, photobleaching, etc.) can be checked.

2.3 Conformational Fluctuation Represented by Fluorescence Fluctuation

In this section, we point out the motivation of the 2bFFS based approach for the study of biomolecule conformational dynamics. As described previously, fluorescence

correlation spectroscopy (FCS) is a technique for analyzing the spontaneous fluctuations in the fluorescence signal arising from a microscopic subvolume of a fluorescent sample.[43-49] FCS measures a temporal correlation function that depends on the time-dependent molecular processes giving rise to the observed fluctuations. In principle, the correlation function can be analyzed to determine accurate values for the relaxation times associated with the different fluctuation process. The time resolution can range from the low nanoseconds to many milliseconds. Some examples of the types of molecular processes studied include hydrodynamic processes, such as translational[50] and rotational[51] diffusion, laminar flow,[45, 52-54] and electrophoretic flow;[42, 55-57] chemical and biochemical processes, such as enzymatic activity;[58] biophysical processes, such as biomolecule conformational fluctuations;[59-65] and photodynamic processes, such as photochemical isomerization[66] and triplet intersystem crossing.[67] FCS is advantageous because of its ability to characterize these processes non-invasively, under conditions of thermodynamic equilibrium. The measurements are carried out *in situ*. Macroscopic perturbation of the system equilibrium to establish a start time for the reaction is not required. This is a key advantage over traditional reaction rate analysis techniques such as Temperature jump (T-jump), Rapid mixing, Fluorescence Recovery After Photobleaching (FRAP), as well as Laser induced T-jump spectroscopy. It will become increasingly important as FCS becomes more widely used to study molecular processes in the cellular environment.[68-71] Reactions taking place in cells are often

coupled to an extremely complex series of chemical processes. Macroscopic perturbation of such a highly coupled system can lead to a complex relaxation process that would be difficult or impossible to analyze. In addition to the problem, it is highly difficult to predict the process in near minimum of the free energy surface, based on the observation under different physical conditions (i.e. temperature) (See Figure 2.9) NMR relaxation methods, though they do not require perturbation of the system, currently lack the sensitivity and spatial resolution for single cell analysis. Therefore, FCS and related techniques offer the best hope for directly observing biochemical and biophysical processes as they are happening inside the cell.

Difficulties with FCS analysis arise, however, when multiple processes are occurring simultaneously in the same system, especially when the relaxation times of the different processes occur on similar time scales. This is often the case when studying the conformational fluctuations of freely diffusing biomolecules. Conformational transformations of a biomolecule give rise to fluorescence intensity fluctuations when the internal molecular motion is coupled to quenching of a dye molecule, or to fluorescence resonance energy transfer (FRET) between donor and acceptor molecules attached to specific regions of the biomolecule. The resulting correlation function thus contains fluctuation contributions characterized by a diffusional relaxation time, τ_D, and a relaxation time for conformational fluctuations, τ_R. Except when $\tau_D >> \tau_R$, the two parameters will be strongly coupled, such that statistically accurate determination of one

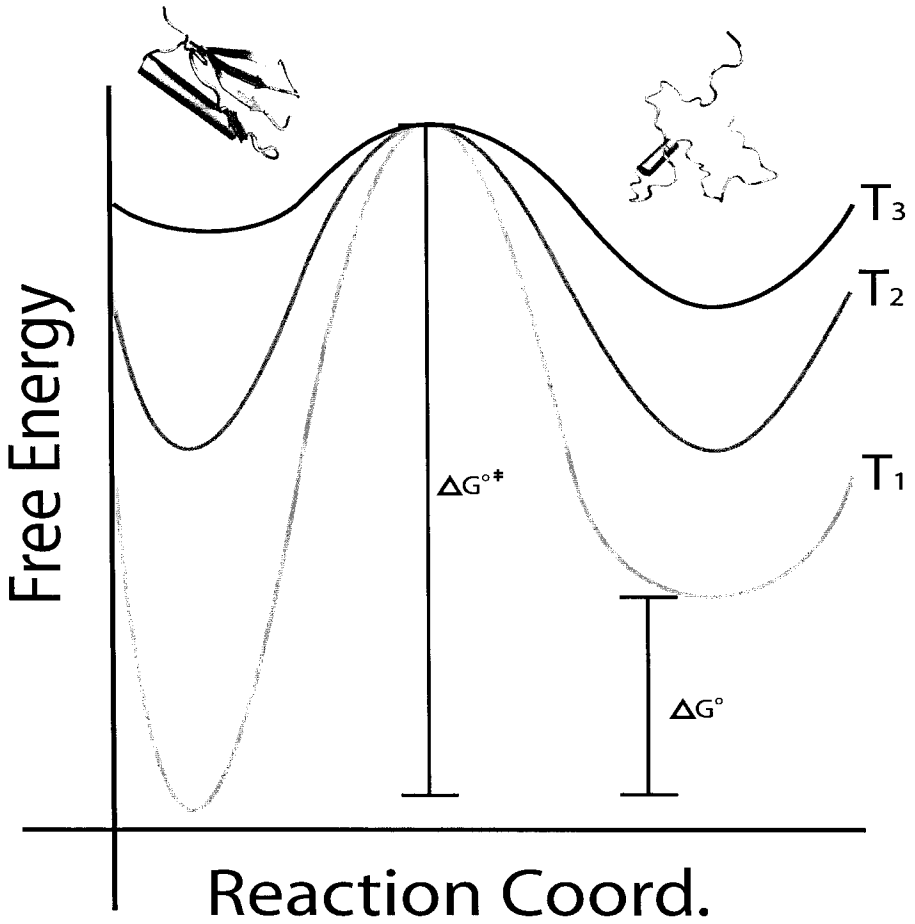

Figure 2.9 Schematic representation of free energy surfaces along with reaction coordinate at various temperature conditions
Under different conditions such as experimental temperature, the reaction of even identical molecules can not be described by same scenario. For complicate system like biomolecule conformational change, perturbation given to the system may cause significant changes on the energy landscape and lead unpredicted behaviors in the reaction. Therefore equilibrium approach such as FCS would be much more preferred.

parameter requires independent determination of the other. The usual way of decoupling the different relaxation times is to design a control molecule that is either missing the quencher or FRET acceptor, or for which the folding and unfolding process is suppressed.[59, 62-64] FCS analysis of the control molecule gives an independent measure of τ_D, which then constrains the correlation function for the molecule of interest, thereby allowing the determination of τ_R. There are several disadvantages to this approach, aside from the ever-present possibility of being misled by the assumptions being made about the nature of the control molecule. For example, the τ_D parameter is related to the decay rate of the correlation function, but not the correlation amplitude, which relates to with analyte concentrations. We have found that maximum precision is achieved only when both τ_D and the correlation amplitude are properly constrained.(See Supple. Info. in Chpater 3) But the amplitude is very difficult to constrain when analyzing different samples separately. Unless the relative concentrations of the control molecule and the molecule of interest are known precisely—a very difficult proposition when dealing with the extremely low concentrations normally encountered in FCS analysis, and especially, when working in the cellular environment—there is no way to constrain the amplitudes with enough precision from an independent measurement of the control molecule. One way of getting around this problem is to simply divide out the measured correlation function of the control molecule, without regard for the concentration differences. But in so doing, one sacrifices the physically relevant

information contained in the correlation amplitudes (*i.e.* the concentration, the equilibrium distribution, and the fluorescence intensity information). Finally, if the ultimate goal is to carry out these types of measurements in the cellular environment, it becomes difficult to imagine how this would be done using the control molecule technique. It would be much more desirable to extract all the relevant information about the dynamic processes of the molecule of interest from a single set of fluorescence intensity measurements, obtained from the sample of interest alone.

Klenerman and co-workers have devised FCS-based analysis techniques that overcome the aforementioned problems in many situations.[60, 61, 65] In one application, the authors used two-color measurements to analyze the anticorrelated fluorescence fluctuations from two different fluorophors attached to a DNA hairpin molecule as they interacted via FRET.[60, 61] The resulting proximity ratio autocorrelation function characterized the relaxation of the FRET efficiency, independent of molecular diffusion. This method is useful when the fluorophors interact via FRET. It does not apply to dye-quencher interactions. FRET, though widely used, is not without complications, and is, thus, not always the preferred mechanism for reporting dynamic motion of biomolecules. There are many situations where it would be preferable to analyze a fluorophor-quencher type interaction. For example, if the donor and acceptor molecules come too close together, rather than experiencing maximum FRET efficiency, they can often simply quench each other. Also, FRET occurs over a fairly broad range of molecular distances,

whereas, quenching interactions are often restricted to direct orbital overlap of the fluorophor and quencher. This can provide greater specificity in honing in on selected molecular contacts.[72] Finally, there are several examples of nucleotide bases, amino acid residues, and other naturally occurring chemical functional groups that can serve as natural fluorescence quenchers,[73, 74] which could be used to advantage in designing fluorescent labeling strategies.

In a more recent study, Klenerman and co-workers extended the dual color correlation analysis strategy to include the characterization of dye-quencher interactions.[65] Intensity ratios from two non-interacting fluorophors, which nonetheless undergo independent quenching and unquenching interactions with different functional groups on the biomolecule of interest, were analyzed. This technique is referred to as ratiometric autocorrelation analysis. Proximity ratio and ratiometric fluorescence autocorrelation analysis have been applied to the folding and unfolding kinetics of DNA hairpin molecules and other nucleic acid constructs. The absence of a diffusion contribution allowed direct measurement of the parameters associated with non-exponential relaxation of the folding and unfolding reaction. However, these techniques only apply to molecules that can be appropriately labeled with FRET pairs, or with non-interacting pairs of dyes. They do not apply to fluctuations caused by the interaction of a single dye-quencher pair.

Another related technique for direct observation folding and unfolding processes is to analyze the time-dependent fluorescence emission from individual isolated molecules.[75-79] There are many examples where this approach has been used successfully. Folding and unfolding reactions have been characterized for individual molecules tethered to a surface, confined in a vesicle, or even freely diffusing in aqueous solution.[63, 80] One of the main limitations of this approach lies in its low temporal resolution. Because of detector dark counts, shot noise, and other sources of background interference, the fluorescence photocounts must be accumulated in sufficiently long sampling intervals in order to observe the intensity fluctuations caused by conformational changes. The minimum sampling times are typically tens to hundreds of microseconds for most of these experiments. If the relaxation process is occurring on a time scale similar to or less than the sampling interval, the single molecule analysis will have an inherent bias towards events taking place on a much longer time scale than the average relaxation time, making it impossible to measure the relaxation time with precision. Correlation analysis analyzes photons accumulated from many single molecule events over a time scale much longer than the measurement time for any one molecule and thus achieves high temporal resolution by averaging many events. In principle, temporal resolution in the low nanoseconds can be achieved, ultimately limited by the fluorescence lifetime of the fluorophors.[81] The ability to resolve heterogeneity at the single molecule level is, unfortunately, sacrificed. But the advantage of obtaining ensemble averaged

kinetics information with excellent temporal resolution and without perturbing the system is retained.

Our goal has been to develop an FCS-based analysis strategy for analyzing biomolecule folding and unfolding kinetics that can extract all the relevant dynamics information from a single set of fluorescence data, obtained from the molecule of interest alone. We have strived for high information content, high precision of the measured parameters, and high temporal resolution. We also desire a technique that is as effective at analyzing fluorophor-quencher fluctuations as it is at analyzing changes in FRET efficiency. Our approach is to detect fluorescence from molecules as they flow sequentially between two spatially offset laser probe volumes. The fluorescence from each probe volume is imaged onto separate detectors, and the detected photocounts are analyzed simultaneously from three different vantage points—cross-correlation analysis of the two detection channels relative to each other, autocorrelation analysis of the two detection channels independently, and photon counting histogram (PCH) analysis. We show that the cross-correlation function characterizes the hydrodynamic properties (flow rate and diffusion rate) of the molecules under study, as well as the correlation amplitude, independent of conformational fluctuations. These parameters can then be used to constrain the measured autocorrelation function, from which precise measurements of the folding and unfolding kinetics are obtained. The relevant kinetic parameters include the relaxation time, τ_R, an amplitude term, B_{eq}, which depends on the equilibrium distribution

and the emission intensities of the folded and unfolded molecules, and a "stretch" parameter,β, which characterizes the degree of heterogeneity in the reaction mechanism. In the Chapter 3, our works investigating the ability of the PCH analysis to resolve the equilibrium distribution of folded and unfolded conformations are discussed. To demonstrate these principles, we analyzed the folding and unfolding kinetics of single-stranded DNA oligonucleotides that form a stem-loop DNA hairpin structure shown in Figure 1.2. Dynamic equilibrium between the folded and unfolded conformations is monitored by way of quenching and unquenching of a dye and a quencher attached to either end of the DNA strand. Assuming two state mechanism for the reaction, our measurements of the relaxation times of the different sized hairpins are complimentary to previous studies on similar systems. In particular, we found evidence for non-exponential relaxation, consistent with the expectation that DNA hairpin formation occurs via a heterogenous reaction mechanism. However, we found the reaction should be described by a three-state mechanism which contains a stable intermediate state along the reaction paths. The DNA hairpin folding mechanism is thought to involve a rapid equilibrium between open and intermediate configurations, and a much slower formation of the fully folded stem-loop hairpin. Hence, we describe a DNA hairpin reaction that is even more complex and more drawn out than previously thought.

2.4 REFERENCE

1. Van Orden, A., K. Fogarty, and J. Jung, *Fluorescence Fluctuation Spectroscopy: A Coming of Age Story.* Appl. Spectrosc., 2004. **58**: p. 122A-137A.
2. Magde, D., W.W. Webb, and E.L. Elson, *Thermodynamic fluctuations in a reacting system- Measurement by fluorescence correlation spectroscopy.* Phys. Rev. Lett., 1972. **29**: p. 705.
3. Magde, D., W.W. Webb, and E.L. Elson, *Fluorescence Correlation Spectroscopy 3. Uniform translation and laminar flow.* Biopolymers, 1978. **17**: p. 361-376.
4. Elson, E.L. and D. Magde, *Fluorescence Correlation Spectroscopy 1. Conceptual Basis and Theory.* Biopolymers, 1974. **13**: p. 1-27.
5. Widengren, J. and R. Rigler, *Fluorescence Correlation Spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces.* Cell. mol. Biol., 1998. **44**: p. 857-879.
6. Schmitz, K.S., *An Introduction to Dynamic Light Scattering by Macromolecules.* 1990, San Diego: Academic Press.
7. Berne, B.J. and R. Pecora, *Dynamic Light Scattering: with Application to Chemistry, Biology and Physics.* 2000, Mineola, NY: Dover Publications.
8. Rigler, R. and U. Mets, *Laser spectroscopy of Biomolecules". Diffusion of Single Molecules through Gaussian Laser Beam.* Proc. SPIE-Int. Soc. Opt Eng., 1992. **1921**: p. 239-248.
9. Rigler, R., J. Widengren, and U. Mets, *Fluorescence Spectroscopy*, ed. Wolfbeis. 1993, Berlin: Springer.
10. Rigler, R., et al., *Fluorescence Correlation Spectroscopy with High Count Rate and Low Background: Analysis of Translational Diffusion.* Eur. Biophys. J., 1993. **22**: p. 169-175.
11. Rigler, R. and U. Mets, *Submillisecond Detection of single Rhodamine Molecules in Water.* J. Fluorescence, 1994. **4**: p. 259-264.
12. Rigler, R., *Fluorescence Correlations, Single Molecule Detection and Large Number Screening. Applications in Biotechnology.* J. Biotechnol., 1995. **41**: p. 177-186.
13. Nie, S.M., D.T. Chiu, and R.N. Zare, *Probing individual molecules with confocal fluorescence microscopy.* Science, 1994. **266**: p. 1018-1021.
14. Shera, E.B., et al., *Detection of Single Fluorescent Molecules.* Chem. Phys. Lett., 1990. **174**: p. 553.

15. Berland, K.M., et al., *Scanning Two-photon Fluctuation Correlation Spectroscopy: Particle Counting Measurements for Detection of Molecular Aggregation.* Biophys. J., 1996. **71**: p. 410.

16. Schwille, P., F.J. MeyerAlmes, and R. Rigler, *Dual-color Fluorescence Cross-Correlation Spectroscopy for Multicomponent Diffusional Analysis in Solution.* Biophys. J., 1997. **72**: p. 1878-86.

17. Kettling, U., et al., *Real-time Enzyme Kinetics Monitored by Dual-Color Fluorescence Cross-Correlation Spectroscopy.* Proc. Natl. Acad. Sci. USA, 1988. **95**: p. 1416-1420.

18. Koltermann, A., et al., *Rapid Assay Processing by Integration of Dual-Color Fluorescence Cross-Correlation Spectroscopy: High Throughput Screening for Enzyme Activity.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 1421-1426.

19. Winkler, T., et al., *Confocal Fluorescence Coincidence analysis: an Approach to Ultra High-Throughput Screening.* Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 1375-1378.

20. Rigler, R. and E.L. Eigen, *Fluorescence Correlation Spectroscopy: Theory and Applications.* 2001, Berlin: Springer-Verlag.

21. Kohl, T., et al., *A Protease Assay for Two-Photon Crosscrrelation and FRET Analysis Based Solely on Fluorescent Protein.* Proc. Natl. Acad. Sci. USA, 2002. **99**: p. 12161.

22. Weidemann, T., et al., *Analysis of Ligand Binding by Two-Colour Fluorescence Cross-Correlation Spectroscopy.* Single Molecules, 2002. **3**: p. 49 - 61.

23. Brinkmeier, M. and R. Rigler, *Flow Analysis by Means of Fluorescence Correlation Spectroscopy.* Techn. Phys., 1995. **41**: p. 205-210.

24. Brinkmeier, M., et al., *Confocal Spectroscopy in Microstructures.* Biophys. Chem., 1997. **66**: p. 229-239.

25. Brinkmeier, M., et al., *Two-Beam Cross-Correlation: a Method to Characterize Transport Phenomena in Micrometer-Sized Structures.* Anal. Chem., 1999. **71**: p. 609-616.

26. Dittrich, P.S. and P. Schwille, *Spatial Two-Photon Fluorescence Cross-Correlation Spectroscopy for Controlling Molecular Transport in Microfluidic Structures.* Anal. Chem., 2002. **74**: p. 4472-4479.

27. Jung, J. and A. Van Orden, *Folding and Unfolding Kinetics of DNA Hairpins in Flowing Solution by Multiparameter Fluorescence Correlation Spectroscopy.* J. Phys. Chem. B, 2005. **109**: p. 3648-3657.

28. Zander, C., J. Enderlein, and R.A. Keller, *Single Molecule Detection in Solution: Methods and Applications*. 2002, Derlin: Wiley-VCH.

29. Chen, Y., et al., *The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy*. Biophys. J., 1999. **77**: p. 553-567.

30. Kask, P., et al., *Fluorescence Intensity Distribution Analysis and Its Application in Biomolecular Detection Technology*. Proc. Natl. Acad. Sci. USA, 1999. **96**(24): p. 13756-13761.

31. Qian, H. and E.L. Elson, *Distribution of Molecular Aggregation by Analysis of Fluctuation Moments*. Proc. Natl. Acad. Sci. USA, 1990. **87**: p. 5479.

32. Qian, H. and E.L. Elson, *On the Analysis of High Order Moments of Fluorescence Fluctuations*. Biophys. J., 1990. **57**: p. 375.

33. Perroud, T.D., B. Huang, and R.N. Zare, *Effect of Bin Time on the Photon Counting Histogram for One-Photon Excitation*. Chemphyschem, 2005. **6**: p. 905-912.

34. Palo, K., et al., *Fluorescence Intensity Multiple Distributions Analysis: Concurrent Determination of Diffusion Times and Molecular Brightness*. Biophys. J., 2000. **79**: p. 2858-2866.

35. Chen, Y., et al., *Dual-Color Photon Counting Histogram*. Biophys. J., 2005. **88**: p. 2177-2192.

36. Kask, P., et al., *Two-Dimentional Fluorescence Intensity Distribution Analysis: Theory and Applications*. Biophys. J., 2000. **78**: p. 1703-1713.

37. Palo, K., et al., *Fluorescence Intensity and Lifetime Distribution Analysis: toward Higher Accuracy in Fluorescence Fluctuation Spectroscopy*. Biophys. J., 2002. **83**: p. 605.

38. Eggeling, C., et al., *Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy*. Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 1556-1561.

39. Fries, J.R., et al., *Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy*. J. Phys. Chem. A, 1998. **102**: p. 6601.

40. Schaffer, J., et al., *Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy*. J. Phys. Chem. A, 1999. **6**: p. 11.

41. Belder, D., et al., Electrophoresis, 2001. **22**: p. 3813-3818.

42. Fogarty, K. and A. Van Orden, *Two-beam fluorescence cross-correlation spectroscopy for simultaneous analysis of positive and negative ions in continuous-flow capillary electrophoresis.* Anal. Chem., 2003. **75**(23): p. 6634-6641.

43. Elson, E.L. and D. Magde, *Fluorescence correlation spectroscopy I. Conceptual basis and theory.* Biopolymers, 1974. **13**: p. 1-27.

44. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence correlation spectroscopy II. An experimental realization.* Biopolymers, 1974. **13**: p. 29-61.

45. Magde, D., W.W. Webb, and E.L. Elson, *Fluorescence correlation spectroscopy III. Uniform translation and laminar flow.* Biopolymers, 1978. **17**: p. 361-376.

46. Eigen, M. and R. Rigler, *Sorting Single Molecules - Application to Diagnostics and Evolutionary Biotechnology.* Proc. Natl. Acad. Sci. USA, 1994. **91**(13): p. 5740-5747.

47. Widengren, J. and R. Rigler, *Review - Fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces.* Cell. Mol. Biol., 1998. **44**(5): p. 857-879.

48. Rigler, R. and E.S. Elson, eds. *Fluorescence Correlation Spectroscopy: Theory and Applications*. Springer Series in Chemical Physics, ed. W. Zinth. Vol. 65. 2001, Springer-Verlag: Berlin.

49. Van Orden, A., K. Fogarty, and J. Jung, *Fluorescence Fluctuation Spectroscopy: A Coming of Age Story.* Appl. Spectrosc., 2004. **58**(5): p. 122A-137A.

50. Rigler, R., et al., *Fluorescence Correlation Spectroscopy with High Count Rate and Low-Background - Analysis of Translational Diffusion.* Eur. Biophys. J. Biophys. Lett., 1993. **22**(3): p. 169-175.

51. Kask, P., et al., *Separation of the rotational contribution in fluorescence correlation experiments.* Biophys. J., 1989. **55**: p. 213-220.

52. Gosch, M., et al., *Hydrodynamic flow profiling in microchannel structures by single molecule fluorescence correlation spectroscopy.* Anal. Chem., 2000. **72**(14): p. 3260-3265.

53. Dittrich, P.S. and P. Schwille, *Spatial two-photon fluorescence cross-correlation Spectroscopy for controlling molecular transport in microfluidic structures.* Anal. Chem., 2002. **74**(17): p. 4472-4479.

54. Kunst, B.H., A. Schots, and A.J.W.G. Visser, *Detection of flowing fluorescent particles in a microcapillary using fluorescence correlation spectroscopy.* Anal. Chem., 2002. **74**: p. 5350-5357.

55. Van Orden, A. and R.A. Keller, *Fluorescence correlation spectroscopy for rapid multicomponent analysis in a capillary electrophoresis system.* Anal. Chem., 1998. **70**(21): p. 4463-4471.

56. Brinkmeier, M., et al., *Two-beam cross-correlation. A method to characterize transport phenomena in micrometer-sized structures.* Anal. Chem., 1999. **71**(3): p. 609-616.

57. LeCaptain, D.J. and A. Van Orden, *Two-beam fluorescence cross-correlation spectroscopy in an electrophoretic mobility shift assay.* Anal. Chem., 2002. **74**(5): p. 1171-+.

58. Kohl, T., et al., *A protease assay for two-photon crosscorrelation and FRET analysis based soley on fluorescent proteins.* Proc. Natl. Acad. Sci. USA, 2002. **99**: p. 12161-12166.

59. Bonnet, G., O. Krichevsky, and A. Libchaber, *Kinetics of conformational fluctuations in DNA hairpin loops.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 8602-8606.

60. Wallace, M.I., et al., *FRET fluctuation spectroscopy: Exploring the conformational dynamics of a DNA hairpin loop.* J. Phys. Chem. B, 2000. **104**: p. 11551-11555.

61. Wallace, M.I., et al., *Non-Arrhenius kinetics for the loop closure of a DNA hairpin.* Proc. Natl. Acad. Sci. USA, 2001. **98**(10): p. 5584-5589.

62. Chattopadhyay, K., et al., *Measurement of microsecond dynamic motion in the intestinal fatty acid binding protein by using fluorescence correlation spectroscopy.* Proc. Natl. Acad. Sci. USA, 2002. **99**: p. 14171-14176.

63. Margittai, M., et al., *Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1.* Proc. Natl. Acad. Sci. USA, 2003. **100**: p. 15516-15521.

64. Slaughter, B.D., et al., *Single-molecule resonance energy transfer and fluorescence correlation spectroscopy of calmodulin in solution.* J. Phys. Chem. B, 2004. **108**: p. 10388-10397.

65. Li, H., et al., *Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy.* Proc. Natl. Acad. Sci. USA, 2004. **101**: p. 14425-14430.

66. Widengren, J. and P. Schwille, *Characterization of photoinduced and back-isomerization of the cyanine dye by fluorescence correlation spectroscopy.* J. Phys. Chem. A, 2000. **104**: p. 6416-6428.

67. Widengren, J., U. Mets, and R. Rigler, F*luorescence correlation spectroscopy of triplet states in solution: A theoretical and experimental study.* J. Phys. Chem., 1995. 9**9:** p. 13368-13379.

68. Schwille, P., F*luorescence correlation spectroscopy and its potential for intracellular applications.* Cell Biochem. Biophys., 2001. 3**4**(3): p. 383-408.

69. Elson, E.L., F*luorescence correlation spectroscopy measures molecular transport in cells.* Traffic, 2001. 2(11): p. 789-796.

70. Levin, M.K. and J.H. Carson, F*luorescence correlation spectroscopy and quantitative cell biology.* Differentiation, 2004. 7**2**(1): p. 1-10.

71. Briddon, S.J., et al., Q*uantitative analysis of the formation and diffusion of A1-adenosine receptor-antagonist complexes in single living cells.* Proc. Natl. Acad. Sci. USA, 2004. **101**(13): p. 4673-4678.

72. Vaiana, A.C., et al., F*luorescence Quenching of Dyes by Tryptophan: Interactions at Atomic Detail from Combination of Experiment and Computer Simulation.* J. Am.Chem. Soc., 2003. 1**25:** p. 14564-14572.

73. Knemeyer, J.P., N. Marme, and M. Sauer, P*robes for detection of specific DNA sequences at the single-molecule level.* Anal. Chem., 2000. 7**2**(16): p. 3717-3724.

74. Vaiana, A.C., et al., F*luorescence quenching of dyes by tryptophan: Interactions at atomic detail from combination of experiment and computer simulation.* J. Am. Chem. Soc., 2003. 1**25**(47): p. 14564-14572.

75. Heyduk, T., M*easuring protein conformational changes by FRET/LRET.* Curr. Opin. Biotechn., 2002. 1**3**(4): p. 292-296.

76. Xie, X.S., S*ingle-molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuation and enzymatic dynamics.* J. Chem. Phys., 2002. 1**17**(24): p. 11024-11032.

77. Zhuang, X.W. and M. Rief, S*ingle-molecule folding.* Curr. Opin. Struct. Biol., 2003. 1**3**(1): p. 88-97.

78. Haran, G., S*ingle-molecule fluorescence spectroscopy of biomolecular folding.* J. Phys.Conden. Matt., 2003. 1**5**(32): p. R1291-R1317.

79. Onoa, B. and I. Tinoco, R*NA folding and unfolding.* Curr. Opin. Struct. Biol., 2004. 1**4**(3): p. 374-379.

80. Schuler, B., E.A. Lipman, and W.A. Eaton, P*robing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy* Nature, 2002. 4**19**: p. 743.

81. Berglund, A.J., A.C. Doherty, and H. Mabuchi, P*hoton statistics of fluorescence resonance energy transfer.* Phys. Rev. Lett., 2002. 8**9:** p. 068101.

Chapter 3

Folding and Unfolding Kinetics of DNA Hairpins

This chapter describes our 2bFFS approach for the kinetics study of the DNA hairpin folding and unfolding reaction as model system. Dynamic equilibrium between the folded and unfolded conformations of single stranded DNA hairpin molecules containing polythymine hairpin loops was investigated using simultaneous two-beam fluorescence cross-correlation spectroscopy and single beam autocorrelation spectroscopy. The hairpins were end-labeled with a fluorescent dye and a quencher, such that folding and unfolding of the DNA hairpin primary structure caused the dye fluorescence to fluctuate on the same characteristic time scale as the folding and unfolding reaction. These fluctuations were observed as the molecules flowed sequentially between two spatially offset, microscopic detection volumes. Cross-correlation analysis of fluorescence from the two detection volumes revealed the

translational diffusion and flow properties of the hairpins, as well as the average molecular occupancy of the two volumes. Autocorrelation analysis of the fluorescence from the individual detection volumes revealed the kinetics of hairpin folding and unfolding, with the parameters relating to diffusion, flow, and molecular occupancy constrained to the values determined from the cross-correlation analysis. This allowed unambiguous characterization of the folding and unfolding kinetics, without the need to determine the hydrodynamic properties by analyzing a separate control sample. The analysis revealed nonexponential relaxation kinetics and DNA size-dependent folding times characteristic of dynamic heterogeneity in the DNA hairpin-forming mechanism.

3.1 INTRODUCTION

As described in previous chapter, conformational fluctuations of biomolecules (i.e. folding and unfolding) play a critical role in their functioning. Due to the complexities in their biological behavior, a non-invasive and unperturbed approach should have many advantages for the study of biomolecular dynamics. Laser induced temperature jump spectroscopy and fluorescence recovery after photobleaching (FRAP) technique are two of the perturbation methods.[1-4] However, any perturbation to the complicated system can leads unpredictable condition which can not be monitored or controlled easily. And the interpretation of the experimental results becomes difficult or arbitrary. Fluorescence correlation spectroscopy (FCS) is proven to be an efficient technique for characterizing many types of molecular processes non-invasively, under conditions of thermodynamic equilibrium.[5-7] It analyzes the spontaneous fluctuations in the fluorescence signal

arising from a microscopic subvolume of a fluorescent sample. The measurements are carried out *in situ*. Macroscopic perturbation of the system equilibrium to establish a start time for the reaction is not required. Difficulties with FCS analysis arise, however, when multiple processes are occurring simultaneously in the same system, especially when the relaxation times of the different processes occur on similar time scales. This is often the case when studying the conformational fluctuations of freely diffusing biomolecules. The resulting correlation function thus contains fluctuation contributions characterized by a diffusional relaxation time, τ_D, and a relaxation time for conformational fluctuations, τ_R. The usual way of decoupling the different relaxation times is to design a control molecule that gives an independent measure of τ_D, which then constrains the correlation function for the molecule of interest, thereby allowing the determination of τ_R.[8, 9] There are several disadvantages to this approach, including the ever-present possibility of false assumptions about the nature of the control molecule. Furthermore, if the ultimate goal is to perform these types of measurements in the cellular environment, it would be much more desirable to extract all the relevant information about the dynamic processes of the molecule of interest from a single set of fluorescence intensity measurements obtained from the sample of interest alone. With this in mind, we developed a new approach: an FCS-based analysis strategy for analyzing biomolecule folding and unfolding kinetics. The approach is to detect fluorescence from molecules as they flow sequentially between two spatially offset laser probe volumes. The fluorescence from each probe volume is imaged onto separate detectors, and the detected photocounts are analyzed simultaneously from three different vantage points: cross-correlation analysis of the two detection channels relative to each other, independent autocorrelation analysis of the two

detection channels, and photon counting histogram (PCH) analysis.[10-12] Cross-correlation analysis of fluorescence from the two detection volumes reveals the transport properties of the molecules, as well as the average molecular occupancy of the two volumes. Autocorrelation analysis of the fluorescence from the individual detection volumes then reveals the relaxation rate of the fluctuations. Finally, the photon counting histogram analysis, which distinguishes molecular species by differences of their molecular brightness, reveals the equilibrium distribution of the different conformers. This multi-parameter approach allows all the relevant dynamics information to be extracted from a single set of fluorescence intensity fluctuation measurements, eliminating the need for separate measurements on control molecules. For the proof of principle, we have investigated the dynamics of DNA hairpin conformational fluctuations represented as in Figure 3.1. (J. Phys. Chem. 2005, 109, 3648).

3.2 EXPERIMENTAL SECTION

3.2.1 Sample Preparation

HPLC purified fluorescently labeled DNA samples were purchased from Qiagen (Alameda, CA). Two different DNA hairpin samples were analyzed. Each hairpin consisted of a five base-pair stem, with the complimentary sequences 5'-AACCC and GGGTT-3', and a loop containing 21 and 30 deoxythymine residues, respectively. The hairpins are identified as hp-T_{21} and hp-T_{30}. They were dual-labeled with Rhodamine 6G (R6G) and 4-{[4-(dimethylamino) phenyl] axo} benzoic acid (dabcyl) at the 5' and 3' ends, respectively.(Figure 3.1.) The sequence of hp-T_{21} and hp-T_{30} were, thus,

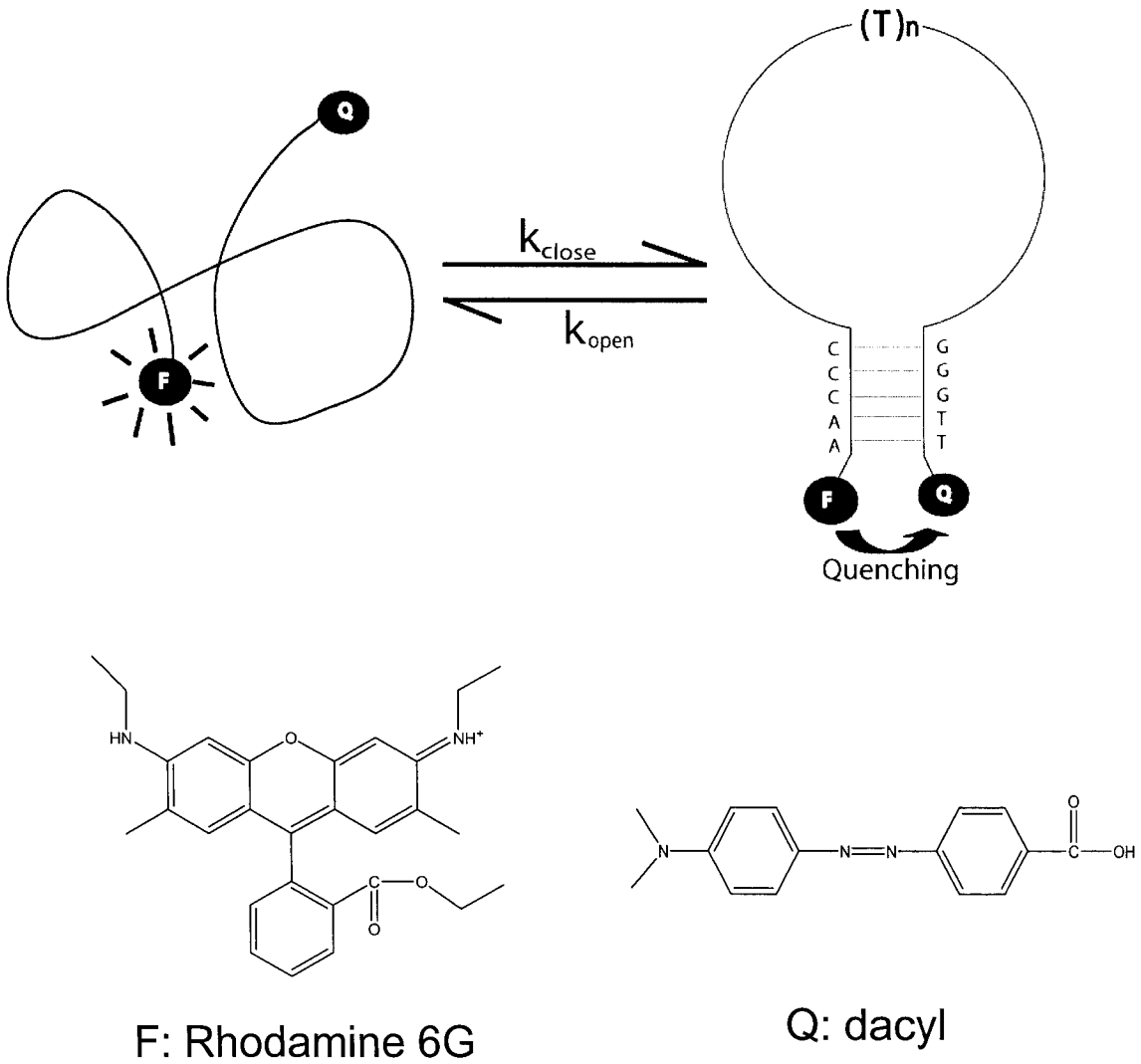

Figure 3.1 The schematic description for the conformational fluctuation of a stem-loop DNA hairpin (hp-T_n) between closed and open states with the characteristic opening rates, k_{op} and closing rate, k_{cl}. F and Q are the fluorophor (R6G) and its quencher (dabcyl), labeled at the each ends of the hairpin stem, respectively. The hairpin fluoresces in the open state, while the fluorescence is quenched by the quencher in the close state.

5'-R6G-AACCC-$(T)_{21}$-GGGTT-dabcyl-3' and 5'-R6G-AACCC-$(T)_{30}$-GGGTT-dabcyl-3', respectively. Dabcyl serves as an efficient quencher of R6G fluorescence when the DNA hairpins are in their folded conformations. We also examined a poly$(dT)_{40}$ oligonucleotide labeled at the 5' end with R6G for comparison purposes. ~5-10-nM DNA solutions were prepared in a pH ~8.0 buffer solution containing 100-mM NaCl, 2.5-mM Tris-HCl, and 250-μM EDTA. We also found use of poly(vinylpyrrolidone) (0.028 wt% in solution, Mw ~106 g/mol, Sigma, St. Louis, MO) as dynamic capillary coating material help preventing the interaction between DNA and glass wall such as capillary interior.[13]

3.2.2 Temperature Dependent Fluorescence Analysis

The quality of the hairpin samples was confirmed by monitoring the fluorescence intensities at the temperatures varying from 5°C to 80°C, on a steady state flourometer (Model ATF 105, AVIV Instrument, Inc. Lakewood, NJ) (Figure 3.2). The sample was excited at 515 nm and fluorescence signal was collected at 565 nm. The resulting melting curves were analyzed to evaluate the equilibrium constants, according to

$$K(T) = \frac{I(80^{\circ}C) - I(T)}{I(T) - I(5^{\circ}C)} \tag{3.1}$$

where $I(T)$ is the fluorescence intensity maximum at temperature T. In this equation, we assumed the reaction occurs between two states, open and closed hairpin states. The fraction of open state, f(T),can be assessed directly by monitoring the fluorescence $I(T)$:

$$f(T) = \frac{I(T) - I_C}{I_O - I_C} \tag{3.2}$$

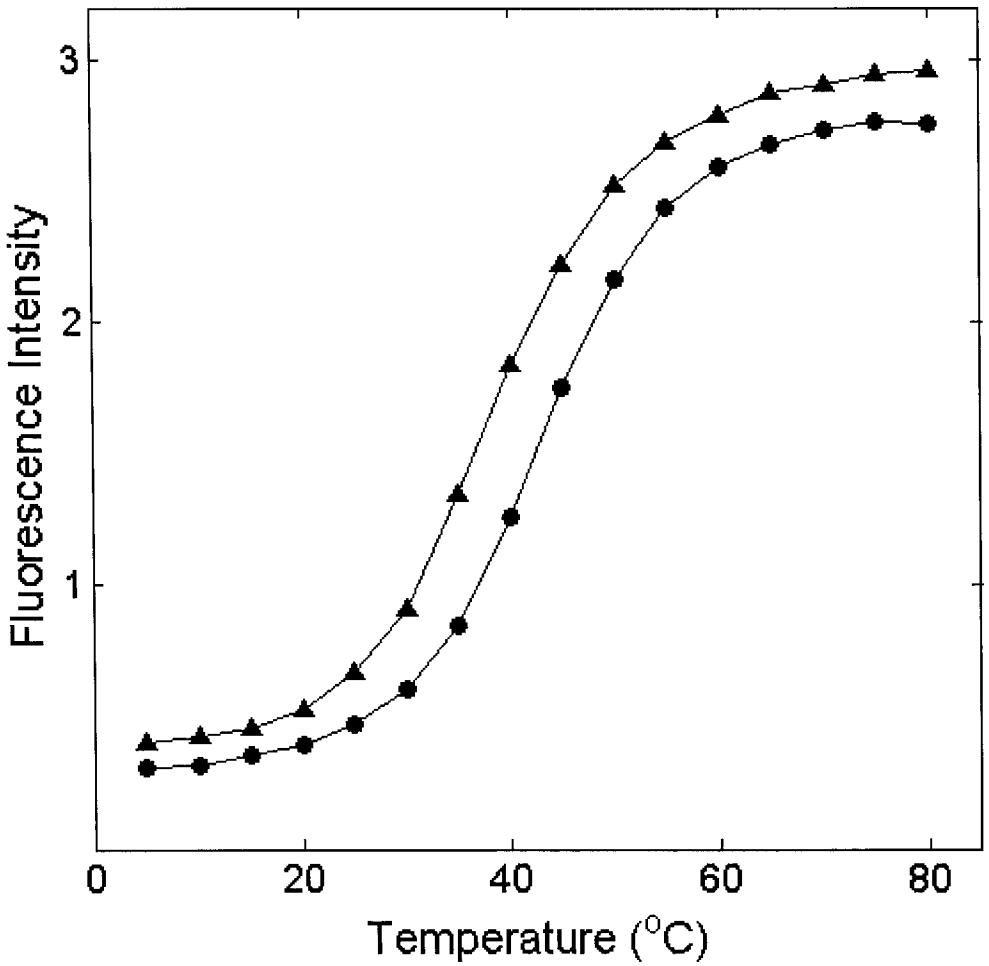

Figure 3.2 Melting profiles of hp-T_{21} (●) and hp-T_{30} (▲) from fluorescence intensity measurements at temperatures varying from 5°C to 80°C .

where I_o is the fluorescence of the open state and I_c is the fluorescence of the closed state. We evaluated *Io* as the fluorescence at 80°C and I_c at 5°C. The equilibrium constant $K(T) = f(T)/[1-f(T)]$ gives the ratio of the chemical rate constants of opening and closing of the hairpin structure.[9]

$$K(T) = \frac{[closed]}{[open]} = \frac{k_{op}(T)}{k_{cl}(T)} \tag{3.3}$$

3.2.3 Instrumentation

The experimental setup was previously described in Chapter 2. (See Figure 2.8) Briefly, the sample solutions were flowed continuously through the capillary by pressurizing the capillary inlet with N_2 gas. A 514.5 nm cw laser beam from an air-cooled, Ar^+ laser (Omnichrome/Melles-Griot, Carlsbad, CA) was used as the excitation source for the R6G labeled DNA molecules. A laser power of ~140-μW per beam was found to offer the best compromise between maximum signal-to-noise ratio, and minimum interference from photo-driven processes (*i.e.,* singlet to triplet intersystem crossing). Fluorescence from each focal region was collected through the same objective and imaged onto the apertures of two 50-μm-core-diameter multimode optical fibers (Thorlabs, Newton, NJ) positioned at the image plane of the objective. The two laser beams formed nearly identical diffraction-limited focal regions, positioned near the center of the inner capillary space (~25-μm from the inner surfaces), separated along the axial dimension of the capillary by a distance, *R*. From autocorrelation analysis of a standard R6G solution, the e^{-2} focal radius in the radial dimension, ω_0, of the focal volume was determined to be 0.306±0.092 μm. The ratio of the radial and axial e^{-2} radii

($\kappa_0=\omega_0/z_0$, where z_0 is the axial radius) was also determined, and found to be 0.104±0.002. The separation distance, R, between the two foci was determined by cross-correlation analysis of the same sample and found to be 2.26 ±0.07-μm. All experiments were carried out at an ambient temperature of 19°C.

3.2.4 Data Analysis

Analysis of the experimental correlation functions was carried out using a least-squares fitting routine from the MatLab Optimization Toolbox (Mathworks, Natick, MA). Different model equations described in following section were used to characterize the correlation information. The residuals from the least-squares fitting routines are calculated to evaluate the fitting quality. The 95% confidence intervals for the fitted parameters were determined using the MatLab Statistics Toolbox.

3.3 ANALYSIS AND RESULTS

3.3.1 Theory

The measured auto- and cross-correlation functions for the DNA hairpin samples contain information about the time-dependent fluorescence intensity fluctuations measured at the two detection volumes (Figure 3.3). These fluctuations are caused by diffusion and unidirectional flow, which alter the number of molecules entering and exiting the probe volumes[9, 14]; folding and unfolding dynamics of the hairpin, which alters the energy transfer efficiency between the dye and the quencher, and, hence, the fluorescence intensity of the dye[8, 9]; and "triplet blinking," which also alters the dye

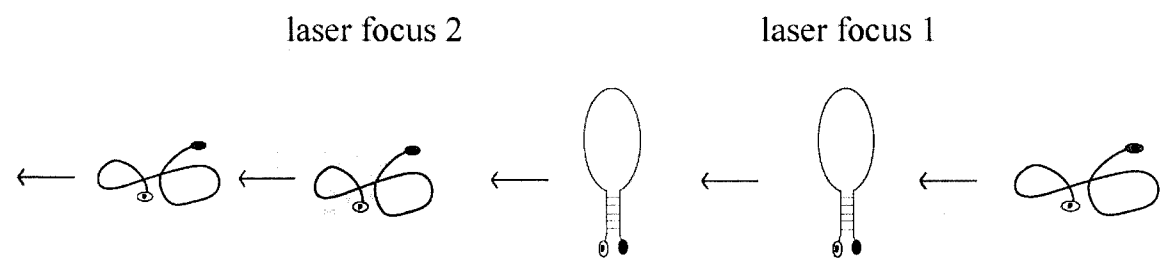

Figure 3.3 Schematic description for the DNA hairpin folding and unfolding reaction inside capillary positioning two laser beam
As flowing along the capillary, the molecule is passing through the laser beams, first laser focus 1, then, laser focus 2. The conformation of hairpin is fluctuating between its open form and closed form. Fluctuations of fluorescence signal in each detector result from the conformational dynamics as well as transport behavior of the molecules.

intensity by placing the dye in a non-radiative triplet state.[15] A derivation of the theoretical forward cross-correlation function, $G_C(\tau)$, that accounts for all these types of intensity fluctuations gives the equation:[16]

$$G_C(\tau)-1=\frac{1}{N}\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2}\left(\frac{1-T+Te^{-\tau/\tau_T}}{1-T}\right)\left(\frac{1-B+Be^{(-\tau/\tau_R)^\beta}}{1-B}\right)\exp\left(-\frac{r^2(1-\tau/\tau_F)^2}{1+\tau/\tau_D}\right)$$
$$=\frac{1}{N}g_D(\tau)g_T(\tau)g_R(\tau)g_{Fc}(\tau) \qquad (3.4)$$

where

$$g_D(\tau)=\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2} \qquad (3.5)$$

$$g_T(\tau)=\frac{1-T+Te^{-\tau/\tau_T}}{1-T}=1+T_{eq}e^{-\tau/\tau_T} \qquad (3.6)$$

$$g_R(\tau)=\frac{1-B+Be^{(-\tau/\tau_R)^\beta}}{1-B}=1+B_{eq}e^{(-\tau/\tau_R)^\beta}\quad, \qquad (3.7)$$

$$\text{and}\quad g_{Fc}(\tau)=\exp\left(-\frac{r^2(1-\tau/\tau_F)}{1+\tau/\tau_D}\right) \qquad (3.8)$$

Here, N is the average number of molecules occupying one of the focal volumes (the occupancy); τ_D is the average transit time of molecules through one of the focal volumes due to translational diffusion; r is the ratio R/ω_0; τ_F is the average transit time for the molecules to flow between the two focal volumes, given by R/V_x, where V_x is the linear flow velocity of the analyte solution; T is the quantum yield, which we have defined $T_{eq}=T/(1-T)$ and τ_T is a time constant for population and depopulation of the non-fluorescent triplet state; B is an amplitude factor that depends on the equilibrium

distribution of the folded and unfolded DNA hairpins, K, and their relative fluorescence intensities, Q, according to[16]

$$B_{eq} = \frac{B}{1-B} = K\left(\frac{1-Q}{1+KQ}\right)^2 \quad , \tag{3.9}$$

where, τ_R is the relaxation time of the folding and unfolding reaction, and β is the stretch parameter.[17] Equation (3.4) assumes the flow axis is parallel to the axis between the two laser beam foci, and that the flow velocity is uniform across the axial and radial dimensions of the focal volume. These conditions are well satisfied when the focal volumes are centered inside the capillary, near the maximum of the parabolic flow velocity profile of the solution, and when slow to moderate flow velocities (millimeters per second to tens of millimeters per second) are used.

$G_C(\tau)$ is a pseudo-Gaussian shaped peak that rises to a maximum near $\tau \approx \tau_F$. This is in contrast to the autocorrelation function which decays from a maximum value at $\tau = 0$. In the cross-correlation function, the hydrodynamic parameters relating to translational diffusion and uniform translation are effectively decoupled, allowing them to be determined independently. Such would not be the case for a single beam autocorrelation analysis. In the cross-correlation function, the peak position is characteristic of τ_F, while the peak dispersion is characterized by the r and τ_D terms. Finally, the peak amplitude is approximately $(1/N)g_D(\tau_F)g_T(\tau_F)g_R(\tau_F)$. The amplitude is referred to as the "degree of correlation". (See Suppl. Info.) It is inversely related to the number of molecules detected by detector 1 that are subsequently detected at detector 2. It decreases with increasing diffusion rate, for example, because the faster the diffusion, the more molecules diffuse out of the beam path before they reach the second detection volume.

Under our experimental conditions, the value of τ_F can be varied between ~30-μs and ~2-ms by adjusting the flow velocity of the analyte solution. For most of the experiments reported here, τ_F was kept between 1- and 2-ms. τ_T, on the other hand, is on the order of only a few microseconds for the R6G fluorophor. Since $\tau_F >> \tau_T$, the $g_T(\tau_F)$ contribution to the cross-correlation amplitude reduces to 1 for all values of τ_F. The triplet blinking rate is so much faster than the transit time between the two laser beams that the molecules have time to blink on and off many times while they flow between beams. Hence, the triplet fluctuations measured at the two detection regions are essentially uncorrelated with respect to each other.

Similarly, under conditions where $\tau_F >> \tau_R$, $g_R(\tau_F)$ makes a negligible contribution to the degree of correlation. This is demonstrated in Figure 3.4, which shows a series of theoretical cross-correlation functions calculated using Equation (3.4). τ_F ranges from ~45-μs to ~1-ms. The set correlation functions depicted as solid bold lines represent control a DNA molecule that has the same diffusion and flow properties as the hairpin DNA, but does not exhibit conformational fluctuations. The only contribution to the cross-correlation peak amplitudes come from $(1/N)g_D(\tau_F)$. The correlation functions depicted as dashed and solid lines represent DNA hairpin molecules with τ_R values of 50- and 150-μs, respectively. Note that when τ_F is greater than ~500-μs, the cross-correlation functions for the control and hairpin DNA samples are virtually indistinguishable. As with triplet blinking, there is effectively no correlation of the conformational fluctuations between the two detection channels under these conditions. At shorter τ_F, enhancement of the cross-correlation amplitudes for the DNA hairpins begins to appear, which becomes more apparent as τ_F decreases. Under these conditions,

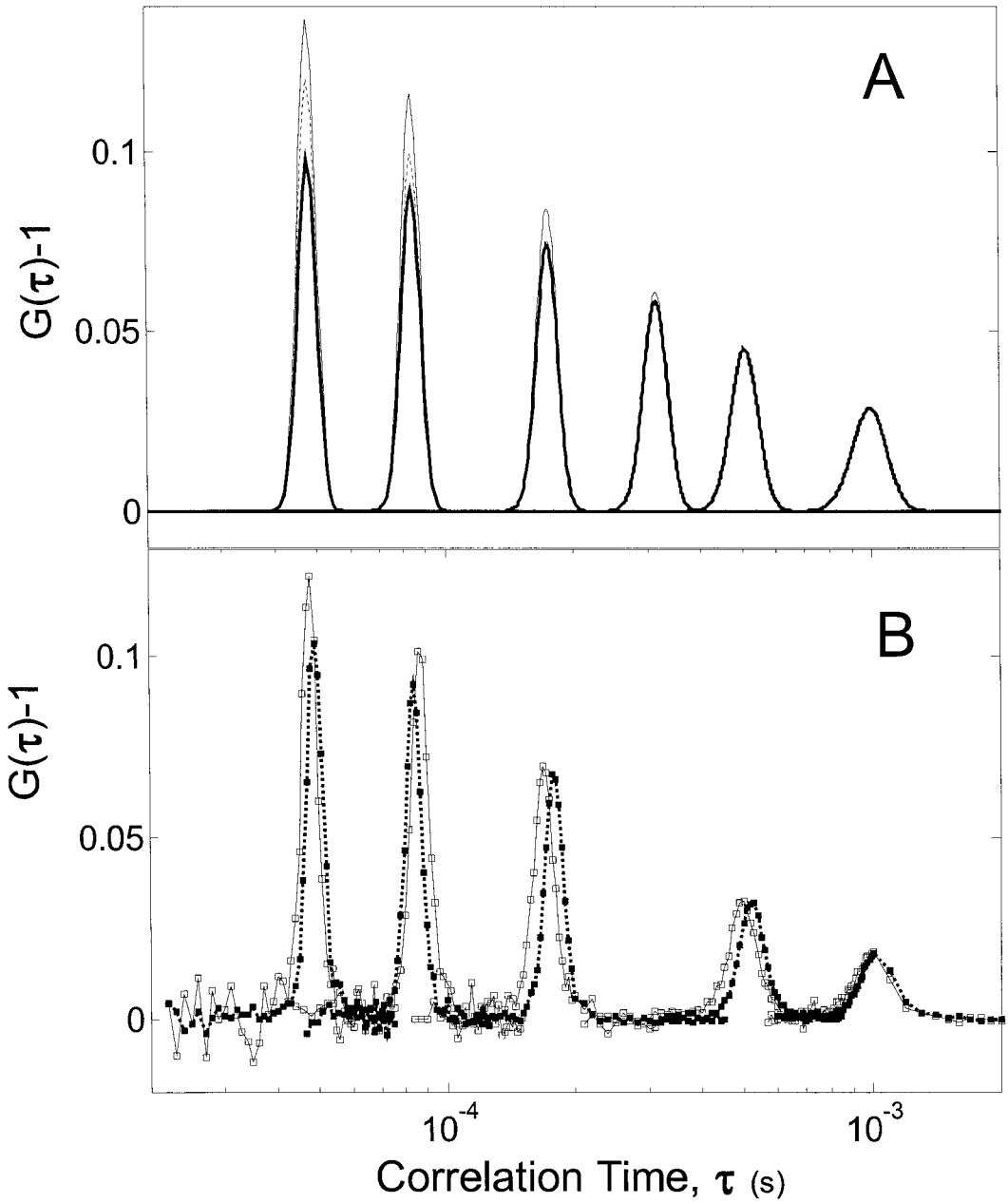

Figure 3.4 (A) A series of theoretical cross-correlation functions calculated using Equation (3.4) with varying τ_F. The solid bold curves represent a DNA sample which does not exhibit conformational fluctuations. The dotted curves and the thin solid curves represent hairpin DNA samples with τ_R= 150 μs and τ_R= 50 μs, respectively. (B) Experimental cross correlation functions measured for poly(dT)$_{40}$ (solid squares connected by dotted curves) and hp-T$_{30}$ (open squares connected by solid curves) samples.

correlation of the conformational fluctuations between detection channels becomes more and more appreciable. These qualitative trends are borne out by the experimental data presented in Figure 3.4B, which compares the cross correlation functions measured for a poly(dT)$_{40}$ control DNA sample with the hp-T$_{30}$ sample over a range of flow velocities. As expected, at transit times above ~500-μs, there is no measurable difference in the degree of correlation of the two samples.

The conclusion drawn from the preceding discussion is that by judicious selection of the flow velocity and/or the distance between focal regions, one can attain conditions where triplet blinking and conformational fluctuations of the DNA hairpins make a negligible contribution to the cross-correlation functions. Hence, the following simplified version of Equation (3.4) can be used to analyze the experimental data:

$$G_C(\tau)-1=a\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\tau/\tau_D}\right)^{1/2}\exp\left(-\frac{r^2(1-\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \tag{3.10}$$

where a is defined as $1/N$. This analysis establishes the values of a, r, τ_D, and τ_F for subsequent analysis of the autocorrelation function obtained from the same set of data. The parameter κ_0 was held fixed to the value determined from autocorrelation analysis of a standard R6G solution in all data analysis procedures. Adjustment of this parameter was found to make a negligible contribution to the values of the other fitted parameters in both the auto- and cross-correlation analyses.

Equation (3.11) represents the theoretical autocorrelation function, $G_A(\tau)$, for molecules undergoing diffusion, flow, triplet blinking, and conformational fluctuations:

$$G_A(\tau)-1=a\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2}\left(\frac{1-T+Te^{-\tau/\tau_T}}{1-T}\right)\left(\frac{1-B+Be^{(-\tau/\tau_R)^\beta}}{1-B}\right)\exp\left(\frac{-(V_x\tau)^2}{\omega_0{}^2(1+\tau/\tau_D)}\right) \tag{3.11}$$

From the definitions of r and τ_F, the substitution $V_x = r\omega_0/\tau_F$ can be made. Equation (3.11) then becomes:

$$G_A(\tau)-1=a\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2}\left(1+T_{eq}e^{-\tau/\tau_T}\right)\left[1+B_{eq}e^{(-\tau/\tau_R)^\beta}\right]\exp\left(-\frac{(r\,\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \tag{3.12}$$

In summary, our analysis consists of measuring a, r, τ_D, and τ_F by fitting the measured cross-correlation functions to Equation (3.10). The autocorrelation functions are then fitted to Equation (3.12) with the aforementioned parameters constrained. In this way, T_{eq}, τ_T, B_{eq}, τ_R, and β are determined by varying these parameters while fitting the autocorrelation function.

3.3.2 Analysis of the Cross-Correlation Functions

Figure 3.5 displays representative auto- and cross-correlation functions obtained for nanomolar solutions of poly(dT)$_{40}$ (Figure 3.5A), hp-T$_{21}$ (Figure 3.5B), and hp-T$_{30}$ (Figure 3.5C). The insets show the experimental cross-correlation functions, along with results of fitting the experimental data to Equation (3.10). The parameters a, r, τ_D, and τ_F obtained from these fits, along with their standard errors, are shown in Table 3.1. All the parameters were well determined from the fitting procedure. The parameter, r, was found to be within experimental error of the ratio R/ω_0 obtained from FCS analysis of a standard R6G solution.

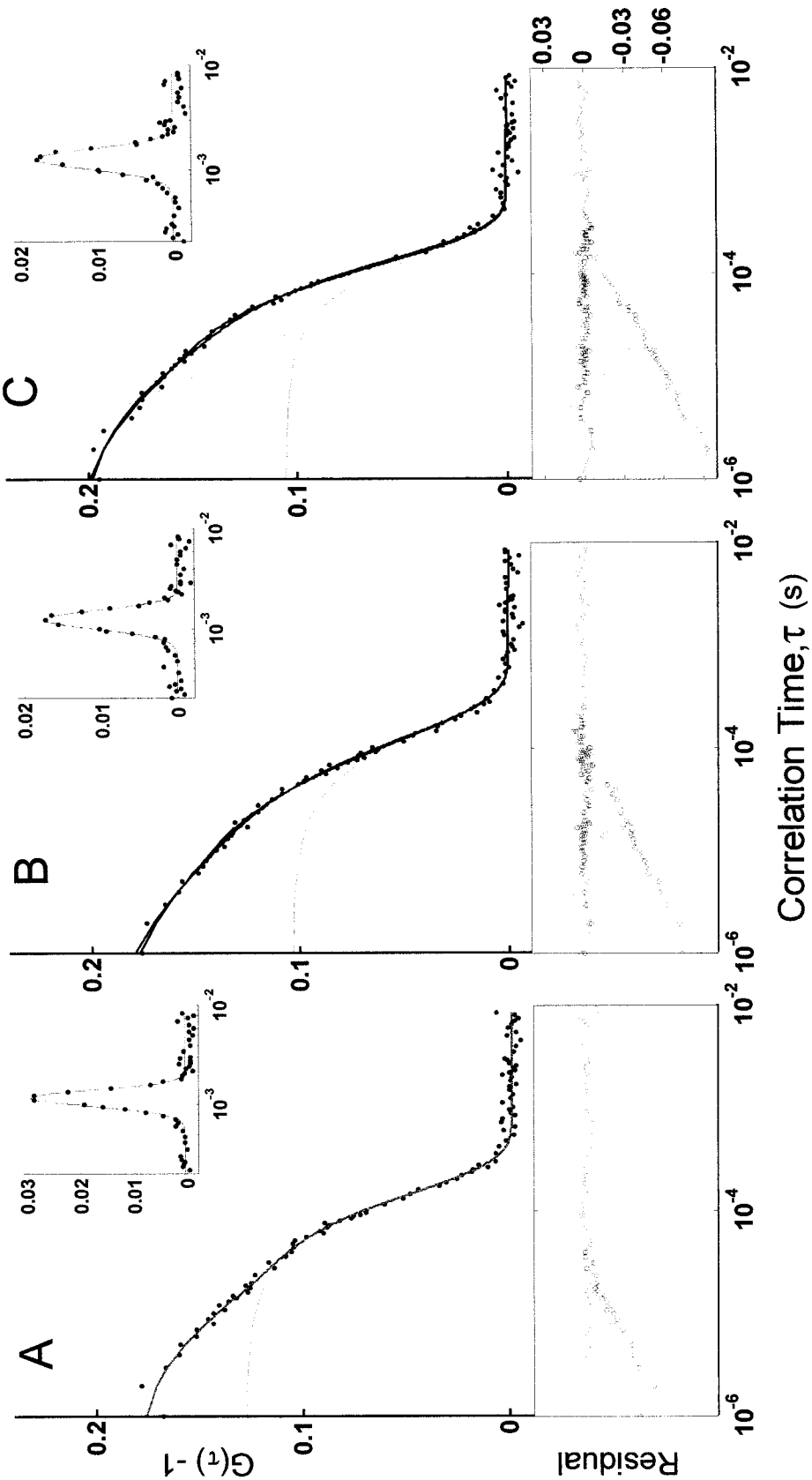

Figure 3.5 The experimental auto-correlation data (black dots) and fitting curves using our models for a poly(dT)$_{40}$ (**A**), a hp-T$_{21}$(**B**), and a hp-T$_{30}$(**C**) DNA sample. (Insets show corresponding cross correlation data and fitting curves.) Red colored curves are calculated autocorrelation functions obtained from the parameters(a, r, τ_d, and τ_F), determined from the cross-correlation analysis. Green colored curves fit the data by a 'triplet model,' which excludes the $g_R(\tau)$ contribution from the Equation (3.12). The experimental data also were analyzed by fitting to Equation (3.12), with a single exponential decay model (blue), and non-exponential decay model (violet). The residuals from the least-squares fitting routines are shown at the bottom. The color scheme for the residual plots is the same as the one used for the autocorrelation functions. The fitted parameters are shown in Table 1.

The τ_D parameters can be used to predict the diffusion coefficients, D, of the DNA molecules using $D = \omega_0^2 / 4\tau_D$. For the poly(dT)$_{40}$ molecule, we obtain $D_{\text{poly(dT)}_{40}}$=6.12(±0.74)×10^{-7} cm^2/s. This can be compared to fluorescence recovery after photobleaching measurements carried out by Nkobo *et al.* to determine the diffusion coefficients of various sized single stranded DNA molecules.[18] According to these measurements, a 40-oligo single stranded DNA molecule is expected to have a diffusion coefficient of ~7.1(±0.6)×10^{-7} cm^2/s under similar solvent conditions. The agreement with our measurement, though imperfect, is still within reason.

For the hairpin DNA molecules, the diffusion coefficients were determined to be $D_{\text{hp-T}_{21}}$=8.83(±1.08)×10^{-7} cm^2/s and $D_{\text{hp-T}_{30}}$=7.08(±0.79)×10^{-7} cm^2/s from the cross-correlation analysis. The diffusion coefficient of hp-T$_{30}$ can be directly compared to that of poly(dT)$_{40}$, since these molecules contain the same number of nucleotide base residues. There is a noticeable (~15-%) enhancement of $D_{\text{hp-T}_{30}}$ relative to $D_{\text{poly(dT)}_{40}}$, which may indicate a greater tendency of the DNA hairpins to exhibit a more compact, rapidly diffusing conformation. Although the statistical relevance of this enhancement is somewhat marginal, it may be taken as an estimate for the difference in diffusion coefficient between the DNA hairpins in their folded and unfolded conformations. It is important to note that Equations (3.4) and (3.12) were derived assuming the diffusion coefficients of the folded and unfolded hairpins were equivalent. However, a ~15-% difference in the diffusion coefficients of the two species is small enough that it will likely constitute only a minor perturbation to the correlation analysis, and will not likely affect the outcome of our measured parameters in any significant way. This issue is currently being investigated in our laboratory.

We can also compare the measured diffusion coefficients of the different sized DNA hairpins. Nkobo *et al.* reported a scaling law for relating the diffusion coefficient of single stranded DNA to the number of nucleotide base residues, M. They found that D scales as $M^{-(0.68\pm0.03)}$.[18] Accordingly, $D_{\text{hp-T}_{21}}/D_{\text{hp-T}_{30}}$ is predicted to be 1.19±0.01, assuming this scaling law holds for hairpin DNA. Our measurements give a ratio of 1.25±0.17, which agrees with the predicted value within experimental error. Hence, the smaller DNA hairpin was found to diffuse at a faster rate than the larger DNA hairpin in a manner consistent with previous studies on the diffusivity of single stranded DNA. In summary, we may conclude that our analysis of the cross-correlation functions successfully provides physically relevant information about the hydrodynamic properties of the molecules under study.

3.3.3 Analysis of the Autocorrelation Functions

We now turn to the analysis of the experimentally determined autocorrelation functions. The red curves overlaying the experimental data in Figure 3.5 are calculated autocorrelation functions based on the parameters a, r, τ_d, and τ_F, that were determined from cross-correlation analysis of the same data. These idealized correlation functions assume no contribution from triplet blinking or conformational dynamics. They were calculated from Equation (3.12), with the parameters T_{eq} and B_{eq} set to zero. Essentially, they represent what the autocorrelation functions would look like for each sample if diffusion and flow were the only processes contributing to the fluorescence fluctuations. The deviation of the experimental measurements from these idealized correlation

functions thus reveal the extent to which other fluctuation processes (*i.e.* triplet blinking and conformational dynamics) are contributing to the actual autocorrelation functions.

For the poly(dT)$_{40}$ sample, the idealized autocorrelation function strongly overlaps the experimental data for lagtimes above ~30-μs. This is further indication that the parameters determined from the cross-correlation analysis well represent the molecular processes occurring at time scales ranging from tens to hundreds of microseconds for this sample. However, at shorter lagtimes, the experimental data deviate from the calculated correlation function, and this deviation becomes increasingly pronounced as the lagtimes decrease. We attribute this deviation to triplet blinking, since poly(dT)$_{40}$ is not expected to undergo any other processes that would significantly alter the fluorescence intensity of the R6G label. We have analyzed the autocorrelation data from the poly(dT)$_{40}$ by fitting to a modified version of Equation (3.12) for which the parameter B_{eq} in $g_R(\tau)$ is set to zero. This form of the autocorrelation function is referred to as the triplet model. As before, a, r, τ_D, and τ_F were constrained to the values determined from the cross-correlation analysis. The fitting of the data was accomplished by varying T_{eq} and τ_T alone. The results are presented in Table 3.1 and shown graphically as the solid green curve in Figure 3.5A. The values thus obtained (τ_T= 9.4 ±1.1-μs, and T_{eq} = 0.587 ±0.004) are larger than expected for the R6G fluorophore by a factor of ~2. Analysis of a standard R6G solution under identical conditions gave rise to corresponding τ_T and T_{eq} values of 4.53 ±0.97 μs and 0.234 ±0.028, respectively. This latter set of values is consistent with previous FCS studies on the triplet intersystem crossing dynamics of R6G in aqueous solution.[19] We, therefore, conclude that the values reported in Table 3.1 for τ_T and T_{eq} of poly(dT)$_{40}$ are the ones best supported by our data. It is possible that

the R6G label is interacting with the single-stranded DNA in a way that somehow stabilizes the triplet state of the fluorophor.

The experimental autocorrelation functions for the hp-T_{21} and hp-T_{30} samples deviate much more strikingly from the idealized ones than does the poly(dT)$_{40}$ autocorrelation function. This can be seen graphically in Figure 3.6. Here, the residuals $[G_{A,\mathrm{calc},i}(\tau_i) - G_{A,\mathrm{exp},i}(\tau_i)$, where $G_{A,\mathrm{calc},i}(\tau_i)$ and $G_{A,\mathrm{exp},i}(\tau_i)$ are the calculated and experimental autocorrelation data points at each $\tau = \tau_i]$ of the idealized and experimental autocorrelation functions for all three DNA samples are plotted. The residuals obtained from fitting the poly(dT)$_{40}$ data to the triplet model are also shown for comparison (grey dots). For the hp-T_{21} sample, the idealized correlation function fits the experimental data reasonably well at lagtimes above ~70-μs; whereas, for hp-T_{30}, a reasonable fit is achieved only after ~100-μs. Clearly there are other fluctuation processes occurring in the DNA hairpin samples which broaden the observed autocorrelation functions from where they would be if diffusion, flow, and triplet blinking were the only contributing factors. Furthermore, the broadening is more pronounced for hp-T_{30} than for hp-T_{21}. To emphasize these points, we attempted to fit the hairpin DNA data to the same triplet model used for fitting the poly(dT)$_{40}$ sample. We constrained τ_T for the DNA hairpins to be within a factor of ~2 of the corresponding value determined for poly(dT)$_{40}$. We then allowed τ_T and T_{eq} to vary within that constraint. The resulting best fits are displayed as solid green curves in Figures 3.5B and 3.5C. The poor quality of these fits shows that the triplet model, by itself, is inadequate to explain the deviation of the experimental autocorrelation functions from the idealized ones for these DNA hairpin samples.

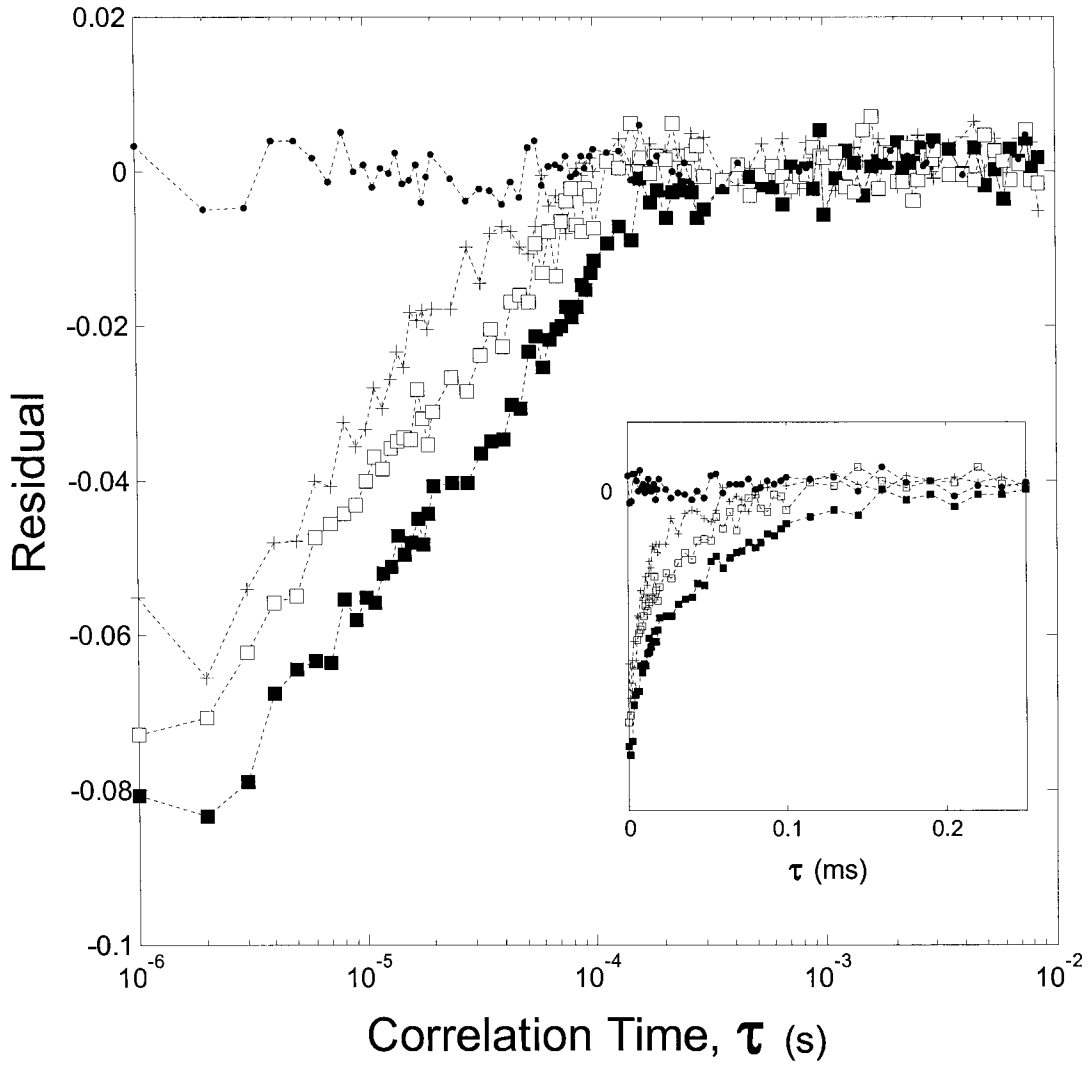

Figure 3.6 Residuals from the analysis in Figure 3.5, comparing the experimental autocorrelation functions with the idealized autocorrelation functions calculated from the cross-correlation parameters (a, r, τ_d, and τ_F). The deviation of hp-T_{21} (open squares), and hp-T_{30} (filled squares) correlation functions are attributed to conformational fluctuations and triplet blinking. The deviation of the poly(dT)$_{40}$ data (plus signs) is attributed to triplet blinking. The residuals obtained from fitting the poly(dT)$_{40}$ data to the triplet model are also shown for comparison (small dots). The inset displays the residual plots on a linear time scale.

The enhanced broadening of the experimental autocorrelation functions for the DNA hairpin samples is attributed to folding and unfolding kinetics of the DNA hairpin structure. To characterize these effects, the experimental autocorrelation functions for the DNA hairpin samples were analyzed by fitting to Equation (3.12), while adjusting the parameters T_{eq}, τ_T, B_{eq}, τ_R, and β. The parameters a, r, τ_D, and τ_F were constrained to the values determined from the cross-correlation analysis. The fitted curves are shown graphically in Figures 3.5B and 3.5C, and the fitted parameters are presented in Table 3.1. Two sets of parameters are presented—one with β constrained to a value of one, and one with β as a free floating parameter. For the latter set of parameters, mean relaxation times, $\langle \tau_R \rangle$, for the folding and unfolding reaction were calculated using

$$\langle \tau_R \rangle = \int_0^\infty \exp\left[-\left(\frac{t}{\tau_R} \right)^\beta \right] d\tau = \left(\frac{\tau_R}{\beta} \right) \Gamma\left(\beta^{-1} \right) , \tag{3.13}$$

where $\Gamma\left(\beta^{-1}\right)$ is the gamma function. All parameters were well determined in both fits. Inclusion of the β parameter resulted in an improved overall precision of the fitted curves to the experimental data as is evident for the lower standard deviations of the nonexponential fitting procedures.

3.4 DISCUSSION

The kinetics of DNA hairpin formation has been studied rather extensively. Some of the most recent experimental studies are those of Libchaber and co-workers,[20, 21] Klenerman and co-workers,[22-24], and Ansari and co-workers.[1-3] Also, a study on the kinetics of end-to-end contact formation in non-hairpin forming single-stranded DNA was recently reported that has strong implications for understanding the mechanism of

Table 1. Parameters from FCS Analysis[a,b]

	poly(dT)$_{40}$	hp-T$_{21}$		hp-T$_{30}$	
Cross Correlation Analysis					
τ_F (ms)	1.311 (12)	1.299 (13)		1.268(13)	
r	7.61 (22)	7.90 (31)		7.77 (39)	
N_{eff} [c]	8.92 (53)	9.49 (57)		9.44 (57)	
τ_d (μs)	383 (31)	266 (16)		332 (22)	
Standard error[d]	6.84 ×10^{-4}	6.92 ×10^{-4}		7.28 ×10^{-4}	
Auto Correaltion Analysis					
		Exponential[§]	Non-exponential[†]	Exponential[§]	Non-Exponential[†]
τ_T (μs)	9.4 (1.1)	3.95 (69)	2.91 (98)	4.94 (78)	3.62 (91)
T_{eq}	0.587 (41)	0.293 (62)	0.306 (92)	0.346 (46)	0.329 (96)
B		0.603 (18)	0.689 (78)	0.754 (30)	0.82 (12)
τ_R (μs)		51.3 (3.1)	45.5 (8.7)	102.7 (9.2)	121 (19)
β			0.72 (14)		0.67 (12)
$<\tau_R>$ (μs)			56 (17)		159 (43)
Standard error[d]	2.54 ×10^{-3}	2.55 ×10^{-3}	2.51 ×10^{-3}	2.57 ×10^{-3}	2.44 ×10^{-3}
Equilbrium Distribution Analysis[e]					
		Exponential[§]	Non-exponential[†]	Exponential[§]	Non-exponential[†]
K_{eq}		28.4 (2.8)		22.5 (2.0)	
Q		0.156 (22)		0.153 (26)	
τ_{op} (ms)		1.49 (17)	1.64 (52)[f]	2.08 (26)	3.73 (97)[f]
τ_{cl} (μs)		53.1 (3.2)	58 (16)[f]	108.0 (9.6)	166 (45)[f]

[a] Parameters are corresponding to the fitting analysis in Figure 5. [b] Values in parentheses are uncertainties (confidence interval) in the last digits. [c] N_{eff} is the effective value of occupancy, and is equal to a^{-1}. [d] Standard errors from each curve fitting are estimated using residual sum-of-squares and the number of data points. [§] Column containing parameters from single exponential fit (β=1 in equation(5)).[†] Column containing parameters from non-exponential fit (equantion(5)). [e] The parameters are calculated for 19°C. [f] Calcualted values using $<\tau_R>$.

DNA hairpin formation.[25] This accumulated research provides us with a reasonably coherent picture of the kinetics and thermodynamics of the DNA hairpin forming process. Some of the key points relevant to the present analysis are summarized as follows. 1) For a given stem sequence, the time constants for hairpin formation increase as a function of the loop size. Hairpins containing a 5-base-pair stem sequence like the ones used in this study exhibit folding times ranging from microseconds to hundreds of microseconds, depending on loop size. This is consistent with our observation of enhanced broadening of the autocorrelation function and the larger measured relaxation time for the hp-T_{30} sample, relative to the hp-T_{21} sample (See Figures 3.5 and 3.6). 2) The folding time depends on the sequence of the hairpin loop, as well as the loop size. Hairpins containing polyadenine loops exhibit longer folding times than equivalent-sized polythymine loops. For example, our measurement shows that the mean relaxation time of hp-T_{30} is ~159-μs. By comparison, the mean relaxation time of a hp-A_{30} hairpin with an equivalent stem sequence has been measured to be ~450-μs under similar temperature and solvent conditions.[22] These much longer folding times are thought to be due, in part, to intramolecular base-stacking interactions in the polyadenine chain. Such interactions are, for the most part, absent in polythymine. 3) DNA hairpin folding exhibits non-Arrhenius kinetics consistent with a heterogenous folding mechanism. The unfolded hairpins must sample a large number of local minima consisting of mismatched base contacts before nucleation of the folded hairpin can proceed. One way this manifests itself in DNA hairpins containing polyadenine loops is by nonexponential relaxation of the folding and unfolding reaction at temperatures below the melting temperature of the stem. A stretched exponential model is needed to characterize the relaxation process under these

conditions.[22-24] Whether similar behavior is observed for hairpins containing polythymine loops is a question that is addressed in the present study.

Our data supports the expectation that DNA hairpin formation occurs via a heterogeneous reaction pathway, in that our analysis gives rise to well determined β parameters, consistent with stretched exponential relaxation, analogous to the observations of Klenerman and co-workers.[22-24] In our experiments, determination of β depended on our ability to constrain the correlation amplitude, a, in addition to the parameters relating to diffusion and flow. When only the diffusion and flow parameters were constrained, the β parameter was undefined.(See Supple. Info..) Hence, procedures that rely on the prior analysis of a control molecule to constrain only the hydrodynamic properties of the molecules, and not the concentration dependent terms, would likely not succeed at determining this important parameter. As described in the previous chapter, a possible interpretations of the β parameter is that $0 \leq \beta \leq 1$, with $\beta \approx 0$ corresponding to a completely random folding mechanism, and $\beta \approx 1$ corresponding to two-state, single barrier folding. Intermediate β values can indicate varying degrees of heterogeneity in the folding mechanism. Our β values show an interesting trend when compared to the β parameters reported by Klenerman and co-workers for DNA hairpins containing polyadenine loops. For the polythymine loops, we find $\beta_{\text{polyT}} \approx 0.7$; whereas, $\beta_{\text{polyA}} \approx 0.5$ has been reported for the same sized polyadenine loops under similar solvent and temperature conditions.[22-24] Hence, the heterogeneity of the folding mechanism, though still present, appears to be less of a factor for the polythymine containing hairpins than for the polyadenine containing hairpins. Ansari and co-workers have argued that the longer folding times of polyadenine containing hairpins are a consequence of enhanced

base-stacking of the polyadenine chain. These interactions serve to increase the roughness of the folding free energy surface by making it more likely for the unfolded hairpin to form mismatched base contacts. For polythymine, these intramolecular interactions are largely absent, giving rise to a smoother free energy landscape. Our observation that β_{polyT} is closer to the limiting case of two-state kinetics than β_{polyA} is consistent with this argument. Finally, a comparison of the β values for the different sized polythymine containing DNA hairpins shows that $\beta_{\text{hp-T}_{21}}$ is ~10-% larger than $\beta_{\text{hp-T}_{30}}$. This may point toward a trend of increasing heterogeneity in the folding mechanism with increasing loop size.

Determination of the folding and unfolding time constants, τ_{cl} and τ_{op}, from the mean relaxation times of the two DNA hairpins requires knowledge of the equilibrium constant, K. Given K, the individual time constants are calculated using

$$\tau_{op} = \langle \tau_R \rangle (K+1) \qquad \tau_{cl} = \frac{\langle \tau_R \rangle (K+1)}{K} \tag{3.14}$$

The correlation analysis alone is not sufficient to determine the equilibrium distribution of folded and unfolded hairpins. The amplitude term, B_{eq}, depends on both the equilibrium distribution and the relative fluorescence intensities of the DNA hairpins in their folded and unfolded states. In the Chapter 4, the work investigating whether the photon counting histogram can be used to resolve these parameters is described. The common measurement of the K is analysis of DNA melting curves. The studies previously reported have been used the melting curve information with assumption of two-state mechanism to describe the folding and unfolding processes. This assumption might be the reason of many inconsistent answers discussed in Supplementary section,

such as loop dependence of the reaction rates and activation energy of the reaction. (see Section 3.6.3) These fundamental questions are still unresolved even after 30 years of DNA hairpin dynamics studies.

3.5 CONCLUSIONS

In summary, we have shown that it is possible to extract relevant information about the hydrodynamic and the reaction rate properties of DNA hairpin molecules undergoing simultaneous translational motion and intramolecular folding and unfolding. This was accomplished by analyzing fluorescence fluctuations observed at two spatially offset detection volumes. Cross-correlation analysis of the two detection volumes revealed information about the hydrodynamic properties, independent of the intramolecular dynamics. The autocorrelation functions of the independent detection volumes could then be constrained to allow determination of the intramolecular dynamics properties. Our measurements compliment previous studies on DNA hairpin folding and unfolding kinetics in that they present evidence for heterogeneity in the folding mechanism in large polythymine containing hairpin loops. This comes from the observation of nonexponential relaxation kinetics and from a larger than expected loop size dependence on the folding time at our laboratory temperature. Similar approach including analysis of the photon counting histogram to measure the equilibrium distributions of the folded and unfolded molecules is discussed in chapter 4. It might be quite interesting to extend the temporal resolution of our experiment into the nanoseconds to tens of nanoseconds time range. Finally, a major longer term goal of these efforts is to study biomolecule folding and unfolding reactions in the cellular environment. In all

these cases, the key to success is to properly constrain the system so the desired kinetics information can be uncovered.

3.6 SUPPLEMENTARY INFORMATION

3.6.1 Precision of kinetic parameters

In this study, the autocorrelation analysis of the fluorescence data revealed the kinetics of hairpin folding and unfolding, with the parameters relating to diffusion, flow, and molecular occupancy constrained to the values determined from the cross-correlation analysis of same data set. We examined the effect of the constraints in the evaluation of kinetic parameters using several fitting methods. In Figure 3.7 the precision of the fitting parameters such as τ_R and β of hp-T_{21} DNA sample are shown together with their 95% confidential intervals. Fitting analysis without any constraint, in other words, simple fitting of the autocorrelation function to the Equation (3.12), can not determine any of the kinetic parameters. The too wide range of confidence intervals makes it arbitral and lack of the corresponding physical meaning. In addition to this, it is shown that maximum precision can be achieved only from the fitting analysis with constraints of both τ_D and the correlation amplitude, a. It demonstrates the critical need of the information on the sample concentration which is inversely related the parameter, a, in the FCS study of biomolecule dynamics. Therefore, the combined approach to FCS analysis of the data set under the same experimental condition and same time would be one of the best hopes to the study of more complex system.

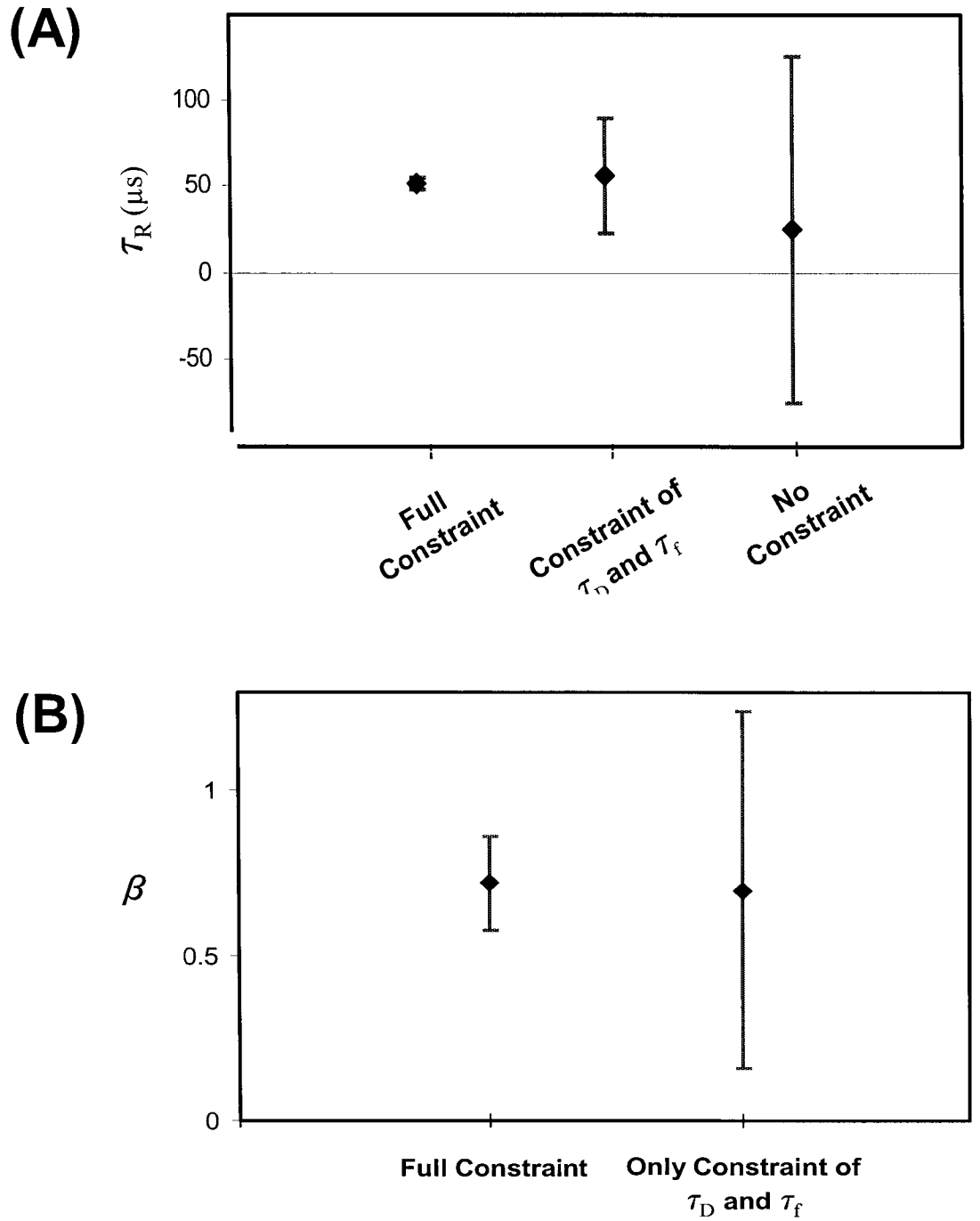

Figure 3.7 The precision of kinetic parameters ((A) τ_R, (B) β) of hp-T_{21} DNA sample from autocorrelation fitting. Error bars represent 95% confidential intervals of the values.

3.6.2 Degree of Correlation

Originally, our approach for studying the kinetics of biomolecule conformational fluctuations started from the application of dual beam fluorescence cross-correlation spectroscopy to the flowing samples. We have analyzed these fluctuations using only cross-correlation data from various flow rates. The experimental apparatus used for this study was similar to the one illustrated in this chapter, except that series of experiments performed under various flow rates controlled by changing gas pressure (Nitrogen gas). It is noted that cross-correlation of the fluorescence from two spatially offset probe volumes results in a pseudo-Gaussian shaped peak. The peak position corresponds to the transit time of the molecules between the two probe volumes, and is a function of the unidirectional molecular flow velocity. The peak width is determined by diffusion in all three dimensions.

We define a parameter referred to as the "*degree of correlation*," which is given by the intensity of the cross-correlation peak. This parameter relates to the number of molecules detected as they pass sequentially through each of the two probe volumes. For example, molecules that pass through the first probe volume, but then diffuse out of the beam path before reaching the second, do not contribute to the degree of correlation. In the absence of quenching and unquenching processes, the degree of correlation relates solely to the diffusion of molecules out of the beam path. Note that the degree of correlation depends on the flow velocity of the solution. When the flow velocity is large compared to the diffusion rate, the degree of correlation will also be large, because molecules detected at the first probe volume will have less time to diffuse out of the beam path before reaching the second probe volume. As the flow velocity decreases, more and

more molecules will diffuse out of the beam path, giving rise to a lower degree of correlation. This can be seen in Figure 3.8B, which presents a series of cross-correlation functions measured over a range of flow velocities for a solution containing single R6G labeled DNA molecules (no quenching/unquenching).

Now if we consider DNA hairpin molecules undergoing conformational fluctuations that give rise to quenching and unquenching of the R6G fluorescence, we expect the conformational dynamics to affect the degree of correlation, as well as diffusion. DNA hairpin molecules adopting a closed conformation will be invisible to the apparatus due to efficient quenching of the R6G fluorescence. Only molecules passing through both laser beams in an open conformation will contribute to the degree of correlation. If, for example, a molecule passes through the first laser beam in an open conformation, and then transforms to a closed conformation by the time it reaches the second laser beam, it will not contribute to the degree of correlation. In general, the degree of correlation will be determined by the relationship between the flow velocity, the diffusion rate, and the relaxation rate for the opening and closing transition. At faster flow speeds, the molecules will have less time to undergo conformational transitions during their transit time between the two laser beams, giving rise to a greater degree of correlation than at slower flow speeds. Figure 3.8A presents a series of cross-correlation peaks measured from a solution containing R6G and dabcyl labeled hairpin DNA at varying flow velocities. Note the much sharper decay in the degree of correlation with decreasing flow velocity, as compared to the data in Figure 3.8B. This is due to the combined effect of diffusion and conformational transitions during the transit times between the two laser beams.

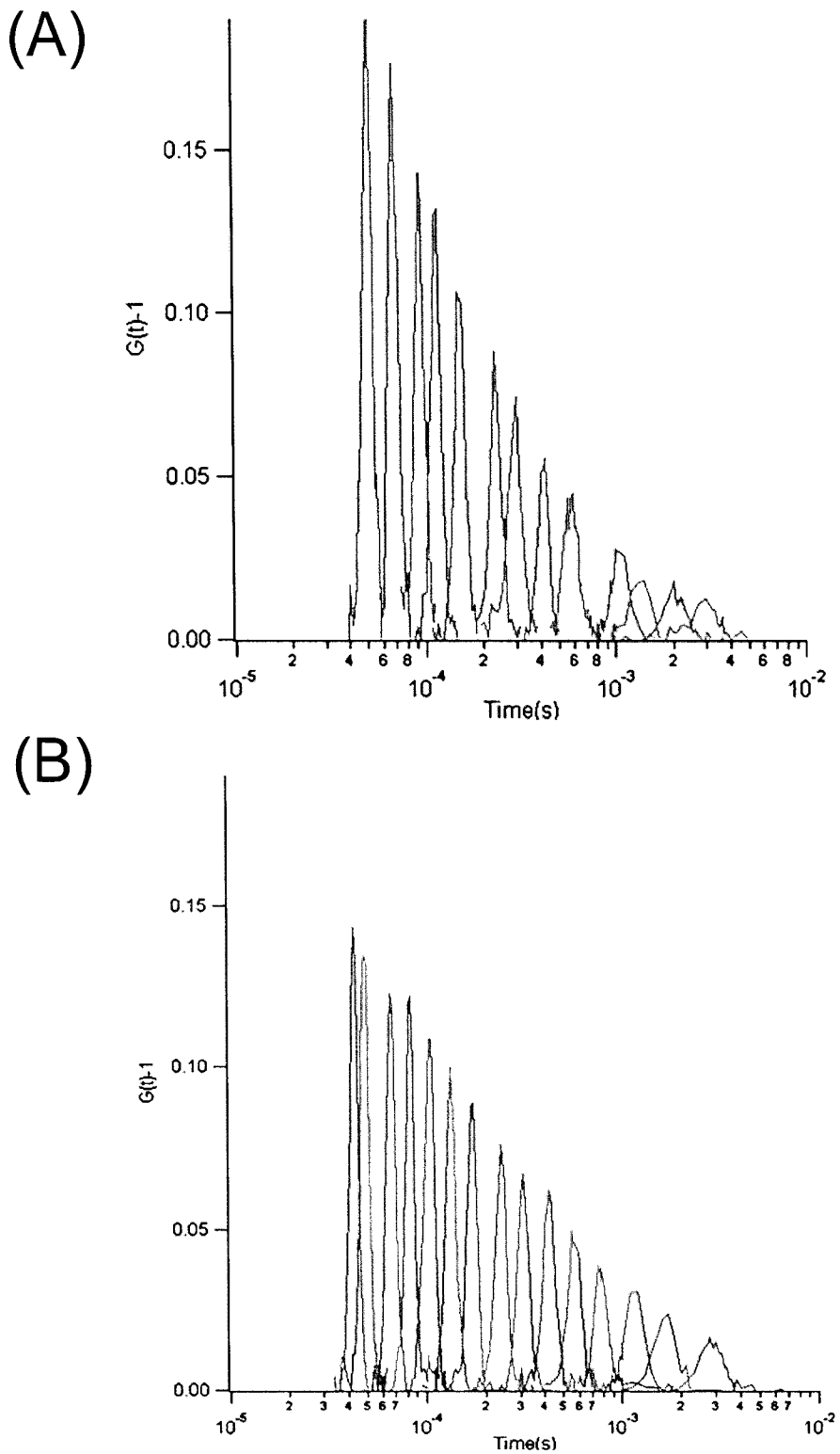

Figure 3.8 Fluorescence cross-correlation functions vs flow velocity for (A) ~10 nM R6g-5'-AACCC-$(T)_{30}$-GGGTT-3'-dabcyl in 100 mM NaCl and (B) ~10 nM R6G labeled T_{40} in 100 mM NaCl.

The data in Figure 3.8A can be analyzed by fitting to the following equation which comes from Equation (3.4):

$$G_C(\tau)-1=\frac{1}{N}\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\,\tau/\tau_D}\right)^{1/2}\left(1+B_{eq}e^{-\tau/\tau_R}\right)\exp\left(-\frac{r^2(1-\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \qquad (3.15)$$
$$=\frac{1}{N}g_D(\tau)g_R(\tau)g_{Fc}(\tau)$$

Here, we applied various gas pressures to achieve a wide range of flow rates. In the data of figure 3.8, τ_F has time scale from ~40 μs to ~40 ms. Therefore, triplet blinking is excluded from the model equation because the process has too fast time scale such that the g_T term goes to 1. The faster process loses its contribution in the cross correlation curve. An advantage of this method, compared to single beam autocorrelation techniques reported previously, is that the relaxation time contribution to the degree of correlation can be decoupled from the diffusion contribution. In single beam autocorrelation analysis, the diffusion term must be determined independently before the relaxation time can be measured. In our experiment, the diffusion time constant, τ_D, is obtained from the width of the cross-correlation peak. The contribution of τ_D to the peak amplitude can be fixed to the value obtained by analyzing the peak width, allowing an independent determination of τ_R. The possible application of this proof-of-concept demonstration to the analysis of biomolecule conformational dynamics is under investigation. A simulation study for the reaction is shown in Figure 3.9. If the molecules are designed to express fluorescence fluctuation associated with target reaction, same analysis can reveal the characteristics of the dynamics in general. The current challenge in the technique is non-homogenous flow profile at higher flow rates. We found that even at the center of square capillary setup, the laser detection volumes experience various flow rates, so that the analysis of τ_D needs to

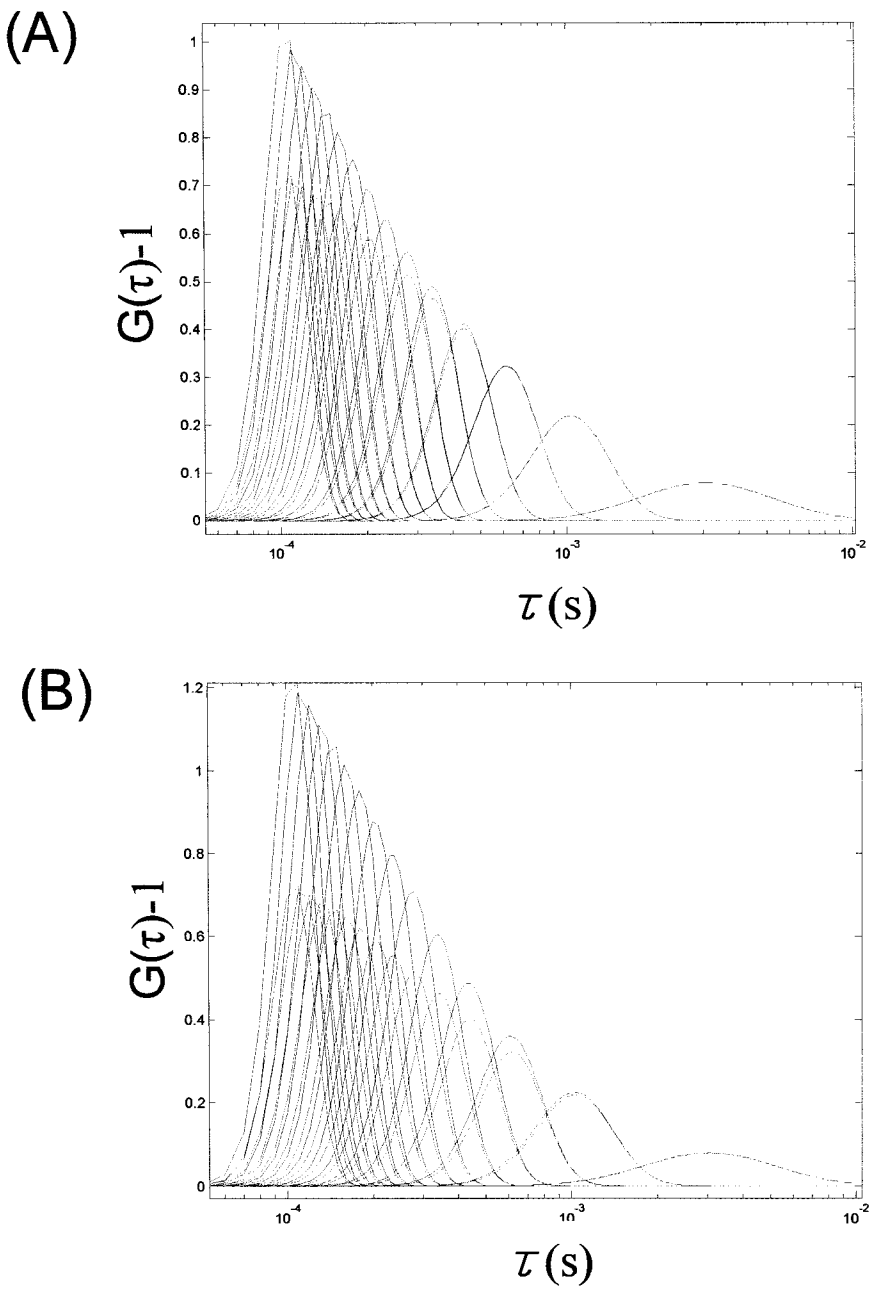

Figure 3.9 Simulated fluorescence cross-correlation functions vs. flow velocity for molecules experiencing conformational fluctuation similar to the reaction of DNA hairpins (red curves). The reaction rates are set to (A) k_1=8000 s^{-1} and k_2=400 s^{-1} (B) k_1=3000 s^{-1} and k_2=400 s^{-1}, respectively. Control cases which do not have the reaction are represented as blue curves for comparison. The peak amplitudes are defined as "degree of correlation".

be corrected by the contribution. Although the modified model equation can be derived from mathematical view of points, its complexities will make the analysis more difficult. One approach to overcome this challenge can be a scanning FCS (SFCS) technique. Rather than pressurizing the solution along the capillary, which causes parabolic flow profiles, sample stage is moving relative to the detection volumes. The two beam SFCS has great potential for the study inside the cell environments where no active flow exists.

3.6.3 Assumption of Two-state Mechanism

In this section, the opening and closing reaction rates are evaluated for the comparison purpose. We measured the K values from the melting curves displayed in Figure 3.2. The equilibrium constants, along with the time constants for folding and unfolding are shown in Table 3.1. Also shown is the fluorescence intensity ratio, Q, for each DNA hairpin, which was extracted from the fitting parameter B_{eq}. It should be noted that the K values from the melting curves, K_{melt} is not equal to K_{eq} which is actual equilibrium constant describing the reaction investigated during the FFS measurement. Our studies in Chapter 4 indicate the significant misleading in the case of using melting curve information for the determination of reaction rates. However, by using the K_{melt}, we compare our data with previous reports which were determined under assumption of two state mechanism for the DNA hairpin formation. The folding and unfolding time constants observed for the two DNA hairpins are in qualitative agreement with FCS measurements reported by Libchaber and co-workers for DNA hairpins with equivalent stem sequences and identical loop sequences.[20, 21] Comparison of the time constants observed for the different hairpins can be used to characterize the folding and unfolding

time dependence on hairpin size at our laboratory temperature. These dependencies are characterized by the quantities $L^{\alpha^{cl}}$ and $L^{\alpha^{op}}$, where L is proportional to the number of nucleotide bases in the hairpin loop, and α^{cl} and α^{op} are the folding and unfolding exponents, respectively. Our measurements show $\alpha^{cl} = 2.96 \pm 0.98$ and $\alpha^{op} = 2.30 \pm 0.93$ for the two hairpin sizes studied. The large errors could be improved by studying a broader range of hairpin sizes, and increasing the number of replicate measurements.

The size dependencies of the folding and unfolding times have been measured for similar DNA hairpins by Libchaber and co-workers[20, 21] and Ansari and co-workers.[1, 3] The Libchaber group used conventional FCS analysis to study the kinetics of DNA hairpin formation under the same solvent conditions as reported here. The hairpin loops consisted of various sized polyadenine and polythymine chains, including polythymine loops equivalent in size to the ones used in our study. Libchaber and co-workers measured the folding and unfolding kinetics over a range of temperatures, from well below to well above the melting temperature of each hairpin. They observed a folding time dependence on hairpin loop size given by $\alpha^{cl} = 2.6 \pm 0.3$, in reasonable agreement with our measurement. However, their analysis differed from ours in that they assumed single-exponential, two-state kinetics for the reaction. Our somewhat larger value may come from our analysis of hairpin folding as a nonexponential process. The Ansari group used laser-induced temperature jump spectroscopy to study the kinetics of three DNA hairpins containing equivalent stem sequences and small polyadenine and polythymine loops of 4, 8, and 12 base residues, for which $\alpha^{cl} \approx 2.2$. These measurements were restricted to elevated temperatures near the melting temperature of the hairpins. The larger folding exponent obtained by Libchaber and co-workers, and by us, is thought to

arise from the increased propensity of DNA hairpins to form misfolded intermediates at temperatures below the melting temperature.

The dependence of the unfolding time on loop size has been somewhat more difficult to characterize, and there remains some controversy surrounding this issue. In an earlier study, Libchaber and co-workers reported unfolding rate parameters that showed a qualitatively similar loop size dependence to the one reported here at our laboratory temperature.[20] However, a later study by the same authors showed no loop size or temperature dependence on the unfolding time.[21] Ansari and co-workers, in their study of smaller hairpins at elevated temperatures, found an enhanced stability of the smaller hairpins that manifested itself as a *decrease* in the unfolding time with hairpin size.[1] For polythymine containing hairpins, they observed $\alpha^{op} \approx -2.3$. Here again, this apparent conflict may arise from the different degrees of heterogeneity in the folding mechanism as a function of hairpin size and temperature. It should be emphasized again the discussion presented here is based on the assumption of K_{melt} is equal to K_{eq}. In Chapter 4, we show that the reaction described here is not the whole process of folding and unfolding, but the fluctuation between open conformation and intermediate state with conformation of misfiled base pairs. Clearly this is an issue that merits further study.

3.6.4 Programming code for the Combined FCS analysis (Matlab)

```
function fitplot1 = fcs_hp   %Modified since 03/20/03
[fname,pname] = uigetfile('*.*','Select data file');

%%%%%%%%%%%%%%%%% Control Panel%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Choose initial values
So2=input('(Wo/Zo)^2=?  ');
tauf0=input('tauf0 in second: ');
r0=input('R/Wo=?  ');
Ampl0=input('1/N=?  ');
taud0=input('taud0 in second:  ');
taut0=input('taut0 in second:  ');
T0=input('T0 :  ');
taur0=input('taur in second:  ');
B0=input('B0:  ');

%Choose fitting mode
fixr = input('r is fixed for CC fitting? 0 if no, 1 if yes,  ');
aftcorr = input('AC fitting with afterpulsing correction? 0 if no, 1 if yes,  ');
hairpin = input('sample ? 0 if control DNA, 1 if hairpin DNA,  ');
bod = input('boundary factor:  ');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Opening data file
Gtemp=[]; Gtemp2=[];
fid=fopen(fname);
for j=1:7
   str=fgetl(fid);
end

%Reading FCS data
cmp=0;
while cmp ~= 1
   str=fgetl(fid);
   cmp=strcmp(str,'');
   if cmp ~= 1
      Gtemp=[Gtemp;str2num(str)] ;
   end
end
for j=1:26
   str=fgetl(fid);
end

%Reading Count rate data
cmp=0;
while cmp ~= 1
   str=fgetl(fid);
   cmp=strcmp(str,'');
   if str == -1
      cmp = 1;
   end
   if cmp ~= 1
      Gtemp2=[Gtemp2;str2num(str)];
   end
```

```
end
status=fclose(fid);

%%%%%%%%%%%%%%%%%%%%%Cross Correlation curve fitting%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%% Getting tauf %%%%%%%%%%%%%%%%%%%%%%%%%%%%
ginf0=0;
tau = Gtemp(46:93,1);        %tau is the lagtimes in seconds
Gtau = Gtemp(46:93,3);        %Gtau contains the cross-correlation data.

tau1=min(tau):1e-6:max(tau);
tol = input('tolerance '); %fitting tolerance
error=0.1;
while error > tol
   pr0=[ginf0;Ampl0;tauf0;taud0;r0];
   l_b=[-.01e-3;Ampl0/bod;tauf0/bod;taud0/bod;r0/bod]; %lower boundary for the fitting parameters.
   u_b=[.01e-3;Ampl0*bod;tauf0*bod;taud0*bod;r0*bod]; %upper boundary for the fitting parameters.
   [pr,resnorm,residual,exitflag,output,lambda,jacobian] = lsqcurvefit(@ccfit1,pr0,tau,Gtau,l_b,u_b,[],So2);
   c1 = nlparci(pr,residual,jacobian); %confidential interval
   Ft=ccfit1(pr,tau1,So2);
   Ampl1=pr(2);
   taud1=pr(4);
   r1=pr(5) ;
   Amplitude=max(Ft);

   close all
   figure(1)
   subplot(3,2,[1 3])
   semilogx(tau,Gtau,'k.')
   hold on
   semilogx(tau1,Ft,'k')
   hold on

   % %%%%%%%%%%%%%%%%%%%%%Fitting with tauf,ginf0 fixed%%%%%%%%%%%%%
   ginf1=pr(1);
   tauf1=pr(3);
   pr0=[Ampl1;taud1;r1];
   l_b=[Ampl1/bod;taud1/bod;r1/bod];
   u_b=[Ampl1*bod;taud1*bod;r1*bod];
   [pr,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@ccfit2,pr0,tau,Gtau,l_b,u_b,[],tauf1,So2,ginf1);
   c2 = nlparci(pr,residual,jacobian);
   Ft2=ccfit2(pr,tau1,tauf1,So2,ginf1);
   Ampl2=pr(1);
   taud2=pr(2);
   r2=pr(3);
   Amplitude=max(Ft2);
   figure(1)
   subplot(3,2,[1 3])
   semilogx(tau1,Ft2,'r')
   hold on

   %%%%%%%%%%%%%%%%%%%%%%%%Fitting with Gmax %%%%%%%%%%%%%%%%%%%%%%
   pr0=[taud2;r2];
   l_b=[taud2/bod;r2/bod];
   u_b=[taud2*bod;r2*bod];
```

```
  [pr,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@ccfit3,pr0,tau,Gtau,l_b,u_b,[],tauf1,So2,ginf1,Amplitude,taud2);
  resnorma=resnorm;
  c3 = nlparci(pr,residual,jacobian);
  Ft3=ccfit3(pr,tau1,tauf1,So2,ginf1,Amplitude,taud2);
  taud3=pr(1);
  r3=pr(2);
  figure(1)
  subplot(3,2,[1 3])
  semilogx(tau1,Ft3,'b')
  hold on
  %%%%%%%%%%%%%%%%%%%%%%%%%%%%Using tau_max%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  Gmax=max(Ft3);
  taumax=-(taud3+tauf1^2./(2*taud3*r3^2));
  taumax=taumax+sqrt((taud3+(tauf1^2/(2*taud3*r3^2)))^2+tauf1^2*(1+1/(r3^2))+2*taud3*tauf1);
  G_taumax=interp1(tau,Gtau,taumax);
  error = Gmax-G_taumax
  tauf0=tauf1;
  Ampl0=Ampl2;
end

if fixr == 0
  tauf=tauf1
  rr=[r1; r2; r3]
  Ampli=[Ampl1; Ampl2]
  Taud=[taud1; taud2; taud3]
else if fixr == 1
      pr0=[Ampl2;taud3;tauf1];
      l_b=[Ampl2/bod;taud3/bod;tauf1/bod];
      u_b=[Ampl2*bod;taud3*bod;tauf1*bod];
      [pr,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@ccfit4,pr0,tau,Gtau,l_b,u_b,[],r3,So2,ginf1);
      c2 = nlparci(pr,residual,jacobian);
      Ft4=ccfit4(pr,tau1,r3,So2,ginf1);
      Ampl3=pr(1);
      taud4=pr(2);
      tauf2=pr(3);
      tauf=tauf2
      rr=[r1; r2; r3]
      Ampli=[Ampl1; Ampl2; Ampl3]
      Taud=[taud1; taud2; taud3; taud4]
      Ampl2=Ampl3;
      taud3=taud4;
      figure(1)
      subplot(3,2,[1 3])
      semilogx(tau1,Ft4,'g')
      hold on
   end
end
Amplitude=Ampl2;
nn_c=1/Amplitude
figure(1)
subplot(3,2,5)
semilogx(tau,residual,'b:+')
hold on
```

```
%%%%%%%%%%%%%%%%%%%%AC fitting%%%%%%%%%%%%%%%%%%%%%%%%
%Auto correlation data
taua = Gtemp(1:93,1);              %tau is the lagtimes in seconds
Gtaua = Gtemp(1:93,5);             %Gtau contains the cross-correlation data.

%%%%%%%%%AFTERPULSING CORRECTION%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if aftcorr ==1
   al=1.55;
   af=1.97;
   I=mean(Gtemp2); %average intensity, '<k>'
   I=I(2);
   Gcorr=al*exp(-af*log10(taua.*1000000))*1.0125/I;
   Gcorr=Gcorr(1:14);
   Gtauatemp=Gtaua(1:14);
   Gtauatemp=Gtauatemp-Gcorr;
   Gtaua=[Gtauatemp;Gtaua(15:93)];
end

%%%%%%%%%%%%%%%%%%%%Fit the curve with taud and r %%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%Only triplet part fitting%%%%%%%%%%%%%%%%%%%%%%%%%%%%
tau1=min(taua):1e-6:max(taua);
p0=[taut0;T0];
lb=[taut0/bod;T0/bod];
ub=[taut0*bod;T0*bod];
[p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@acfit1,p0,taua,Gtaua,lb,ub,[],tauf,Amplitude,So2,taud3,r3);
Fmc=acfit1(p,tau1',tauf,Amplitude,So2,taud3,r3);
cac = nlparci(p,residual,jacobian);
res=residual.*residual./93;
kai2=sum(res)

figure(1)
subplot(3,2,[2 4])
semilogx(taua,Gtaua,'k.')
hold on
semilogx(tau1,Fmc,'r')
hold on
subplot(3,2,6)
semilogx(taua,residual,'r-')
hold on
taut=p(1);
T=p(2);
Tt=T/(1+T);
%%%%%%%%%%%%%%%%%%%%Fit the curve without triplet part %%%%%%%%%%%%%%%%%%%%%%%%
ginfa0=0;
p0=[ginfa0];
lb=[ginfa0-1e-6];
ub=[ginfa0+1e-6];
[p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@acfit2,p0,taua,Gtaua,lb,ub,[],tauf,Amplitude,So2,taud3,r3);
Fmd=acfit2(p,tau1',tauf,Amplitude,So2,taud3,r3);
res=residual.*residual./93;
kai2=sum(res)

figure(1)
subplot(3,2,[2 4])
```

```
semilogx(tau1,Fmd,'g')
hold on
subplot(3,2,6)
semilogx(taua,residual,'g:+')
hold on

%%%%%%%%%%Fit with Conformational fluctuation part%%%%%%%%%%%%%%%%%%%%%%%%
if hairpin == 1
p0=[taut0;T0;taur0;B0];
lb=[taut0/bod;T0/bod;taur0/bod;B0/bod];
ub=[taut0*bod;T0*bod;taur0*bod;B0*bod];
[p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@acfit3,p0,taua,Gtaua,lb,ub,[],tauf,Amplitude,So2,taud3,r3);
Fmc=acfit3(p,tau1',tauf,Amplitude,So2,taud3,r3);
cac = nlparci(p,residual,jacobian)
aaa=cac(3,:)*1000000
Nac = interp1(tau1',Fmc,0,'pchip','extrap');
Nac=1/Nac
res=residual.*residual./93;
kai2=sum(res)

figure(1)
subplot(3,2,[2 4])
semilogx(taua,Gtaua,'k.')
hold on
semilogx(tau1,Fmc,'b')
hold on
subplot(3,2,6)
semilogx(taua,residual,'b:+')
hold on
taut=p(1);
T=p(2);
taur=p(3)
B=p(4)
end

%Correction for PCH analysis
taut=taut;
T=T;
tau=input('timebin of PCH:  ') ;
gamma_t=(1+2*T*taut/tau*(1-taut/tau*(1-exp(-tau/taut))))
Npch=nn_c/gamma_t
Tt=T/(1+T);
N=nn_c*(1-Tt);
FF=[];
for j=1:93
   F4=ginfa0+Amplitude*exp(-
((taua(j)*r3/tauf).^2)./(1+taua(j)/(taud3))).*((1+taua(j)/(taud3)).*sqrt(1+So2*taua(j)/(taud3)));
   FF=[FF;F4];
end

figure(1)
subplot(3,2,[2 4])
semilogx(taua,FF,'g')
hold on
```

```
%%%%%%%%%%%%%%%Fit the curve with control model %%%%%%%%%%%%%%%%%%%%%%%%%
[p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@mfinal0831,p0,taua,Gtaua,lb,ub,[],tauf,Amplitude,So2,taud0,r0,ginfa0);
resnormb=resnorm;
Fmc=mfinal0831(p,tau1',tauf,Amplitude,So2,taud0,r0,ginfa0);

figure(3)
semilogx(tau1,Fmc,'g')
hold on
figure(4)
semilogx(taua,residual,'g:+')

%%%%%%%%%%%%%%%%%% Recording Data %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[fname,pname] = uiputfile('*.rtf','Save Correlation Data');
if fname~=0
   filename=strcat(pname,fname);
   fid=fopen(filename,'w');
   count=fprintf(fid,'%s %s\n','File name:',fname);
   count=fprintf(fid,'%s %s\n','From Cross correlation curve fitting','');
   count=fprintf(fid,'%s %6.5f %s\n','N:',1/Amplitude,'');
   count=fprintf(fid,'%s %6.5f %s
(%6.5f %6.5f) %6.2f %s\n','Tauf:',tauf*1e3,'ms',cj(3,1)*1e3,cj(3,2)*1e3,(cj(3,2)-tauf)/tauf,'');
   count=fprintf(fid,'%s %6.5f %s
(%6.2f %6.2f) %6.2f %s\n','Taud0:',taud0*1e6,'us',c2(1,1)*1e6,c2(1,2)*1e6,(c2(1,2)-taud0)/taud0,'');
   count=fprintf(fid,'%s %6.5f %s (%6.2f %6.2f) %6.2f %s\n','r0(R/wo):',r0,'',c2(2,1),c2(2,2),(c2(2,2)-
r0)/r0,'');
   count=fprintf(fid,'%s %6.5e %s\n %s\n','resnorm:',resnorma,'','');
   count=fprintf(fid,'%s %s\n','From Auto correlation curve fitting','');
   count=fprintf(fid,'%s %6.5f %s
(%6.2f %6.2f) %6.2f %s\n','Taut:',taut*1e6,'us',ci(1,1)*1e6,ci(1,2)*1e6,(ci(1,2)-taut)/taut,'');
   count=fprintf(fid,'%s %6.5f %s (%6.2f %6.2f) %6.2f %s\n','T/(1-T):',T,'',ci(2,1),ci(2,2),(ci(2,2)-T)/T,'');
   count=fprintf(fid,'%s %6.5e %s\n','resnorm:',resnormb,'');

   sta=fclose(fid);
end
end

%%%%%%%%%%%%%%Subroutines %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function F= ccfit1(pr,tau1,So2)
F=pr(1)+pr(2)*exp(-(pr(5).^2)*((1-
tau1/pr(3)).^2)./(1+tau1/pr(4)))./((1+tau1/pr(4)).*sqrt(1+So2*tau1/pr(4)));
end

function F=ccfit2(pr,tau1,tauf1,So2,ginf1)
F=ginf1+pr(1)*exp(-(pr(3).^2)*((1-
tau1/tauf1).^2)./(1+tau1/pr(2)))./((1+tau1/pr(2)).*sqrt(1+So2*tau1/pr(2)));
end

function F=ccfit3(pr,tau1,tauf1,So2,ginf1,Amplitude,taud2)
F=ginf1+Amplitude*((1+tauf1/taud2).*sqrt(1+So2*tauf1/taud2))*exp(-(pr(2).^2)*((1-
tau1/tauf1).^2)./(1+tau1/pr(1)))./((1+tau1/pr(1)).*sqrt(1+So2*tau1/pr(1)));
end

function F=ccfit4(pr,tau1,r3,So2,ginf1)
F=ginf1+pr(1)*exp(-(r3.^2)*((1-tau1/pr(3)).^2)./(1+tau1/pr(2)))./((1+tau1/pr(2)).*sqrt(1+So2*tau1/pr(2)));
end
```

```
function F=acfit1(p,tau1,tauf,Amplitude,So2,taud3,r3)
F=Amplitude*exp(-
((tau1.*r3./tauf).^2)./(1+tau1/(taud3))).*((1+tau1/(taud3)).*sqrt(1+So2*tau1/(taud3))).*(1+p(2)*exp(-
tau1/p(1)));
end

function F=acfit2(p,tau1,tauf,Amplitude,So2,taud3,r3)
F=Amplitude*exp(-((tau1.*r3./tauf).^2)./(1+tau1/(taud3))).*((1+tau1/(taud3)).*sqrt(1+So2*tau1/(taud3)));
end

function F=acfit3(p,tau1,tauf,Amplitude,So2,taud3,r3)
F=Amplitude*exp(-
((tau1.*r3./tauf).^2)./(1+tau1/(taud3))).*((1+tau1/(taud3)).*sqrt(1+So2*tau1/(taud3))).*(1+p(2)*exp(-
tau1/p(1))).*(1+p(4)*exp(-tau1/p(3)));
end
```

3.7 REFERENCES

1. Shen, Y., S.V. kuznetsov, and A. Ansari, *Loop Dependence of the Dynamics of DNA Hairpins.* J. Phys. Chem. B, 2001. **105**: p. 12202-12211.
2. Kuznetsov, S.V., et al., *A Semiflexible Polymer Model Applied to Loop Formation in DNA Hairpins.* Biophys. J., 2001. **81**: p. 2864-2875.
3. Ansari, A., S.V. kuznetsov, and Y. Shen, *Configurational Diffusion Down a Folding Funnel Describes the Dynamics of DNA Hairpins.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 7771-7776.
4. Lippincott-Schwartz, J. and e.a. J. F. Presley, *ER-to-Golgi Transport Visualized in Living Cells.* Nature, 1997. **389**: p. 81-85.
5. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence Correlation Spectroscopy 2. An Experimental Realization.* Biopolymers, 1974. **13**: p. 29-61.
6. Elson, E.L. and D. Magde, *Fluorescence Correlation Spectroscopy 1. Conceptual Basis and Theory.* Biopolymers, 1974. **13**: p. 1-27.
7. Van Orden, A., K. Fogarty, and J. Jung, *Fluorescence Fluctuation Spectroscopy: A Coming of Age Story.* Appl. Spectrosc., 2004. **58**: p. 122A-137A.
8. Bonnet, G., O. Krichevsky, and A. Libchaber, *Kinetics of Conformational Fluctuation in DNA Hairpin Loops.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 8602-8606.
9. Goddard, N.L., et al., *Sequence Dependent Rigidity of Single Stranded DNA.* Phys. Rev. Lett., 2000. **85**: p. 2400-2403.
10. Chen, Y., et al., *The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy.* Biophys. J., 1999. **77**: p. 553-567.
11. Kask, P., et al., *Fluorescence Intensity Distribution Analysis and Its Application in Biomolecular Detection Technology.* Proc. Natl. Acad. Sci. USA, 1999. **96**(24): p. 13756-13761.
12. Perroud, T.D., et al., *Photon Counting Histogram for One-Photon Excitation.* Chemphyschem, 2003. **4**: p. 1121-1123.
13. LeCaptain, D.J. and A. Van Orden, *Two-Beam Fluorescence Cross-Correlation Spectroscopy in an Electrophoretic Mobility Shift Assay.* Anal. Chem., 2002. **74**: p. 1171-1176.
14. Brinkmeier, M., et al., *Two-Beam Cross-Correlation: a Method to Characterize Transport Phenomena in Micrometer-Sized Structures.* Anal. Chem., 1999. **71**: p. 609-616.
15. Widengren, J., U. Mets, and R. Rigler, *Fluorescence Correlation Spectroscopy of Triplet States in Solution: a Theoretical and Experimental Study.* J. Phys. Chem. B, 1995. **99**: p. 13368-13379.
16. Rigler, R. and E.L. Eigen, *Fluorescence Correlation Spectroscopy: Theory and Applications.* 2001, Berlin: Springer-Verlag.

17. Ying, L., M.I. Wallace, and D. Klenerman, *Two-State Model of Conformational Fluctuation in a DNA Hairpin-Loop.* Chem. Phys. Lett., 2001. **334**: p. 145-150.

18. Nkobo, A.E., et al., *Diffusion coefficient of DNA molecules during free solution electrophoresis.* Electrophoresis, 2001. **22**: p. 2424-2432.

19. Widengren, J., ? Mets, and R. Rigler, F*luorescence correlation spectroscopy if triplet states in solution: A theoretical and experimental study.* J. Phys. Chem., 1995. 9**9:** p. 13368-13379.

20. Bonnet, G., O. Krichevsky, and A. Libchaber, K*inetics of conformational fluctuations in DNA hairpin loops.* Proc. Natl. Acad. Sci. USA, 1998. 9**5:** p. 8602-8606.

21. Goddard, N.L., et al., S*equence dependent rigidity of single-stranded DNA.* Phys. Rev. Lett., 2000. 8**5:** p. 2400-2403.

22. Wallace, M.I., et al., F*RET fluctuation spectroscopy: Exploring the conformational dynamics of a DNA hairpin loop.* J. Phys. Chem. B, 2000. 1**04:** p. 11551-11555.

23. Wallace, M.I., et al., N*on-Arrhenius kinetics for the loop closure of a DNA hairpin.* Proc. Natl. Acad. Sci. USA, 2001. 9**8**(10): p. 5584-5589.

24. Li, H., et al., M*easuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy.* Proc. Natl. Acad. Sci. USA, 2004. 1**01:** p. 14425-14430.

25. Wang, X. and W.M. Nau, K*inetics of end-to-end collision in short single-stranded nucleic acids.* J. Am. Chem. Soc., 2004. 1**26:** p. 808-813.

Chapter 4

Equilibrium Distribution of DNA Hairpin Conformers

In this chapter, a new approach to characterize the equilibrium distribution between different conformational states of a biomolecule is discussed. The folding of a dye-quencher labeled DNA hairpin molecule was investigated using fluorescence auto- and cross-correlation spectroscopy (FCS) and photon counting histogram analysis (PCH). The auto- and cross-correlation measurements revealed the flow and diffusion times of the DNA molecules through two spatially offset detection volumes, the relaxation time of the folding reaction, and the total concentration of DNA molecules participating in the reaction. The PCH measurements revealed the equilibrium distribution of DNA molecules in folded and unfolded conformations and the specific fluorescence brightnesses of the fluorophor in each conformational state. These measurements were

carried out over a range of NaCl concentrations, from those that favored the unfolded form of the DNA hairpin to those that favored the closed form. DNA melting curves obtained from each sample were also analyzed for comparison. The FCS and PCH measurements revealed that the total concentration of the reactants being probed was depleted as the conditions were adjusted to favor the closed form of the DNA hairpin. It was also found that the equilibrium distribution measured by FCS and PCH agreed with the distribution determined from the melting curve analysis only for reaction conditions that favored the open form of the DNA. As the NaCl concentration was raised, the melting curves showed the expected shift in the equilibrium distribution toward the folded hairpin; whereas, the FCS and PCH measurements showed a nearly constant equilibrium distribution. These observations suggest a three-state mechanism for the reaction, involving a stable intermediate form of the DNA hairpin. The reaction being probed by FCS and PCH is suggested to be a rapid equilibrium between open and intermediate conformations. The final, closed form of the DNA hairpin is suggested to be stable on a much longer time scale than that of the FCS and PCH measurements, and thus serves as a sink into which the observed reactants are depleted as the reaction progresses.

4.1 INTRODUCTION

As described in Chapter 1, DNA stem-loop hairpins are model systems for secondary structure formation in single stranded DNA and RNA.[1-5] In recent years, there has been a resurgence of interest relating to the kinetics and energetics of DNA hairpin folding, fueled in part by the development of new experimental techniques for

probing the microsecond time scale on which these reactions occur.[6-16] Experimental methods that have been utilized include laser-induced temperature jump spectroscopy[6, 9, 11, 16] and fluorescence correlation spectroscopy (FCS)[7, 8, 12-15]. These experiments consistently observe chemical relaxation processes attributed to DNA hairpin folding and unfolding. The relaxation time observed for a given hairpin depends on the size of the hairpin loop, the number of base pairs in the stem, and the loop sequence, and can range from a few microseconds to hundreds of microseconds. It has generally been assumed that these relaxation processes describe a two-state reaction in which the DNA molecules spend most of their time in a fully folded stem-loop hairpin structure or a fully unfolded random coil configuration. Partially folded, or misfolded intermediates are thought to be transient and short-lived compared to the overall folding reaction.[7, 9, 17]

In the previous study described in Chapter 3, we characterized the relaxation times of DNA hairpin folding reactions using a new "multi-parameter" approach to FCS.[15] This technique analyzes the fluorescence observed from a flowing sample of DNA hairpin molecules at two spatially offset detection volumes. Simultaneous cross-correlation and auto-correlation analysis of the fluorescence from the two detection volumes resolves the transport properties of the molecules from their chemical relaxation properties, and thus allows more information to be drawn from the system than could be obtained using conventional single-beam FCS alone.

The present study combines this multi-parameter FCS technique with photon counting histogram (PCH) analysis[18-20], which monitors the absolute and relative concentrations of the reactants taking part in the observed chemical relaxation process, in

addition to the reaction time. We studied the folding of a stem-loop DNA hairpin molecule containing five base pairs in the stem and twenty-one thymine residues in the loop region. The reaction conditions were varied from those favoring the open form of the DNA hairpin, to those favoring the closed form. Our results do not support a two-state reaction mechanism for the folding of this DNA hairpin. For example, one of our key observations is that the overall reactant concentration was dramatically depleted as the reaction conditions were altered in favor of the closed form of the hairpin. This observation cannot be explained from the standpoint of a two-state reaction mechanism. If only two states were present, the overall reactant concentration would not depend on the reaction conditions. Only the relative concentrations would have this dependence. We can explain this and other observations presented below by assuming the chemical relaxation process being probed by FCS and PCH involves an intermediate form of the DNA hairpin that is stable on the time scale of the observed reaction. The fully folded configuration is not directly observed in our FCS and PCH measurements, but serves as a stable sink into which the reactants become transformed as the reaction progresses. Our observations are suggestive of a three-state DNA hairpin folding mechanism involving a rapid equilibrium between open and intermediate configurations, and a much slower formation of the fully folded stem-loop hairpin. Hence, we describe a DNA hairpin reaction that is even more complex and more drawn out than previously thought. (*J. Am. Soc. Chem.* **2005** Submitted.)

4.2 EXPERIMENTAL SECTION

4.2.1 Sample preparation and Instrumentation

Details of the experiment were described in the last two chapters. Briefly, DNA hairpin was dual-labeled with Rhodamine 6G (R6G) and 4-{[4-(dimethylamino) phenyl] axo} benzoic acid (dabcyl) at the 5' and 3' ends, respectively. The hairpin sequence was 5'-R6G-AACCC-$(T)_{21}$-GGGTT-dabcyl-3'. DNA solutions with nanomolar DNA concentrations were prepared in a pH ~8.0 buffer solution containing 2.5-mM Tris-HCl, 250-μM EDTA, and various concentration of NaCl. The quality of the hairpin samples was confirmed by monitoring the fluorescence intensities at the temperatures varying from 5°C to 85°C, on a steady state flourometer.

The FCS and PCH experimental setup is shown in Figure 4.1. It is similar setup previously described in the Chapter 3, with some modification. The sample solutions were flowed continuously through the capillary by pressurizing the capillary inlet with N_2 gas. Fluorescence fluctuation spectroscopy (FFS) measurements were performed using a home-built single molecule confocal fluorescence microscope.[21] Two split laser beams (514.5 nm, 40-μW per beam before focus lens) were focused into the capillary by a 100× 1.25 NA oil immersion microscope objective. Fluorescence from each focal region was collected through the same objective; passed through the 530-nm long pass dichroic mirror; split by 50/50 beam splitter; then imaged into two 50-μm confocal pinholes separately, (Thorlabs, Newton, NJ) which provide axial resolution by rejecting out-of-focus light. The output of the pinholes was band-pass filtered and focused onto the active areas of two single-photon-counting avalanche photodiode detectors. We found better collection efficiency as well as better signal to noise ratio can be achieved by using pinholes to image signal fluorescence, instead of using optical fibers. The two laser beams formed nearly identical diffraction-limited focal regions, positioned near the

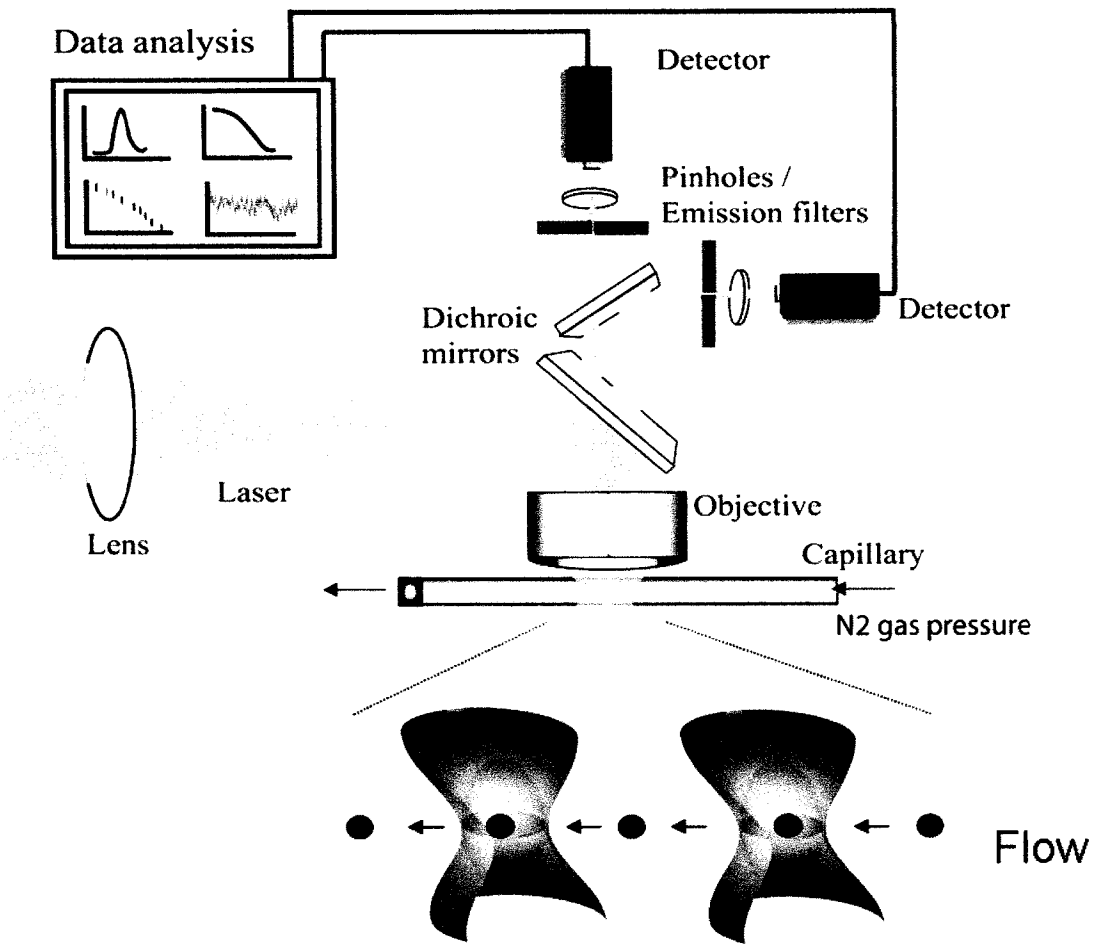

Figure 4.1 Schematic representation of the two-beam FFS experiment
The optical setup was designed to position the two laser beam in the center of the squared capillary, separated by a distance of R (shown in bottom). The sample solution flows through the capillary under applied gas pressure. Fluorescence signal from each laser focal volume is collected and used for the simultaneous analysis of Auto- and Cross-Correlation as well as PCH.

center of the inner capillary space (~25-μm from the inner surfaces), separated along the axial dimension of the capillary by a distance, R (~3-μm). The optimum position of the two laser beam foci relative to the z-position of the capillary and the flow axis could be confirmed by cross-correlation analysis.[22, 23] The entire experiment was placed inside an enclosure to block stray laser light. All experiments were carried out at an ambient temperature of ~19°C.

4.2.2 Data Acquisition

The photocounts from the two detectors were recorded using the two channels of a 800-MHz gated photon counter (PMS-400, Becker & Hickl GmbH, Berlin, Germany), interfaced to a Pentium computer. A sampling time of 1-μs per sampling interval was used for each channel. A LabView computer program, written in-house, was used to perform real-time FFS analysis on the detected photons as they were being accumulated. The program simultaneously autocorrelated the photocounts observed from the individual detection volumes, and cross-correlated the photocounts from the two detection volumes relative to each other. The cross-correlation functions were calculated after first re-binning the accumulated photocounts into 9-μs sampling intervals to achieve improved signal-to-noise ratio. All correlation functions were calculated using the "multiple-τ" concept, in which the intervals between successive lagtimes increase as the lagtime increases. The rebinned photocounts of 9-μs sampling intervals were also used to construct photon counting histograms. The histograms obtained from each detection channel were identical within experimental error. The user interface of the LabView program continuously displays a running average of the correlation functions and the

cumulative histograms. The experimental photon counting distribution, $\{p_k\}$ was analyzed by fitting the normalized histogram to a theoretical PCH model, $\{\bar{p}_k\}$, devised by Perroud *et al.*[20] The PCH fitting quality is determined by estimating reduced χ^2_{red} based on a binomial distribution statistics.[18, 24]

$$\chi^2_{red} = \frac{\sum_{k_{min}}^{k_{max}} r_k^2}{k_{max} - k_{min} - d} \tag{4.1}$$

where, d is number of fitting parameters and r_k is a normalized residual given by

$$r_k = \frac{N(p_k - \bar{p}_k)}{\sigma_k}, \tag{4.2}$$

where, N is the total number of experimental data points in the histogram and σ_k is a standard deviation given by

$$\sigma_k = \sqrt{N p_k (1 - p_k)}. \tag{4.3}$$

Analyses of both experimental correlation functions and photon counting distribution were carried out using a Levenberg-Marquardt algorithm (least-squares fitting routine) from the MatLab Optimization Toolbox.

4.3 ANALYSIS AND RESULTS

4.3.1 FCS and PCH Analysis

We begin by describing our efforts to evaluate our FCS and PCH measurements assuming a DNA hairpin folding reaction that goes to completion via a two-state reaction process:

$$\bar{N}_1 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \bar{N}_2 \tag{4.4}$$

Here, $\bar{N}_1$ and $\bar{N}_2$ refer to the average number of DNA hairpin molecules occupying the detection volume of the optical microscope in their open and closed conformations, respectively. Figure 4.2A displays auto- and cross-correlation functions (inset) obtained from DNA hairpin samples containing constant DNA concentration and varying NaCl concentration. The cross-correlation functions, $G_{C,2S}(\tau)$, were analyzed by fitting to Equation (4.5)[22, 23, 25-29]:

$$G_{C,2S}(\tau)-1=\frac{1}{\bar{N}_1+\bar{N}_2}\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\tau/\tau_D}\right)^{1/2}\exp\left(-\frac{r^2(1-\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \qquad (4.5)$$
$$=\frac{1}{\bar{N}_{CC}}g_D(\tau)g_{Fc}(\tau)$$

where

$$g_D(\tau)=\left(\frac{1}{1+\tau/\tau_D}\right)\left(\frac{1}{1+\kappa_o^2\tau/\tau_D}\right)^{1/2} \qquad (4.6)$$

$$g_{Fc}(\tau)=\exp\left(-\frac{r^2(1-\tau/\tau_F)}{1+\tau/\tau_D}\right) \qquad (4.7)$$

The parameter τ_D is the average diffusion time of the DNA molecules through the detection volume. κ_0 is the ratio (ω_0/z_0) of the radial, ω_0, and axial, z_0, focal radii and has been held constant at the value 0.104±0.002 determined previously. r is the ratio (R/ω_0) of the distance, R, between the two detection volumes and the focal radius, ω_0. τ_F is the average flow time of the DNA molecules between the two focal volumes and is equal to R/V_x, where V_x is the linear flow velocity of the analyte solution at the point of detection. Finally, $\bar{N}_{CC}$ represents the average number of DNA hairpin molecules in both open and closed conformations occupying the detection volume ($\bar{N}_{CC}=\bar{N}_1+\bar{N}_2$).

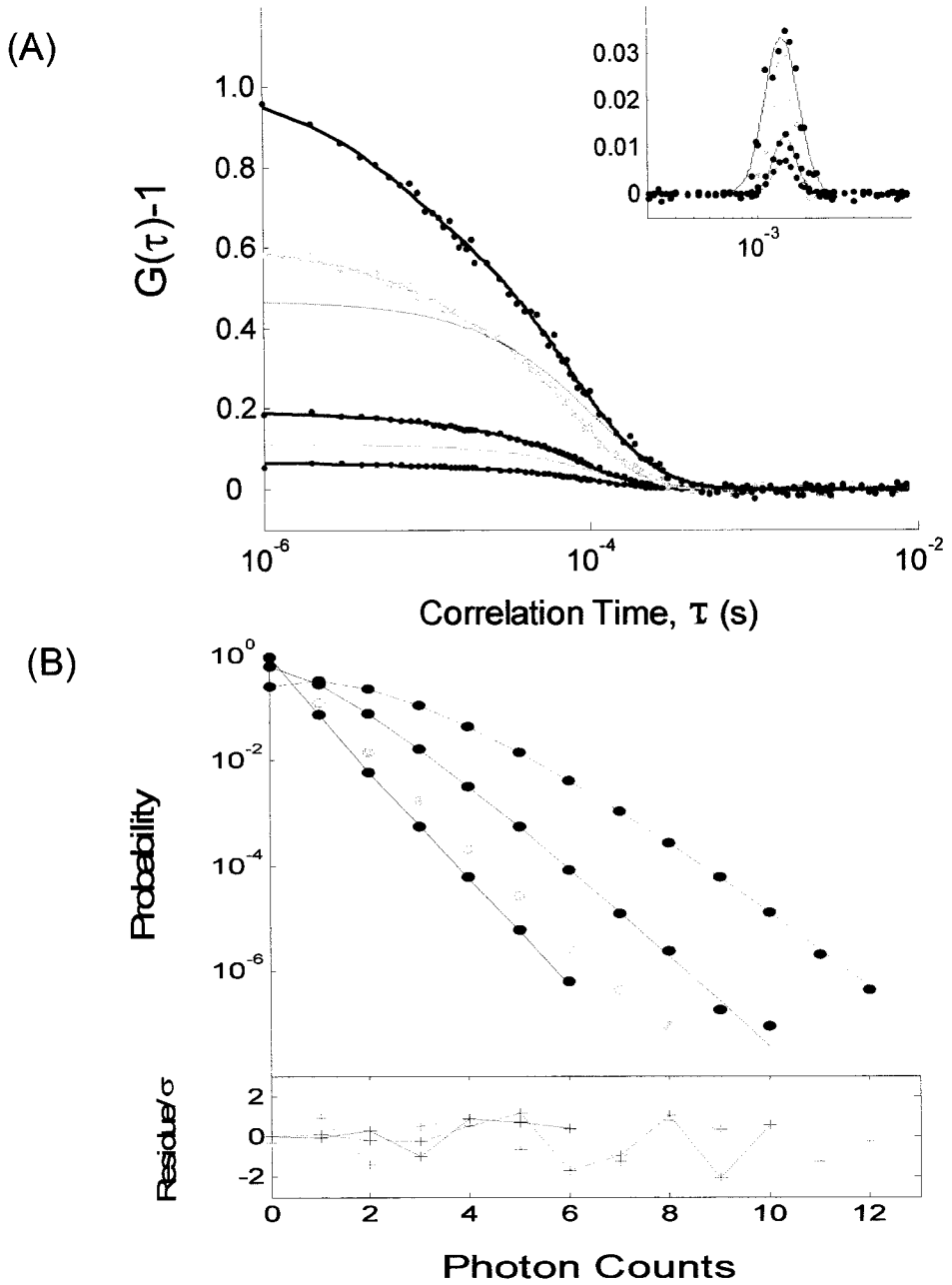

Figure 4.2 Experiment data (dots) and corresponding fitting curves (solid lines) from simultaneous FCS and PCH measurements vs. NaCl concentration (blue: 0-, green: 25-, red: 100-, and black: 500-mM NaCl). Panel (A) shows the observed auto-correlation functions, with the cross correlation functions shown in the inset. The dashed curves are predicted auto-correlation functions based on the cross-correlation analysis. Panel (B) shows the photon counting histograms obtained from the same samples using 9-μs sampling intervals. The DNA concentration was held constant at ~25-nM for all the samples.

For a two-state chemical reaction with a relaxation time faster than the flow time (τ_F), the cross-correlation function is only sensitive to the diffusion and flow properties of the DNA molecules.[15] Fluorescence fluctuations caused by the reaction are not sufficiently correlated with respect to the two detection volumes under these conditions. Cross-correlation analysis thus gives the parameters τ_D, τ_F, r, and $\bar{N}_{CC}$, without contributions from chemical relaxation. The dotted curves in Figure 4.2A predict how the autocorrelation functions would appear under conditions of pure diffusion and flow of the DNA molecules, based on the measured cross-correlation parameters. Any deviation between the predicted and observed autocorrelation functions is attributed to relaxation caused by triplet blinking at shorter time scales, and quenching and unquenching of the fluorescent dye, induced by DNA conformational fluctuations, at longer time scales.

Parameters relating to the chemical relaxation of the DNA hairpin structure were obtained by fitting the experimental autocorrelation functions to Equation (4.8), which corresponds to the theoretical autocorrelation function, $G_{A,2S}(\tau)$, describing two-state chemical relaxation.[25, 30-36]

$$\begin{aligned} G_{A,2S}(\tau)-1 &= \frac{1}{\bar{N}_{CC}} g_D(\tau)\left(1+T_{eq}e^{-\tau/\tau_T}\right)\left[1+B_{eq}e^{(-\tau/\tau_R)^\beta}\right]\exp\left(-\frac{(r\,\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \\ &= \frac{1}{\bar{N}_{CC}} g_D(\tau) g_T(\tau) g_R(\tau) g_{Fa}(\tau) \end{aligned} \tag{4.8}$$

where

$$g_T(\tau) = \frac{1-T+Te^{-\tau/\tau_T}}{1-T} = 1+T_{eq}e^{-\tau/\tau_T}\ , \tag{4.9}$$

$$g_R(\tau) = \frac{1-B+Be^{-\tau/\tau_R}}{1-B} = 1+B_{eq}e^{-\tau/\tau_R}\ , \text{ and} \tag{4.10}$$

$$g_{Fa}(\tau) = \exp\left(-\frac{(r\,\tau/\tau_F)^2}{1+\tau/\tau_D}\right) \quad . \tag{4.11}$$

Here, T is the quantum yield and τ_T is the time constant for triplet blinking; τ_R is the relaxation time of the two-state reaction; β accounts for the possibility of "stretched" exponential relaxation[12]; and B is an amplitude term that depends on the equilibrium distribution of the reactants and their specific brightnesses. These parameters were determined by constraining τ_D, τ_F, r, and $\bar{N}_{CC}$ to the values obtained in the cross-correlation analysis, and adjusting τ_T, T_{eq}, τ_R and B_{eq} to achieve the best fit. Representative results from our auto- and cross-correlation analyses are presented in Table 4.1. Relaxation times (τ_R) for the DNA hairpin folding reaction were observed to be ~60-μs. These relaxation times are consistent with previous kinetics experiments on the folding of DNA hairpins with the same size and composition as the one studied here.[7, 8, 15]

What is of interest to us here is the equilibrium distribution of folded and unfolded DNA hairpin molecules as a function of NaCl concentration. It is known that NaCl stabilizes the folded form of the hairpin.[37] This is evident from the melting curves displayed in Figure 4.3B, which show the characteristic rise in melting temperature with added NaCl. If a two-state reaction is assumed, the melting curves can be analyzed to determine the equilibrium constant $K_{melt} = \bar{N}_2/\bar{N}_1$ using Equation (4.12):

$$K_{melt}(T) = \frac{I(85^\circ\mathrm{C}) - I(T)}{I(T) - I(5^\circ\mathrm{C})} \tag{4.12}$$

where $I(T)$ is the fluorescence intensity at temperature T. We have evaluated K_{melt} for each sample at our laboratory temperature of 19°C (See Figure 4.3A). As expected, K_{melt}

TABLE 1: Parameters from FFS Analysis[a,b]

PCH Analysis[c]			FCS Analysis	
Model:	one species	two species	Cross Correlation	
$\overline{N}_{app}$	1.54(03)	1.44(55)	r	11.8(06)
ϵ_{app}	0.539(31)	0.48(18)	$\overline{N}_{cc}$	4.26(05)
			τ_F (ms)	1.47(15)
$\overline{N}_1$	1.27(12)	0.72(16)	τ_D (μs)	214(13)
$\overline{N}_2$	2.98(22)	0.98(25)		
ϵ_1	0.591(17)	0.61(14)	Auto Correlation	
ϵ_2	0.074(14)	0.26(12)		
			T_{eq}	0.198(24)
F	0.732(54)	0.71(16)	B_{eq}	0.987(97)
χ^2	1.75	1.69	τ_T (μs)	4.12(71)
			τ_R (μs)	62.7(5.1)
K_{FFS}	2.33(22)	1.36(46)		
Q	0.125(24)	0.42(22)	$[G(0)-1]^{-1}$	1.65(01)

[a] Paramters correspond to the fitting analysis in Figure 1 (100mM NaCl). [b] Values in parentheses are the standard deviation for the last digits. [c] Values underlined are calculated using fitting parameters.

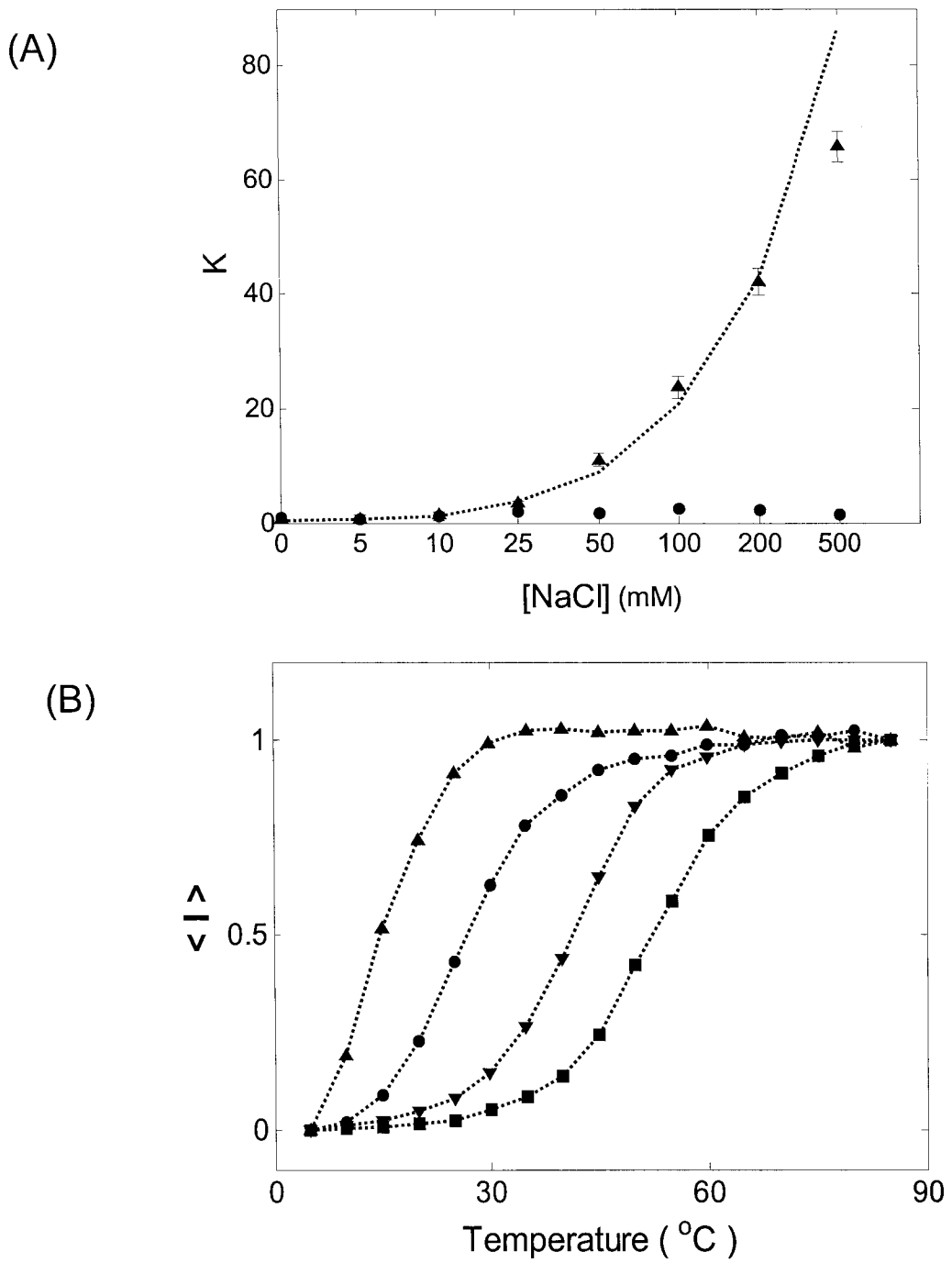

Figure 4.3 (A) Equilibrium constants of DNA hairpin samples vs. NaCl concentration and (B) corresponding melting profiles [data sets with NaCl concentrations of 0- (▲), 25- (●), 100- (▼), and 500-mM (■) are shown.] In panel (A), K_{melt} (▲) represents the equilibrium constants evaluated from the melting curves according to Equation (4.12). K_{FFS} (●) represents the equilibrium constants determined from our FCS and PCH analysis. The dotted line in panel (A) is $K_{melt,3S}$ calculated according to Equation (4.25).

rises precipitously with added NaCl.

Our initial goal was to simply characterize this behavior using our fluorescence fluctuation spectroscopy data itself, without reference to the DNA melting curves. The correlation amplitude parameters $\bar{N}_{CC}$ and B_{eq} contain information about the reaction equilibrium. In particular, B_{eq} has the form:

$$B_{eq} = K_{FFS}\left[\frac{1-Q}{1+K_{FFS}Q}\right]^2 = \frac{\bar{N}_1\bar{N}_2(\varepsilon_1-\varepsilon_2)^2}{(\bar{N}_1\varepsilon_1+\bar{N}_2\varepsilon_2)^2} \tag{4.13}$$

K_{FFS} denotes the equilibrium constant as determined by our fluctuation spectroscopy measurements; ε_1 and ε_2 are the specific brightnesses of R6G when the DNA hairpin molecules are in their open and closed conformations, respectively, and Q is the ratio $\varepsilon_2/\varepsilon_1$. The B_{eq} parameter is sometimes used to estimate K_{FFS} based on the assumption that $Q \approx 0$. However, this is often a very severe approximation that must be used with great care. As shown in Figure 4.4A, when $Q > \sim 0.1$, the measured value of B_{eq} can support a broad range of K_{FFS} values. B_{eq} essentially loses its dependence on K_{FFS} under these conditions. It is also seen that B_{eq} only takes on values greater than 1 when $Q << \sim 0.1$. Our B_{eq} measurements (Figure 4.4B) are in the regime where $B_{eq} < \sim 1$, and $Q > \sim 0.1$. This means that K_{FFS} and Q cannot be resolved using our FCS measurements alone. Additional information is needed to resolve these parameters.

Our motivation for combining PCH with FCS was to overcome this limitation in the study of biomolecule folding reactions. The PCH is a probability distribution for detecting k photons arising from the optical detection volume during a specified sampling

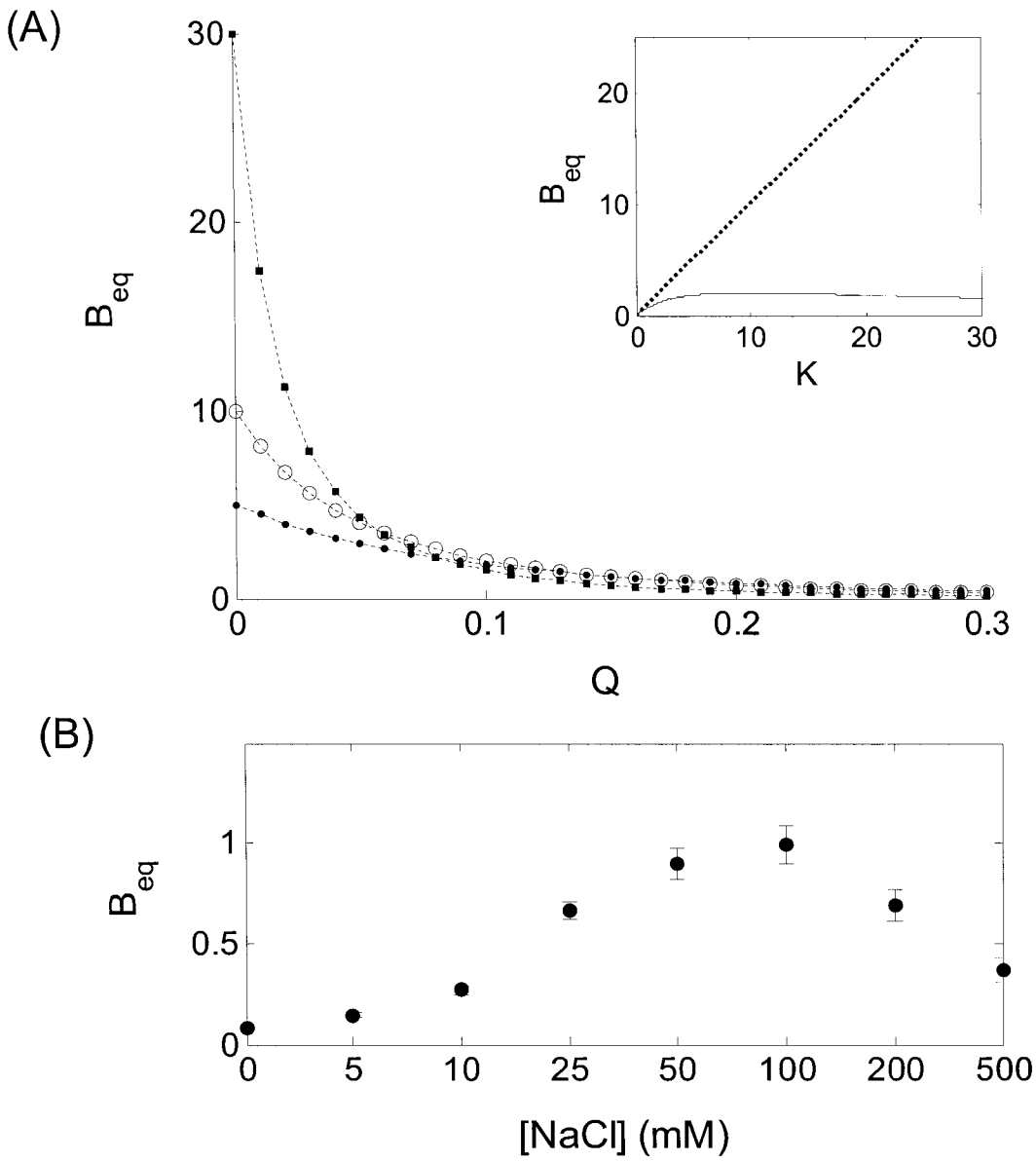

Figure 4.4 (**A**) Simulations of the amplitude term B_{eq} as function of Q and K (**inset**), and (B) experimental B_{eq} values from FCS analysis of DNA hairpin samples Panel (A) shows three equilibrium constants (•: K=5, ○: K=10, ■: K=30). Note that at Q=0, $B_{eq} = K$. As Q deviates from 0, the values of B_{eq} become considerably different from K. For instance, the difference between cases of Q=0 (dotted line) and Q=0.1 (solid line) is shown in the inset. According to observed values in panel (**B**), the expected Q is greater than ~0.1 in our experiments.

interval.[18, 19] The probability depends on the average number of fluorescent molecules being probed by the optical microscope and their specific brightnesses. This is precisely the information needed to resolve the K_{FFS} and Q terms contained in the B_{eq} parameter. Figure 4.2B shows photon counting histograms obtained from the same DNA samples represented by the corresponding FCS data in Figure 4.2A. We used sampling intervals of 9-μs, chosen to be small compared to the relaxation time of the DNA hairpin folding reaction. Hence, the system is considered static on the time scale of the PCH counting interval. Two approaches for obtaining the desired information from this data were investigated. To simplify matters, we note that ε_1 is the specific brightness of unquenched R6G, which we have measured separately by PCH analysis of R6G labeled DNA in the absence of quencher (See Suppl. Info.). The challenge becomes one of determining $\bar{N}_1$, $\bar{N}_2$, and ε_2.

The first PCH analysis approach we investigated was to determine $\bar{N}_1$, $\bar{N}_2$, and ε_2 directly by fitting the observed PCH to a two-species model of the probability distribution. The model we used was devised by Perroud *et al.*[20], who devised a way to correct for the out-of-focus emission observed whenever one-photon fluorescence excitation is employed. For a one-component system, the probability for detecting k photon counts from $\bar{N}$ identical molecules with specific brightness, ε, is given by

$$P(k;\bar{N},\varepsilon)=\sum_{N=0}^{\infty} p^{(N)}(k;\mathrm{q},\varepsilon)\cdot Poisson(N,\mathrm{q}\bar{N}) \tag{4.14}$$

where

$$p^{(1)}(k;\mathrm{q},\varepsilon)=\frac{1+\mathrm{F}_2}{(1+\mathrm{F}_1)^2}\left[p_G^{(1)}(k;\mathrm{q},\varepsilon)+\frac{1}{\mathrm{q}k!}\sum_{j=k}^{\infty}\frac{(-1)^{j-k}}{(j-k)!(2j)^{3/2}}\varepsilon^j F_j\right] \tag{4.15}$$

and, according to Chen *et al.*[18]

$$p_G^{(1)}(k;V_0,\varepsilon)=\frac{\pi\omega_0^{\,2}z_0}{V_0k!}\int_0^{\infty}\gamma(k,\varepsilon e^{-2x^2})dx \tag{4.16}$$

where, $\gamma(a,x)$ is the incomplete gamma function. q is the ratio between an ideal Gaussian shaped detection volume defined in FCS, V_{FCS}, and a reference volume, V_0 (*i.e.* V_0 = qV_{FCS}). It is found that the probability distribution $P(k;\bar{N},\varepsilon)$ is independent of the choice of V_0, as long as V_0 is large enough to obtain a positive value of $P(0;\bar{N},\varepsilon)$.[18, 24] We found q=6 to be an appropriate choice for our experiment. A first-order correction, in which $F_j = F$ for $k = 1$, and $F_j = 0$ for $k > 1$ was used. F is the ratio between out-of-focus and in-focus emission. For our optical setup, F is found to be ~0.7, in good agreement with the PCH measurements of Huang *et al.* using a similar optical setup.[24] The probability distribution for a two-component system is obtained from the above equations using $P(k;N_1,\varepsilon_1,N_2,\varepsilon_2)=P(k;N_1,\varepsilon_1)\otimes P(k;N_2,\varepsilon_2)$

$$P(k;\bar{N}_1,\varepsilon_1,\bar{N}_2,\varepsilon_2)=P(k;\bar{N}_1,\varepsilon_1)\otimes P(k;\bar{N}_2,\varepsilon_2) \tag{4.17}$$

Our PCH measurements could be analyzed by fitting to Equation (4.17). The parameters $\bar{N}_1$, $\bar{N}_2$, ε_2, and F, were determined, with the parameter ε_1 held constant (See Table 4.1).

An alternative approach to the PCH analysis was to apply a single-species model that depended on an *apparent* number of molecules, $\bar{N}_{app}$, with *apparent* specific brightness, ε_{app}.[38, 39] $\bar{N}_{app}$ and ε_{app} are weighted averages characterizing the distribution of fluorescent species of differing brightness present in the sample. For a two-component system, these parameters are given by the Equations (4.18-4.19)

$$\bar{N}_{app}=\frac{\left(\bar{N}_1\varepsilon_1+\bar{N}_2\varepsilon_2\right)^2}{\bar{N}_1\varepsilon_1^{\,2}+\bar{N}_2\varepsilon_2^{\,2}} \text{ and} \tag{4.18}$$

$$\varepsilon_{app} = \frac{\bar{N}_1 \varepsilon_1^{\,2} + \bar{N}_2 \varepsilon_2^{\,2}}{\bar{N}_1 \varepsilon_1 + \bar{N}_2 \varepsilon_2} \quad . \tag{4.19}$$

Our PCH data could be analyzed by fitting to the single-component distribution (Equation (4.14)) with $\bar{N} = \bar{N}_{app}$ and $\varepsilon = \varepsilon_{app}$. The parameters $\bar{N}_{app}$, ε_{app}, and F were obtained from the analysis.

Table 4.1 compares the PCH parameters obtained using the different analysis procedures for a representative set of data. It is found that the results are essentially equivalent in that they give the same values for $\bar{N}_{app}$, ε_{app}, and F. They also agree with the corresponding parameters derived from FCS within experimental error. In particular, we notice that $\bar{N}_{app}$ has the same functional form as the inverse amplitude of the autocorrelation function, $\left[G_{A,2S}(0)-1\right]^{-1}$, and we find good consistency between these parameters for both analysis procedures. The advantage of the two-component PCH analysis is that the equilibrium constant, K_{FFS}, can be determined directly from the PCH data alone. The disadvantage is that four adjustable parameters are needed to fit the data, which means the parameters were not as well determined as one would wish (~20- to 30-% relative standard deviations for $\bar{N}_1$ and $\bar{N}_2$). The single-component analysis requires fewer adjustable parameters, so the parameters were determined more precisely (<10-% relative standard deviations). However, K_{FFS} could not be obtained directly because the one-component PCH analysis did not provide enough information to uniquely determine $\bar{N}_1$ and $\bar{N}_2$. We found the best precision for determining K_{FFS} was attained by combining the information from the one-component PCH with the FCS results. Using Equations (4.13) and (4.18-4.19), along with the known value of ε_1, we obtain

$$\bar{N}_1 = \frac{\bar{N}_{app}\varepsilon_{app}^{\ 2}}{\left[\varepsilon_1^{\ 2} + (\varepsilon_1 - \varepsilon_{app})^2 / B_{eq}\right]} \text{ and} \tag{4.20}$$

$$K_{FFS} = \frac{\bar{N}_{CC} - \bar{N}_1}{\bar{N}_1} \tag{4.21}$$

Table 4.1 shows the values of $\bar{N}_1$, $\bar{N}_2$, and K_{FFS} derived from the above equations for a representative sample.

4.3.2 Failure of the Two-State Reaction Mechanism

In the past, it was assumed that the complete DNA hairpin folding reaction and the sub-millisecond chemical relaxation processes observed by FCS represented one and the same reaction.[7, 8, 12-15] If this were the case, then K_{melt} and K_{FFS} would describe the same reaction process. We are now in a position to evaluate that assumption. Figure 4.3A compares the behavior of K_{melt} and K_{FFS} as a function of NaCl concentration. There is good agreement between the two sets of constants at NaCl concentrations below ~25-mM, at which the open form of the DNA hairpin is favored at our laboratory temperature. At higher concentrations, however, the deviation becomes quite dramatic. The melting curves predict a precipitous rise in the equilibrium constant with increasing NaCl concentration as already discussed; whereas, the equilibrium constant determined by PCH and FCS stays relatively constant. According to our observations, the assumption that K_{melt} and K_{FFS} describe the same reaction is only valid under conditions that favor the open form of the hairpin. When the closed form of the hairpin is favored, K_{melt} and K_{FFS} represent different reactions. This is our first indication that the DNA hairpin folding reaction may be more complicated than previously recognized.

Additional insight come from the behavior of the FCS and PCH parameters themselves. Figure 4.5A compares the observed behavior of $\bar{N}_{app}$, ϵ_{app}, $[G_A(0)-1]$ (the autocorrelation amplitude), $\langle I \rangle$ (the average photocount rate), and $\bar{N}_{CC}$ with their expected behavior assuming a two-state reaction characterized by the equilibrium constant, K_{melt} (dotted curves). The x-coordinates correspond to $K_{melt}/(K_{melt} + 1)$ for each NaCl concentration, which represents the fraction of folded DNA hairpins that would be observed in a two-state reaction. The expected behaviors of these parameters were calculated using the fraction and the measured values of ε_1 and ε_2. (See Supple. Info. for the details of calculation.) These trends show that, in a two-state reaction $\bar{N}_{app}$ should slowly decrease to a minimum value, while $[G_{A,2S}(0)-1]$ should have the inverse of this trend. ϵ_{app} should decrease as the brighter species is depleted in favor of the darker species, and $\bar{N}_{CC}$ should remain constant. What we have observed is diametrically opposed to these expectations. At low NaCl concentrations, all the parameters are in good agreement with their expected values. But as NaCl is added, $\bar{N}_{app}$ drops precipitously, as $[G_A(0)-1]$ rises inversely. $\bar{N}_{CC}$ abruptly drops, while ϵ_{app} remains relatively constant.

Naturally, we wondered if this apparently anomalous behavior could be due to experimental error or to misinterpretation of our measurements. To address this possibility, we constructed a model two-state system[40], and subjected it to the same set of measurements as described above. The model system consisted of solutions containing poly(dT)$_{40}$ ssDNA, singly labeled with either R6G- or Cy3- and mixed together to achieve different fluorophor ratios. PCH analysis of pure component solutions was used to measure the molecular brightnesses, ε_{R6G} and ε_{Cy3}. ε_{R6G} was identical to ε_1, presented

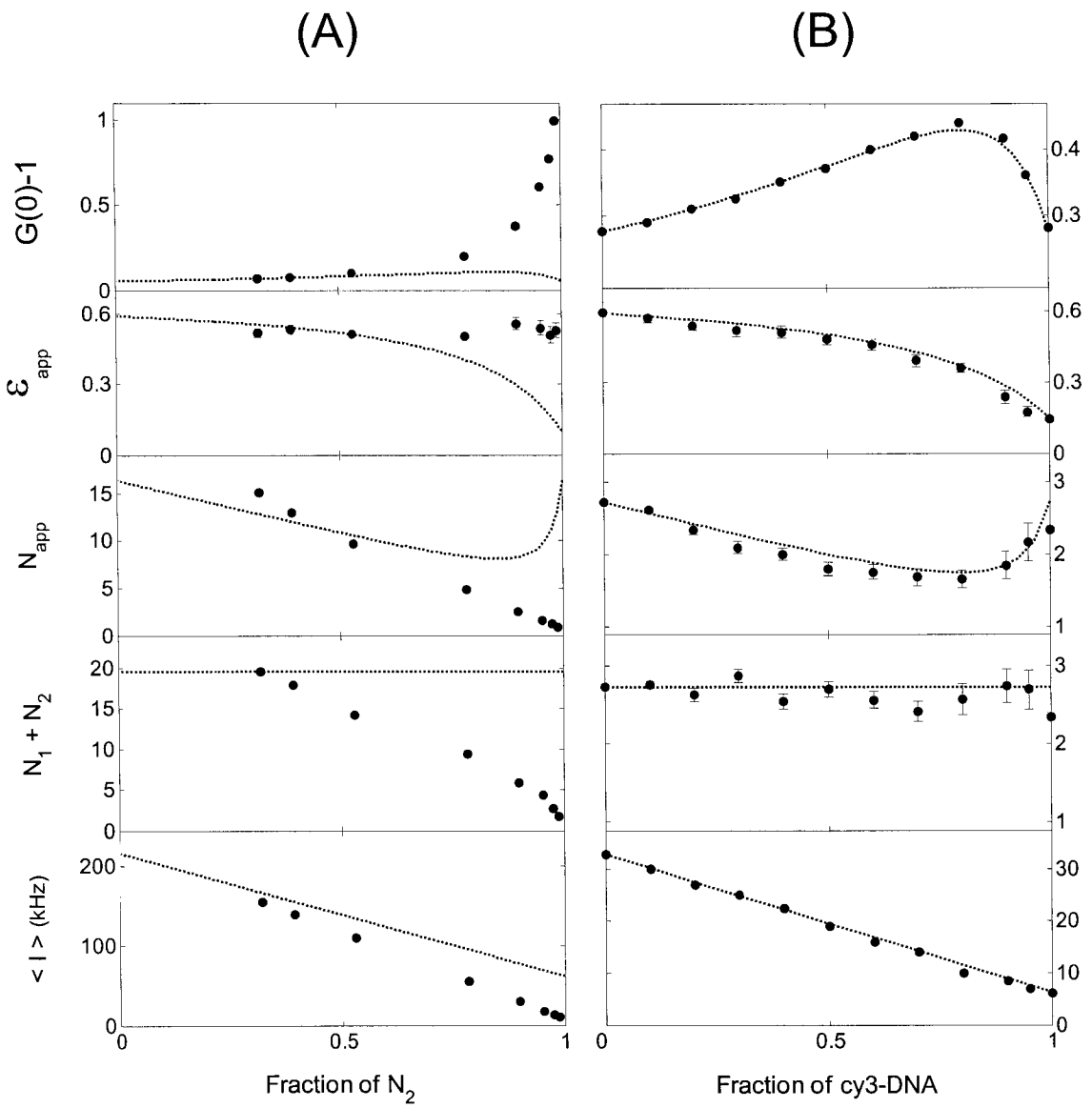

Figure 4.5 FCS and PCH parameters of DNA hairpin samples (**A**) and a model two-state system (**B**) The DNA hairpin samples contained ~25 nM DNA concentrations and varying concentrations of NaCl. The model system contained Cy3- and R6G-singly labeled poly(dT)$_{40}$ single-stranded DNA prepared in different ratios. The total concentration of Cy3 and R6G were held constant at ~4.5 nM. In panel (**A**), the *x*-coordinate represents the expected fraction, $f = K_{melt}/(1+K_{melt})$, of DNA molecules in the folded conformation, assuming a two-state reaction. $\bar{N}_1 + \bar{N}_2$ was obtained from the cross-correlation analysis. In panel (**B**), $\bar{N}_{app}$ and ε_{app} were evaluated using fitting parameters from the two-species PCH analysis according to Equations (4.18-4.19). The expected behaviors of the parameters, assuming a two-state system for each, are represented by dotted lines in both panels. (See Supple. Info. for details)

above. ε_{Cy3} was measured under the same experimental conditions. Hence, the ratio $\varepsilon_{Cy3}/\varepsilon_{R6G}$ is ~0.25 and comparable to our Q value measured for the DNA hairpins.(See Supple. Info.) Figure 4.5B shows how $\bar{N}_{app}$, ε_{app}, $[G_A(0)-1]$, $\langle I \rangle$ and $\bar{N}_{Cy3} + \bar{N}_{R6G}$, as determined by FCS and PCH analysis, changed with increasing fraction of Cy3-labeled ssDNA. The concentrations of the different components were adjusted so that the total DNA concentration remained constant. These parameters behave in precisely the way one would expect for a two-state system. What's more, these observations are completely analogous to the expected behavior of DNA hairpin molecules undergoing a two-state reaction. We conclude that our experiment is robust in its ability to characterize this type of behavior.

We also considered the possibility of detector saturation artifacts skewing the results for higher count rate samples, as well as changes in the background photocount rate due to added NaCl. Detector saturation artifacts could be investigated by carrying out the same measurements on samples in which the DNA hairpin concentrations were adjusted to achieve the maximum count rate observed in our experiments for each sample. The same was done for the model two-state system. The model system exhibited precisely the behavior that would be expect for a two-state system, while the DNA hairpin samples exhibited behavior that was anomalous for a two-state system (See Suppl. Info.). Also, analysis of solvent blanks vs. NaCl concentration showed no change in the background photocount rate. It is concluded that if the DNA hairpins were reacting according to a two-state mechanism, this would be manifested in the behavior of the FCS and PCH parameters. The fact that the FCS and PCH parameters deviate so strongly from

the expected behavior leads us to consider alternative reaction mechanisms for DNA hairpin folding.

4.3.3 The Three-State Reaction Mechanism

The observations presented thus far can be explained in a straightforward manner if we make the following assumptions. First, it is apparent that the concentration of DNA molecules detectable by PCH and FCS decreases as the DNA hairpin folding reaction progresses. This is evident from the way $\bar{N}_{app}$, ϵ_{app}, $\bar{N}_{CC}$, and the correlation amplitudes are affected by the NaCl concentration, and it suggests that some of the DNA is being converted into a non-fluorescent dark state that does not contribute to the FCS and PCH measurements. Second, it is apparent that this dark state is being formed, and is stable, on a longer time scale than the FCS correlation time. Our FCS experiment can only analyze chemical reactions that occur on a faster time scale than the molecular transport time through the optical detection region. Anything taking place on a longer time scale than ~1-ms is static with respect to FCS. What our results suggest is that the undetectable dark state is static on the FCS time scale and is thus not taking part in the sub-millisecond chemical relaxation process observed by FCS. As far as our FCS and PCH measurements are concerned, this non-fluorescent dark state is a sink into which DNA molecules disappear and never return.

This implies that the DNA hairpin folding reaction is not going to completion on the time scale probed by FCS, as has previously been thought, and that some other reaction is responsible for the observed sub-millisecond relaxation process. How can we be sure this is the case? In fact, *all* our observations point to this key assertion, but it can

be explained most clearly with reference to the way the cross-correlation amplitude rises with added NaCl. As already discussed, sub-millisecond chemical relaxation processes observed by FCS do not contribute to the cross-correlation amplitude because they occur too rapidly to remain correlated on the FCCS time scale. Hence, the cross-correlation amplitude is inversely proportional to the *total* concentration of all species taking part in the relaxation. If the observed relaxation represented the complete hairpin folding reaction, then the overall concentration of DNA species involved in the reaction would not change on the FCS time scale. Consequently, the cross-correlation amplitude would remain essentially constant as the reaction progressed (Actually, there would be a slight decrease in the amplitude due to the dependence on τ_D). The only way we can explain the dramatic rise in the cross-correlation amplitude with added NaCl is to assume that the sub-millisecond chemical relaxation observed by FCS does *not* represent the complete folding reaction. Rather there must be at least one intermediate step along the way. It is this intermediate reaction that causes the observed chemical relaxation, not the complete folding reaction. As the folding reaction progresses, the DNA species taking part in the observed chemical relaxation are depleted, lowering the total concentration of detectable DNA, raising the cross-correlation and auto-correlation amplitudes, and lowering the apparent number of molecules probed by PCH. At present, we cannot rule out the possibility that the complete reaction may be occurring on a similar time scale to the cross-correlation measurements (~1-ms to ~10-ms) or longer, but it is clearly not occurring any faster than this.

The simplest reaction mechanism that supports our observations is the three-state reaction:

$$\bar{N}_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \bar{N}_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \bar{N}_3 \,, \tag{4.22}$$

$$K_1 = \frac{\bar{N}_2}{\bar{N}_1}, \quad K_2 = \frac{\bar{N}_3}{\bar{N}_2} \quad . \tag{4.23}$$

$\bar{N}_1$ still refers to the completely unfolded DNA conformation for which the R6G fluorescence is unquenched. $\bar{N}_2$ now refers to a reaction intermediate that is stable on the sub-millisecond time scale of the FCS experiment. $\bar{N}_3$ is the completely folded DNA hairpin. We believe this final state represents a dark state in which the R6G fluorescence is efficiently quenched and undetectable by PCH and FCS. The auto- and cross-correlation functions for this three-state reaction mechanism are given by the Equation (4.24) (See Supple. Info. for a complete derivation.):

$$G_{3S}(\tau) - 1 = \frac{1}{\bar{N}_{total}} \left(1 + B_1 e^{\lambda_2 \tau} + B_2 e^{\lambda_3 \tau}\right) g_D(\tau) g_T(\tau) g_F(\tau) \tag{4.24}$$

where $\bar{N}_{total}$ is the total concentration of DNA in all its forms ($\bar{N}_{total} = \bar{N}_1 + \bar{N}_2 + \bar{N}_3$); B_1, B_2, λ_2, and λ_3 are different functions of the rate and equilibrium constants; and $g_D(\tau)$, $g_T(\tau)$, and $g_F(\tau)$ are defined in Equations (4.6-4.7), (4.9) and (4.11).

Our observations require that k_1, $k_{-1} >> k_2$, k_{-2}. Accordingly, the chemical relaxation process observed by our FCS experiment arises from the two-state $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction; whereas, the $\bar{N}_2 \rightleftharpoons \bar{N}_3$ reaction occurs on a time scale not observable by FCS (*i.e.* slower than the diffusion and flow time of molecules through the optical probe region). This implies that the DNA molecules fluctuate between open and intermediate configurations many times before the correctly folded stem-loop hairpin structure is fully formed. Under these conditions, $G_{3S}(\tau)$ takes on approximately the same functional forms

as $G_{2S}(\tau)$ in Equations (4.5) and (4.8) (See Suppl. Info.). The key differences are that $\bar{N}_2$ now refers to the reaction intermediate, K_{FFS} refers to K_1, τ_R is the relaxation time of the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction, and ε_2 is the specific brightness of R6G in the intermediate state. The equations for $\bar{N}_{app}$ and ϵ_{app} (Equations (4.18-4.19)) are redefined similarly.

All of the observations presented above can now be explained. K_{FFS} and K_{melt} exhibit different dependences on NaCl because K_{FFS} only describes the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction; whereas, K_{melt} describes the complete folding reaction (It should be noted that Equation (4.12) assumes a two-state reaction mechanism, so K_{melt} may not represent the actual equilibrium constant for the complete reaction). At low NaCl concentrations, the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction dominates, so K_{FFS} and K_{melt} are essentially equivalent. As NaCl is added, the $\bar{N}_2 \rightleftharpoons \bar{N}_3$ reaction becomes dominant, such that K_{melt} rises, while K_{FFS} changes more slowly. NaCl apparently does not exert as large an influence on the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction equilibrium as it does on the complete folding reaction. The behavior of $\bar{N}_{app}$, ε_{app}, $[G_A(0)-1]$, and $\bar{N}_{CC}$, can also be explained. The FCS and PCH measurements are only sensitive to $\bar{N}_1$ and $\bar{N}_2$. At low NaCl concentrations, $\bar{N}_1$ and $\bar{N}_2$ are the dominant species, so the PCH and FCS parameters behave as expected for a two-state reaction. As the reaction progresses, $\bar{N}_1$ and $\bar{N}_2$ are depleted, which causes the observed drop in $\bar{N}_{app}$ and $\bar{N}_{CC}$, and the corresponding rise in the correlation amplitudes. Also, because K_{FFS} changes slowly with added NaCl, the change in the relative concentrations of $\bar{N}_1$ and $\bar{N}_2$ is correspondingly slow, which explains why ε_{app} is not affected by the addition of NaCl.

This three-state reaction can be quantified in the following way. At 0-M NaCl, it is assumed that only $\bar{N}_1$ and $\bar{N}_2$ are appreciable.[37] Hence, the total amount of DNA for all the samples is approximately $\bar{N}_{total}$ = $\bar{N}_1$ (0-M NaCl) + $\bar{N}_2$ (0-M NaCl), which is measured by FCS and PCH. The same procedure can then be used to measure $\bar{N}_1$ (x-M NaCl) and $\bar{N}_2$ (x-M NaCl) for all subsequent samples. $\bar{N}_3$ (x-M NaCl) is obtained for each sample using $\bar{N}_3 = \bar{N}_{total} - \left(\bar{N}_1 + \bar{N}_2\right)$. Figure 4.6 shows the results of these measurements. Our analysis can also explain the observed melting curves. We define a parameter $K_{melt,3S}$ that can be calculated for each sample using data from our FCS and PCH analysis for comparison with the K_{melt} values obtained directly from the melting curves. By analogy to Equation (4.12), we have:

$$K_{melt,3S,i} = \frac{\langle I \rangle_{\max} - \langle I \rangle_i}{\langle I \rangle_i - \langle I \rangle_{\min}} \tag{4.25}$$

where $\langle I \rangle_i$ is the average fluorescence count rate observed in our FCS and PCH experiment for a given NaCl concentration. $\langle I \rangle_{\max}$ and $\langle I \rangle_{\min}$ are the fluorescence count rates that would be observed from samples in which all the DNA was present in a completely open or a completely closed form, respectively. (See Supple. Info. for the details of calculation) . As seen in Figure 4.3A, the $K_{melt,3S}$ and K_{melt} values are in remarkably good agreement.

4.4 DISCUSSION

Previous studies on DNA hairpin formation have assumed a folding free energy landscape comprising two potential minima, corresponding to the random coil and the

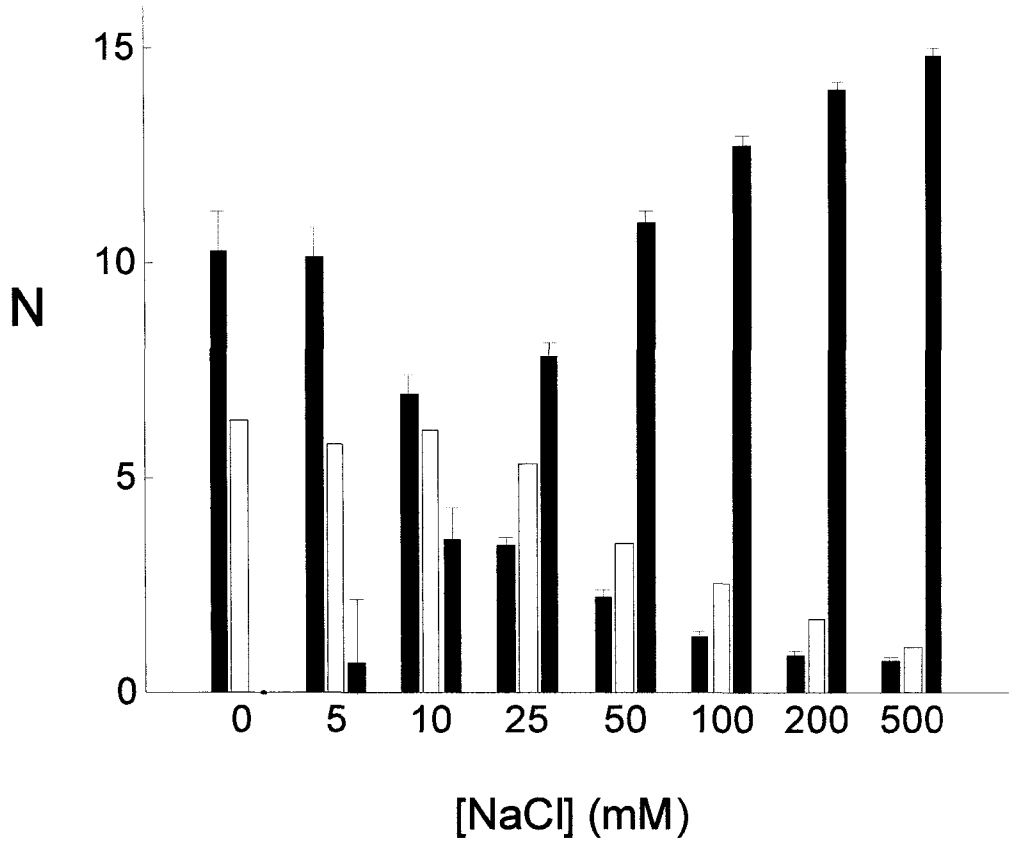

Figure 4.6 Equilibrium distribution of the three DNA hairpin conformations ($\bar{N}_1$: blue, $\bar{N}_2$: green, $\bar{N}_3$: red) determined by the three-state reaction mechanism. The unfolded and intermediate forms of the DNA hairpin are favored at low NaCl concentration, while the folded form is favored at high NaCl concentration.

fully folded stem-loop hairpin, and a transition state ensemble of partially folded or misfolded intermediates.[6-9, 11-16] These intermediates states cause roughness of the free energy surface, but are short-lived on the time scale of the folding reaction. Recent atomistic molecular dynamics simulations identified some plausible intermediate structures in the unfolding of a small hairpin forming RNA molecule that may be present in the transition state ensemble.[41, 42] These simulations, together with a statistical mechanics model developed by Ansari *et al.*[6, 9, 11], suggest a heterogeneous transition state ensemble consisting of collapsed hairpin-like structures in which the loop is fully or partially formed and the stem is held together by single native or nonnative base-pair contacts that are only partially hydrogen bonded.

Our observations suggest the DNA hairpin folding mechanism may need to be expanded to accommodate longer-lived reaction intermediates that do not lead directly to the completely folded stem-loop hairpin structure. Candidates for such species can also be found in the simulation studies of Sorin *et al.*[41, 42] From large-scale, parallel molecular dynamics simulation, which involves sampling a large number of constant temperature trajectories that total more than 500-μs of simulations for an all-atom model of RNA hairpin 5'-GGGC[GCAA]GCCU-3', with continuum representation of solvent effects. To investigate the folding process at 300K, the simulations were started from the fully extended, denatured state. The majority of the folding trajectories observed by these authors terminated in "collapsed" structures containing native-like hairpin loops, but with stem structures that were misfolded in a variety of ways. This "collapsed ensemble" was distinct from the transition state ensemble in that a much longer reaction time (~8-μs) for

forming the collapsed structures was observed, indicating higher stability of these structures compared to those of the transition state.

We propose that the DNA folding reaction being probed by our FCS experiment corresponds to the formation of a collapsed ensemble of long-lived intermediates analogous to the ones observed in the molecular dynamics simulations of Sorin *et al.*[41, 42] This seems consistent with the observations that the simulated collapse occurred on the same time scale as the experimentally observed reaction time for a hairpin of this size, and that most of the simulated folding reactions resulted in these types of structures. It may explain why we observe a finite value of ε_2 in our PCH measurements. Misfolding of the stem sequence may give rise to partial or intermittent contact between the R6G fluorophor and the dabcyl quencher, thereby lowering the fluorophor-quencher energy transfer efficiency. The formation of the completely folded stem structure occurs on a time scale that is too slow to be observed by FCS. Yet once formed, this structure is highly stable on the FCS time scale. The correctly folded hairpin structure results in efficient quenching of the R6G fluorescence, which is why this form of the DNA appears to us as a sink into which DNA molecules disappear from view.

There are some sets of experimental observations reported in the literature that may be interpreted in a way that supports this view of DNA hairpin folding. Grunwell, *et al.* used single-pair fluorescence resonance energy transfer (FRET) to monitor the opening and closing of individual immobilized DNA hairpin molecules that caused abrupt transitions between high and low FRET efficiency states.[43] The experiment was sensitive to the tens of milliseconds time scale, but not to the sub-millisecond time scale probed by FCS. The authors studied a larger sized hairpin structure with a longer stem,

and with adenine residues in the loop instead of thymine. Nonetheless, the opening and closing times they observed were more than an order of magnitude longer than expected, based on extrapolating solution phase kinetics measurements to this larger hairpin size. These unusually long reaction times have generally been attributed to interactions of the DNA molecules with the surface. However, our observations suggest an alternative way of interpreting these measurements. The high FRET state corresponds to the fully formed stem-loop hairpin structure; whereas, the low FRET state may be averaged over the open state and an intermediate state. If our proposed mechanism is correct, much longer transition times in and out of the high FRET state are to be expected. Fluctuations between the open state and the intermediate state may be below the time resolution of the spFRET experiment, and would thus not be observed. Another observation can be found in the Liphardt et al.'s work.[44] They have used mechanical force to induce the unfolding and refolding of single RNA molecules, including a simple RNA hairpin, a molecule containing a three-helix junction, and a domain of a ribozyme. By imposing a constant force on the molecule, they were able to monitor the end-to-end distance between the two ends of the hairpin and to watch the distance hop back and forth between two values characteristic of the fully unfolded and the fully folded hairpin. They determine the folding and unfolding rates from the average lifetimes in the two states, and found that, at the critical force for which the opening and closing rates are same (~14pN in the presence of 10mM Mg^{2+}), the folding times are ~1s. The very slow folding times observed in these measurements, compared to the folding times of tens of microseconds observed in the FCS and T-jump measurements has been explained as effects of applied force and longer stem. However, based on our observations, we are

interpreting these measurements in other way. The disparity between these single molecule experiments and the solution phase kinetics experiments may simply be due to the fact that the two types of experiments are probing different aspects of the folding reaction. The solution phase kinetics experiments are probing the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction; whereas, the single molecule experiment is probing the $\bar{N}_2 \rightleftharpoons \bar{N}_3$ reaction.

How do our observations impact the conclusions drawn from previous kinetics experiments regarding the energetics of DNA hairpin loop formation? On the one hand, the reactions being probed are still considered to be loop-forming reactions and are therefore valid models from which to gain insight into DNA hairpin folding. However, many of the conclusions drawn from these experiments are based on the temperature and loop size dependence of the rate constants for hairpin loop formation (k_1 in Equation (4.4)). These constants were calculated using the measured reaction times, combined with equilibrium constants derived from analysis of the melting curves. We have shown that when conditions favor the closed form of the hairpin, the equilibrium constant derived from the melting curve does not describe the reaction being probed by the kinetics experiment. Hence, rate constants determined in this way may be in error. Just how severe might these errors be? For our DNA hairpin sample under conditions of 100-mM NaCl, we obtain $k_1 = 1.12(\pm 0.16)\times 10^4$ s^{-1}, based on our τ_R and K_{FFS} measurements for this sample. If we use K_{melt} in our calculation, we obtain $k_1 = 1.53(\pm 0.21)\times 10^4$ s^{-1}. Although the use of K_{melt} may overestimate the rate constant somewhat, k_1 does not have a strong dependence on the equilibrium constant when $K > 1$. When $K < 1$, the reactions probed by the kinetics experiments are the same as the ones described by the melting

curves. Hence, we believe conclusions drawn from the behavior of k_1 can still be supported.

However, conclusions regarding the reverse reaction are a different matter, because the k_{-1} rate constant depends much more strongly on the reaction equilibrium when $K > 1$. For example, our 100-mM NaCl sample gives $k_{-1} = 4.79(\pm 0.50)\times 10^3\ s^{-1}$ and $k_{-1} = 6.46(\pm 0.74)\times 10^2\ s^{-1}$ using K_{FFS} and K_{melt}, respectively. One area of controversy in the literature on the unfolding of DNA hairpins has to do with the loop size dependence of k_{-1}. Some studies have found that k_{-1} has no dependence on loop size.[8] Other studies show an increase in k_{-1} with loop size, and still others show a decrease.[11, 15] We believe these conflicts come from the fact that some experiments were carried out under conditions that favored the closed form of the hairpin, while others were carried out under conditions in which the open form was favored. Conclusions drawn from the k_{-1} parameter are only valid when conditions favor the open structure. Even then, we believe the reaction described by k_{-1} pertains to the dissociation of a misfolded stem structure, not the unwinding of a fully intact set of base pairs. Works with different stem size-(different number of base pairs) and loop size-DNA hairpin are currently under investigation in our group.

4.5 CONCLUSION

In this Chapter, we present for the first time to our knowledge, its application to the study of biomolecule conformational fluctuation between different species, which causes the changes of molecular brightness. We have shown that more detailed information about the DNA hairpin folding reaction can be obtained by employing a

fluorescence fluctuation spectroscopy approach that combines auto-correlation, cross-correlation, and PCH measurements. These measurements support a three-state reaction mechanism for DNA hairpin formation that consists of a rapid equilibrium between open and intermediate forms of the DNA, and a fully folded form that is stable on the time scale of our FCS experiment.(Figure 4.7) Further insight into this proposed mechanism could be gained by carrying out these measurements on DNA hairpins with varying stem and loop sizes and compositions. For example, a smaller sized hairpin with a shorter stem sequence may exhibit folding times for both the intermediate reaction and the final folding reaction that could be measured on the FCS time scale simultaneously. Also, it is suggested that the experimental techniques described in this paper could be employed in the study of more complicated biomolecule folding reactions involving RNA, DNA aptamers, and proteins.

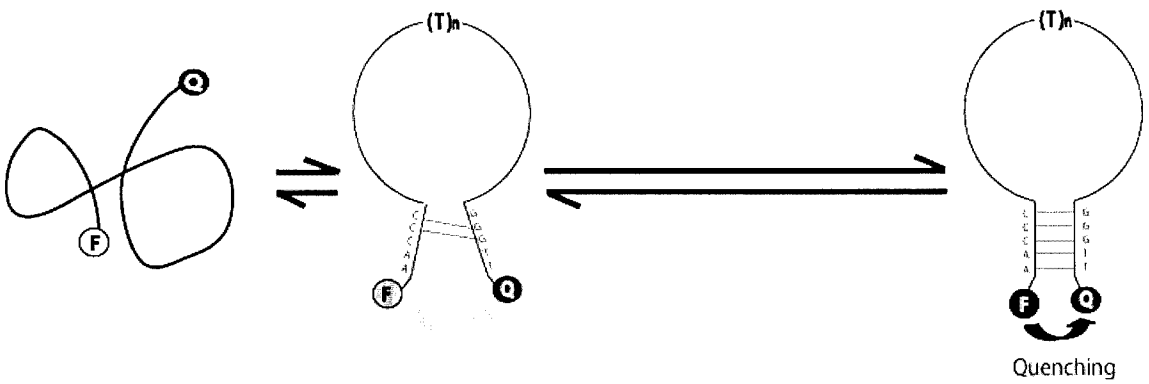

Figure 4.7 Schematic description of a three-state mechanism of DNA hairpin formation based on our observation

4.6 SUPPLEMENTARY IMFORMATION

4.6.1 Derivation of three-state FCS model

For a simple fluorescent molecule, such as poly(dT)$_{40}$ ssDNA labeled with a fluorescence dye, the fluorescence correlation function can be expressed as,

$$G^{\circ}(\tau)-1=\frac{1}{\bar{N}}g_D(\tau)g_T(\tau)g_F(\tau) \tag{4.26}$$

For a molecule experiencing another reaction between two states, such as fluorescence fluctuation caused by folding and unfolding conformational dynamics,

$$\bar{N}_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \bar{N}_2 \tag{4.4}$$

the whole correlation function can be expressed as[30, 32, 34]

$$G(\tau)-1=\left(G^{\circ}(\tau)-1\right)g_R(\tau)\ . \tag{4.27}$$

The conformational relaxation term $g_R(\tau)$ is,

$$g_R(\tau)=1+B_{eq}\exp\left(-\tau/\tau_R\right), \tag{4.10}$$

$$\text{where } B_{eq}=K\left(\frac{1-Q}{1+KQ}\right)^2 \text{ and } K=\frac{\bar{N}_2}{\bar{N}_1}=\frac{k_1}{k_{-1}},\ Q=\frac{\varepsilon_2}{\varepsilon_1} \tag{4.28}$$

In the case of slow flow rate ($\tau_F >> \tau_R,\ \tau_T$), the $g_R(\tau)$ term as well as $g_T(\tau)$ term in cross correlation function are reduced to 1.[15]

Therefore, the cross-correlation function for our experimental condition can be expressed as

$$G_C(\tau)-1=\frac{1}{\bar{N}_1+\bar{N}_2}g_D(\tau)g_{Fc}(\tau), \text{ and} \tag{4.29}$$

the corresponding auto-correlation function can be,

$$G_A(\tau)-1=\frac{1}{\bar{N}_1+\bar{N}_2}\left(1+B_{eq}\exp(-\tau/\tau_R)\right)g_D(\tau)g_T(\tau)g_{Fa}(\tau) \tag{4.30}$$

We found the increase of cross correlation peak amplitude as function of NaCl concentration can not be explained by the $1/\left(\bar{N}_1+\bar{N}_2\right)$ term.

Now, as the simplest reaction mechanism that supports our observations, we consider the FCS model of DNA hairpins experiencing a conformational change between three states.

$$\bar{N}_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \bar{N}_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \bar{N}_3 \tag{4.22}$$

where,

$$K_1=\frac{\bar{N}_2}{\bar{N}_1}=\frac{k_1}{k_{-1}},\quad K_2=\frac{\bar{N}_3}{\bar{N}_2}=\frac{k_2}{k_{-2}},\quad Q_1=\frac{\varepsilon_2}{\varepsilon_1},\quad Q_2=\frac{\varepsilon_3}{\varepsilon_2} \tag{4.31}$$

We assume the diffusion coefficients of all species are equal. Also, we assume $\epsilon_3 \approx 0$, so that our detector can not see the $\bar{N}_3$ species.

In similar way, the whole correlation function can be expressed as $G(\tau)-1=\left(G(\tau)^\circ-1\right)g_R(\tau)$ and then, the conformational relaxation term $g_R(\tau)$ can be derived by solving the following first order linear differential equations and applying the procedure developed by Elson, E. and Magde, D [30-32]

$$\frac{d}{dt}\begin{pmatrix}\bar{N}_1\\ \bar{N}_2\\ \bar{N}_3\end{pmatrix}=\begin{bmatrix}-k_1 & k_{-1} & 0\\ k_1 & -(k_{-1}+k_2) & k_{-2}\\ 0 & k_2 & -k_{-2}\end{bmatrix}\begin{pmatrix}\bar{N}_1\\ \bar{N}_2\\ \bar{N}_3\end{pmatrix} \tag{4.32}$$

The eigenvalues of the differential equation matrix above are calculated as,

$$\begin{aligned}\lambda_1&=0\\ \lambda_{2,3}&=-\frac{(k_1+k_{-1}+k_2+k_{-2}\mp\sigma)}{2}\end{aligned} \tag{4.33}$$

where,

$$\sigma = \left[k_1^{\,2} + 2k_1k_{-1} - 2k_1k_2 - 2k_1k_{-2} + k_{-1}^{\,2} + 2k_{-1}k_2 - 2k_{-1}k_{-2} + k_2^{\,2} + 2k_2k_{-2} + k_{-2}^{\,2} \right]^{1/2} \quad (4.34)$$

The resulting correlation function is derived as,

$$g_R(\tau) = 1 + B_1 \exp(\lambda_2 \tau) + B_2 \exp(\lambda_3 \tau), \quad (4.35)$$

and the two amplitude terms are expressed as

$$B_{1,2} = \gamma \mp \delta, \quad (4.36)$$

$$\text{where, } \gamma = \frac{K_1\left((1-Q_1)^2 + (1+K_1Q_1^{\,2})K_2\right)}{2(1+K_1Q_1)^2}, \text{ and } \delta = \frac{-K_1(k_2\alpha + k_4\beta)}{2\sigma(1+K_1Q_1)^2} \quad (4.37)$$

$$\alpha = \left((A - K_1^{\,2}K_2)Q_1^{\,2} - 2(A + K_1K_2)Q_1 + A - K_2\right) \quad (4.38)$$

$$\beta = \left((A-2)Q_1^{\,2} + 2(1+K_2)Q_1 - (1+K_2)^2\right) \quad (4.39)$$

$$A = (1 + K_1 + K_1K_2) \quad (4.40)$$

$$\delta = \frac{-K_1\left((A - K_1^{\,2}K_2)Q_1^{\,2} - 2(A + K_1K_2)Q_1 + A - K_2\right)}{2(1+K_1)(1+K_1Q_1)^2}. \quad (4.41)$$

From the correlation function, we found some characteristic values such as auto-correlation amplitude, $[G_A(0)\text{-}1]$:

$$\begin{aligned} G_A(0) - 1 &= \frac{1 + B_1 + B_2}{\bar{N}_{tatal}} = \frac{1 + 2\gamma}{\bar{N}_1 + \bar{N}_2 + \bar{N}_3} \\ &= \frac{A(K_1Q_1^{\,2} + 1)}{(\bar{N}_1 + \bar{N}_2 + \bar{N}_3)(1 + K_1Q_1)^2} = \frac{\bar{N}_1\varepsilon_1^{\,2} + \bar{N}_2\varepsilon_2^{\,2}}{(\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2)^2} = \frac{1}{\bar{N}_{app}} \end{aligned} \quad (4.42)$$

if $Q_1 = 0$, so that, only $\bar{N}_1$ is detectable, the amplitude becomes

$$G_A(0) - 1 = \frac{A}{\bar{N}_{tatal}} = \frac{(1 + K_1 + K_1K_2)}{\bar{N}_{tatal}} = \frac{1}{\bar{N}_1}. \quad (4.43)$$

and the relaxation time for the whole reaction of conformational change:

$$k_R = k_1 + k_{-1} + k_2 + k_{-2} = -(\lambda_2 + \lambda_3) = \frac{1}{\tau_R} \tag{4.44}$$

where, $-\lambda_2 < -\lambda_3 < k_R$.

The whole correlation function has following functional form.

$$G(\tau) - 1 = \frac{1}{\bar{N}_{tatal}} \left(1 + B_1 \exp(\lambda_2 \tau) + B_2 \exp(\lambda_3 \tau)\right) g_D(\tau) g_T(\tau) g_F(\tau) \tag{4.45}$$

Now, we suggest that the second reaction is very slower process (i.e. k_2, $k_{-2} << k_1$, k_{-1}). In this case, $\lambda_2 \approx 0$, $\lambda_3 \approx -k_R = -(k_1 + k_{-1})$, and $g_R(\tau)$ becomes

$$g_R(\tau) = 1 + B_1 + B_2 \exp(-k_R \tau) \tag{4.46}$$

and the correlation functions become,

$$G(\tau) - 1 = (1 + B_1)\left(1 + \frac{B_2}{1 + B_1} \exp(-k_{R1} \tau)\right)\left(G^{\circ}(\tau) - 1\right) \tag{4.47}$$

$$G_C(\tau) - 1 = \frac{1 + B_1}{\bar{N}_{tatal}} g_D(\tau) g_{Fc}(\tau) = \frac{1}{\bar{N}_1 + \bar{N}_2} g_D(\tau) g_{Fc}(\tau) \tag{4.48}$$

$$G_A(\tau) - 1 = \frac{1}{\bar{N}_1 + \bar{N}_2} \left(1 + B_{eq} \exp(-k_R \tau)\right) g_D(\tau) g_T(\tau) g_{Fa}(\tau) \tag{4.49}$$

where, $B_{eq} = K_1 \left(\frac{1 - Q_1}{1 + K_1 Q_1}\right)^2$. (4.50)

Now, we find the cross correlation amplitude depends on $\bar{N}_1 + \bar{N}_2$, not on $\bar{N}_{tatal}$ ($= \bar{N}_1 + \bar{N}_2 + \bar{N}_3$), and can explain the increase of peak amplitude and all the behaviors of parameters observed in our FFS experiments.

4.6.2 Combined approach of PCH and FCS analysis

The photon counting histogram (PCH) analysis is based on the observation volume profile, $W(r)$. The probability of detecting k photons ($k > 0$) from one fluorescent molecule in a sufficiently large reference volume V_0, is given by:

$$p^{(1)}(k;V_o,\varepsilon)=\frac{1}{V_o}\int Poisson\left[k,\varepsilon\cdot W(\vec{r})\right]d\vec{r} \tag{4.51}$$

In the theory for one-photon excitation FCS, $W(r)$ is approximated by a 3D Gaussian function $W_G(r)$:

$$W_G(\vec{r})=\exp\left(-2\frac{x^2+y^2}{{\omega_0}^2}-2\frac{z^2}{{z_0}^2}\right) \tag{4.52}$$

Chen et al. proposed the PCH integration using the 3D Gaussian approximation.[45]

$$p_G^{(1)}(k;V_0,\varepsilon)=\frac{\pi{\omega_0}^2 z_0}{V_0 k!}\int_0^\infty \gamma(k,\varepsilon e^{-2x^2})dx \tag{4.53}$$

However, the PCH model simply based on this approximation is found to be inadequate for fitting experimental data obtained from a confocal setup with one-photon excitation. Perroud et al. proposed a modified model, which is based on the correction to the observation volume profile for the out-of-focus emission by introducing additional F_j terms[20]:

$$p^{(1)}(k;\mathrm{q},\varepsilon)=\frac{1+\mathrm{F}_2}{(1+\mathrm{F}_1)^2}\left[p_G^{(1)}(k;\mathrm{q},\varepsilon)+\frac{1}{\mathrm{q}k!}\sum_{j=k}^{\infty}\frac{(-1)^{j-k}}{(j-k)!(2j)^{3/2}}\varepsilon^j F_j\right] \tag{4.15}$$

Fluorescent intensity is given by,

$$<I>=C\varepsilon\int W(\vec{r})d\vec{r}=C\varepsilon\omega_1 \tag{4.54}$$

And the observation volume is

$$V = \frac{\omega_1^2}{\omega_2}, \text{ where } \omega_j = \int [W(\vec{r})]^j \, d\vec{r} \tag{4.55}$$

The corresponding particle number in V is

$$\bar{N} = \frac{\omega_1^2}{\omega_2} C \tag{4.56}$$

Therefore, the fluorescent intensity is given by

$$< I >= \gamma_2 \bar{N} \varepsilon, \tag{4.57}$$

$$\text{where, } \gamma_2 = \frac{\omega_2}{\omega_1} = \frac{(1+F_2)\omega_2^{3DG}}{(1+F_1)\omega_1^{3DG}} = \frac{(1+F_2)}{(1+F_1)} \gamma_2^{3DG} \text{ and } \gamma_2^{3DG} = \frac{1}{2\sqrt{2}} \tag{4.58}$$

The PCH for the multiple independent species can be obtained by convoluting the PCHs of individual species:

$$P(k; \bar{N}_1, \varepsilon_1, ..., \bar{N}_n, \varepsilon_n) = P(k; \bar{N}_1, \varepsilon_1) \otimes P(k; \bar{N}_2, \varepsilon_2) \otimes \cdots \otimes P(k; \bar{N}_n, \varepsilon_n) \tag{4.59}$$

In $P(k; N_1, \varepsilon_1, ..., N_n, \varepsilon_n) = P(k; N_1, \varepsilon_1) \otimes P(k; N_2, \varepsilon_2) \otimes \cdots \otimes P(k; N_n, \varepsilon_n)$

$$< I >= \gamma_2 \sum_{i=1}^{n} \bar{N}_i \varepsilon_i \tag{4.60}$$

Originally, we used a simple assumption, which is a two-state model: The conformational states of our DNA hairpins can be distinguished according to the status of their molecular brightness, "bright" and "dark". In our experimental condition, we could resolve a binary mixture of molecules that differ in their molecular brightness. (See Section 4.6.3). If the two species model could resolve the different conformational states, we can directly calculate the equilibrium distribution of the different conformational states, K and their relative fluorescence intensities, Q. However, we could not directly resolve the two different states of DNA hairpins using two-species PCH model.[46](See Section 4.6.4) Our FCS analysis shows the clear evidence for the fluorescence intensity

fluctuation caused by conformational changes. Therefore, we conclude that our PCH analysis determine an apparent brightness, ϵ_{app} and apparent number of molecules, $\bar{N}_{app}$, which are defined by[38, 39]

$$\bar{N}_{app} = \frac{\left(\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2\right)^2}{\bar{N}_1\varepsilon_1^{\,2} + \bar{N}_2\varepsilon_2^{\,2}} \tag{4.18}$$

$$\varepsilon_{app} = \frac{\bar{N}_1\varepsilon_1^{\,2} + \bar{N}_2\varepsilon_2^{\,2}}{\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2} \tag{4.19}$$

By definition, $\bar{N}_{app}$ is equal to the inverse value of auto-correlation amplitude at time 0.

$$G_A(0) - 1 = \frac{\bar{N}_1\varepsilon_1^{\,2} + \bar{N}_2\varepsilon_2^{\,2}}{\left(\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2\right)^2} \tag{4.61}$$

In our experiments, these two values are well matched within ~7% error. The small difference may be caused by geometrical factors, such as deviation of observation profile, W(r) from the 3D-Gaussian function. There are also some ways of correction for it, which take non-ideal detector properties (i.e. afterpulsing, dead time) or other photophysical properties (i.e. triplet blinking) with effect of detection time bin into account.[47-49] We found the small difference between $\bar{N}_{app}$ and $[G_A(0)\text{-}1]^{-1}$ have negligible effect in our distribution analysis. (See Table 4.3 in Section 4.6.4)

The fluorescent intensity in the case of two component system is given by

$$< I >= \gamma_2 \bar{N}_{app}\varepsilon_{app} = \gamma_2(\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2). \tag{4.62}$$

From this equation, we have following relationship;

$$\bar{N}_{app}\varepsilon_{app} = \bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2. \tag{4.63}$$

In the FCS analysis, one can only determine an amplitude factor, B_{eq}, which is a function of K_{FFS} and Q. In order to determine actual K_{FFS} from FCS analysis, the simultaneous determination of Q is essential.

We developed a new distribution analysis using combined measurement of FCS and PCH.

First, the amplitude factor B_{eq} is defined by,

$$B_{eq} = K_{FFS}\left[\frac{1-Q}{1+KQ}\right]^2 = \frac{\bar{N}_1\bar{N}_2(\varepsilon_1-\varepsilon_2)^2}{(\bar{N}_1\varepsilon_1+\bar{N}_2\varepsilon_2)^2} \tag{4.13}$$

From the Equations (4.13), (4.18-4.19), and (4.63), the analytical expressions of the parameters ($\bar{N}_1$, $\bar{N}_2$, ϵ_1, ϵ_2, then, K_{FFS} and Q) can be derived as following;

$$\bar{N}_1 = \frac{\bar{N}_{app}\varepsilon_{app}^{\ 2}}{\left(\varepsilon_1^{\ 2} + \frac{(\varepsilon_1-\varepsilon_{app})^2}{B_{eq}}\right)}, \tag{4.20}$$

$$Q = 1-\left[\frac{(\varepsilon_1-\varepsilon_{app})\bar{N}_{app}\varepsilon_{app}}{\varepsilon_1\left(\bar{N}_{app}\varepsilon_{app}-\bar{N}_1\varepsilon_1\right)}\right], \tag{4.64}$$

$$K_{FFS}Q = \frac{\left(\bar{N}_{app}\varepsilon_{app}-\bar{N}_1\varepsilon_1\right)}{\bar{N}_1\varepsilon_1}, \tag{4.65}$$

$$K_{FFS} = \frac{K_{FFS}Q}{Q} \tag{4.66}$$

In order to get a numeric value of the parameters, we need at least one known parameter. In this analysis, we use the molecular brightness of R6G-DNA as ϵ_1. We found the R6g-DNA has same molecular brightness in the solutions of different NaCl concentration. Alternatively, once $\bar{N}_1$ is determined in Equation (4.20), the equilibrium constant K_{FFS} can also be determined using $\bar{N}_{CC} = \bar{N}_1 + \bar{N}_2$,

$$K_{FFS} = \frac{\bar{N}_{CC} - \bar{N}_1}{\bar{N}_1} \tag{4.21}$$

The evaluation of both procedures gives consistent results.

4.6.3 PCH analysis of model two-state system

PCH analysis was performed using our model two-state system, which consists of R6G-DNA and Cy3-DNA. First, PCH analysis of pure single component solutions measures the number of molecules $\bar{N}_{R6g}$, $\bar{N}_{Cy3}$, and molecular brightness, ε_{R6g}, ε_{Cy3}. Then, a mixture containing both components with equal volume was prepared and analyzed under same experimental condition. With two-species fitting, the numbers of molecule and corresponding molecular brightness show good agreement to the expected values from preparation of sample mixture; molecular brightness unchanged and the numbers of molecule halved upon mixing. Much better χ^2 value from two-species fitting indicates the presence of two species in the sample. With one-species fitting, the unreasonable F value or much bigger χ^2 when F is fixed shows that one-species model fails.

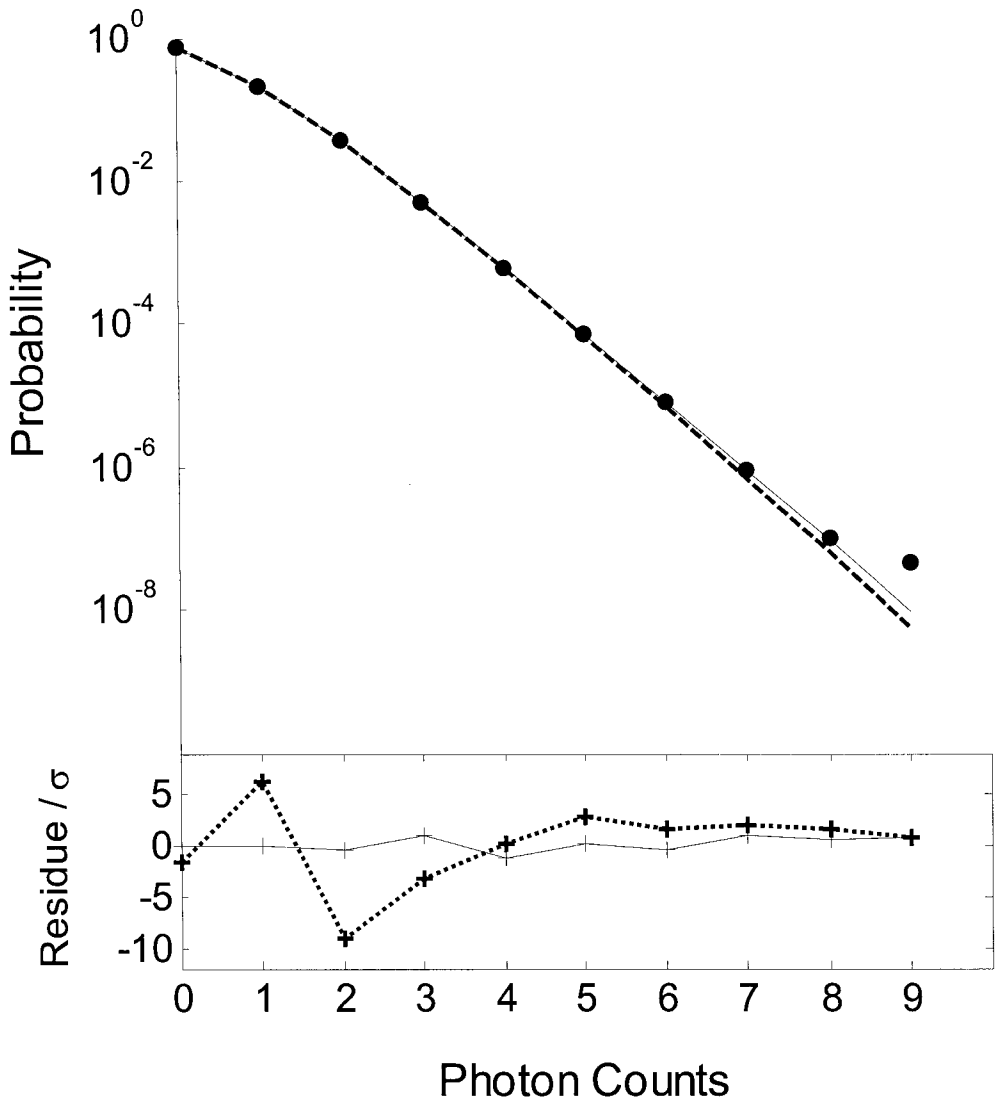

Figure 4.8 PCH analysis of model two-state system
Experimental PCH data and fitting curves with one-species model (dotted line) and two-species model (solid line). The corresponding normalized residues are shown below.

Table 4.2 Parameters from PCH analysis of model two-state system

	N	ϵ	F	$\chi 2$
r6g-poly(T)	2.19 (03)	0.590 (15)	0.732 (36)	0.663
cy3-poly(T)	9.78 (33)	0.148 (04)	0.519 (29)	1.15
mixture	1.10 (36) 5.68 (66)	0.564 (77) 0.149 (07)	0.653 (33)	1.19
*(one species model)	4.74 (06)	0.392 (28)	1.15 (05)	2.01
	4.67 (08)	0.306 (05)	0.653 *(fixed)	18.4

4.6.4 PCH analysis of DNA hairpin sample

PCH analysis was performed using our DNA hairpin samples in various concentration of NaCl solution. Table 4.3 shows a representing data for DNA hairpin sample in 100mM NaCl buffer solution. Both one-species model and two-species model give similar χ^2 from fitting, therefore, we could not directly resolve different species using two-species fitting.[46] We also test the effect of detector non-ideality (i.e. dead time, under our experimental condition, the afterpulse effect is negligible.) and faster processes (i.e. triplet-state formation). From triplet blinking correction, a little bit better match between $\overline{N}_{app}$ and the inverse of $[G_A(0)\text{-}1]$ was observed. However, negligible effects are observed, overall.

Table 4.3 Parameters from PCH analysis of DNA hairpin sample

	N	ϵ	F	χ^2
one species	1.54 (03)	0.539 (31)	0.732 (54)	1.75
(dead time correction)*	1.50 (02)	0.559 (06)	0.793 (17)	2.23
(triplet blinking correction)*	1.66 (31)	0.514 (23)	---	---
two species	0.72 (16) 0.98 (25)	0.61 (14) 0.26 (12)	0.71 (16)	1.69
$[G(0)-1]^{-1}$	1.65 (01)			

4.6.5 Expected behavior of parameters

Based on the information in previous section, we calculate the behavior of parameters in Figure 4.5, which would be expected in a two-state system[40].

In Figure 4.5A, We represent the fraction of folded DNA hairpins (f) on the x-coordinates, which corresponds to K_{melt} for each NaCl concentration as following:

$$f = \frac{K_{melt}}{(1+K_{melt})} = \frac{\bar{N}_2}{\bar{N}_1 + \bar{N}_2} \tag{4.67}$$

Then, the fraction, f and the measured values of ϵ_1 (= 0.59) and ϵ_2 (~0.1) were applied to the Equations (4.68-4.71) to predict the behavior of corresponding parameters, such as:

$$G_A(0) - 1 = \frac{\bar{N}_1\varepsilon_1^2 + \bar{N}_2\varepsilon_2^2}{\left(\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2\right)^2} = \bar{N}_{CC}\frac{(1-f)\varepsilon_1^2 + f\varepsilon_2^2}{\left((1-f)\varepsilon_1 + f\varepsilon_2\right)^2} \tag{4.68}$$

$$\bar{N}_{app} = \frac{1}{\bar{N}_{CC}}\frac{\left((1-f)\varepsilon_1 + f\varepsilon_2\right)^2}{(1-f)\varepsilon_1^2 + f\varepsilon_2^2} \tag{4.69}$$

$$\varepsilon_{app} = \frac{(1-f)\varepsilon_1^2 + f\varepsilon_2^2}{(1-f)\varepsilon_1 + f\varepsilon_2} \tag{4.70}$$

$$< I >= \gamma_2\bar{N}_{CC}\left((1-f)\varepsilon_1 + f\varepsilon_2\right) \tag{4.71}$$

and, $\bar{N}_1 + \bar{N}_2 = \bar{N}_{CC}$

Because we used constant DNA concentration, $\bar{N}_{CC}$ is fixed to a constant value equal to the sum of $\bar{N}_1$ and $\bar{N}_2$ which is calculated from FFS analysis of 0M NaCl case.

For model two-state system, (Figure 4.5B), the x-coordinates represent fraction of Cy3-DNA which is given according to the procedure of sample preparation. The predicted behavior of parameters was calculated in similar way. However, it should be noted that

the model two-state system does not have a chemical reaction given by Equation (4.4), thus, the correlation functions in Equations (4.29) and (4.30) are changed as following:

$$G_C(\tau)-1=\frac{1}{\bar{N}_{app}}g_D(\tau)g_{Fc}(\tau), \text{ and} \tag{4.72}$$

$$G_A(\tau)-1=\frac{1}{\bar{N}_{app}}g_D(\tau)g_T(\tau)g_{Fa}(\tau) \tag{4.73}$$

Therefore, the observed $\bar{N}_{CC}$ behaves same way as $\bar{N}_{app}$. The experimental data of $\bar{N}_{app}$, ε_{app}, and $\bar{N}_1+\bar{N}_2$ were determined using fitting parameters from two-species PCH analysis.

4.6.6 FFS experiments with varying sample concentration

DNA hairpin samples were in the various concentration of NaCl solution(0, 5, 10, 25, 50, 100, 200, 500mM). Each sample concentration was adjusted to achieve constant fluorescence signal counts (~39kHz). Samples for the model two-state system were prepared by mixing two stock solutions containing only one component (R6g-DNA or Cy3-DNA), each. The concentrations of both stock solutions were adjusted to achieve same signal counts rate (~40kHz), thus, the resulting mixtures showed same strength of signal. Figure 4.7 shows the experimental results. The predicted behavior of parameters was calculated in similar way as one described in previous section (Equations (4.68)-(4.71)). One difference is that because each sample has different concentration, the value of $\bar{N}_{CC}$ is not fixed to a constant but obtained from measured parameter from cross correlation curve fitting. The model system exhibited precisely the behavior that would be expect for a two-state system, while the DNA hairpin samples exhibited behavior that

was anomalous for a two-state system. From the observation, we could exclude the possibility of detector saturation artifacts in our experiments.

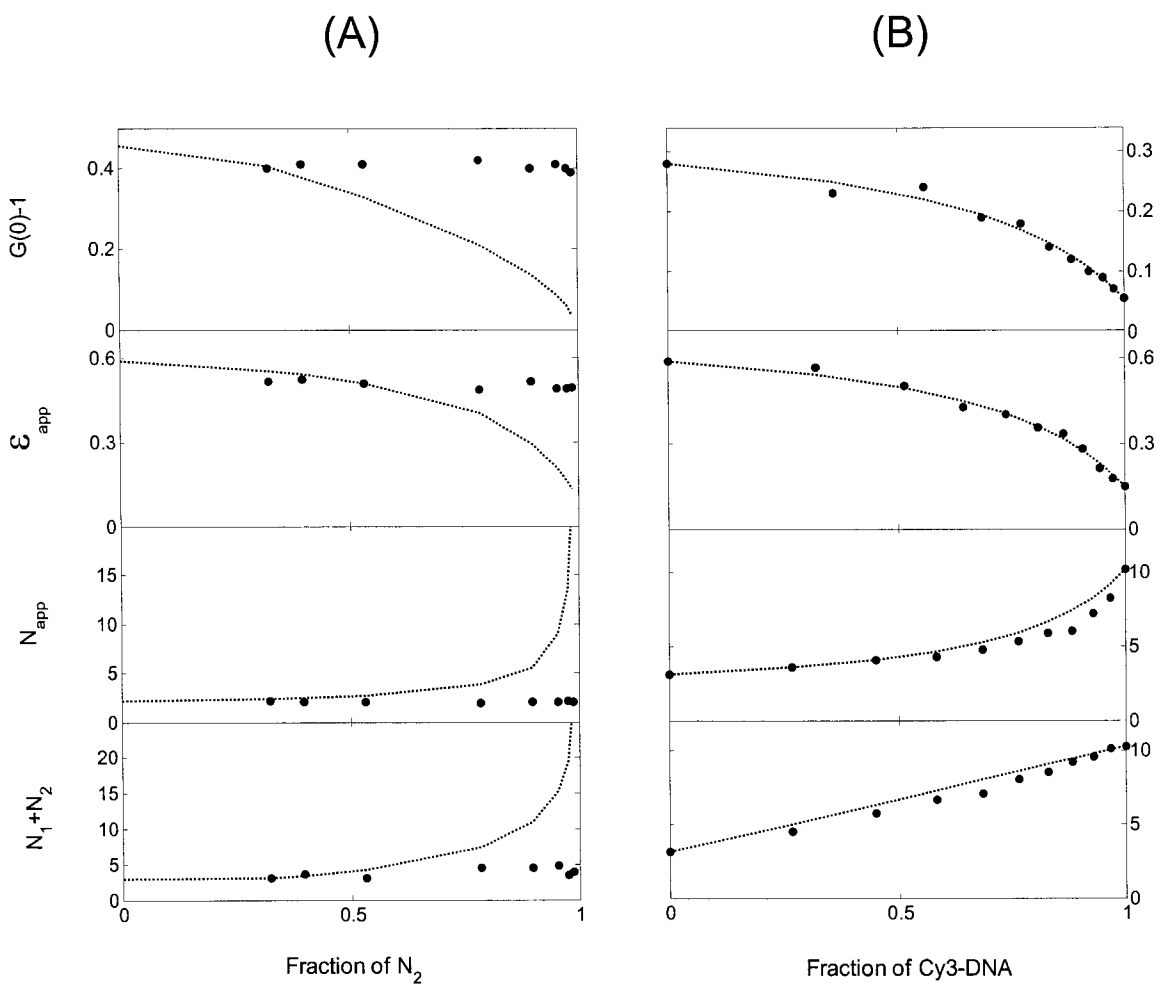

Figure 4.9 FFS analysis of experiments with varying sample concentration
FFS data of DNA hairpin sample (**A**) and model two-state system (**B**) are shown with curves representing expected behavior of each parameter (dotted lines).
Signal intensity $< I >$ was equal for all the samples in each experiment.

4.6.7 Calculation of $K_{melt,3S}$

In order to explain the observed melting curves of DNA samples, we define a parameter $K_{melt,3S}$ as

$$K_{melt,3S,i} = \frac{\langle I \rangle_{\max} - \langle I \rangle_i}{\langle I \rangle_i - \langle I \rangle_{\min}} \quad , \tag{4.25}$$

where, $\langle I \rangle_i$ is the average fluorescence count rate observed in our FCS and PCH experiment for a given NaCl concentration. $\langle I \rangle_{\max}$ and $\langle I \rangle_{\min}$ are the fluorescence count rates that would be observed from samples in which all the DNA was present in a completely open or a completely closed form, respectively. Therefore, $\langle I \rangle_{\max}$ is given by $\gamma_2 \bar{N}_{total} \varepsilon_1$, where $\gamma_2 = \frac{1}{2\sqrt{2}} \frac{(1+F_2)}{(1+F_1)}$. In same way, $\langle I \rangle_{\min}$ can be given by $\gamma_2 \bar{N}_{total} \varepsilon_3$. However, the molecular brightness of a completely closed form, ε_3 is not determined, thus, $\langle I \rangle_{\min}$ is taken to be approximately equal to the photocount rate calculated from $\langle I \rangle_i$ of the 500-mM NaCl sample and corresponding melting curve:

$$< I >_{\min} = < I >_{NaCl-500mM} \times \left[\frac{I(5^o C)}{I(19^o C)} \right]_{NaCl-500mM} \tag{4.74}$$

In Figure 4.3A, the calculated $K_{melt,3S}$ are compared with K_{melt} values obtained directly from the melting curves and show remarkably good agreement.

4.6.8 Checking salt effects on dye property

In order to verify that our observations are not artifacts caused from interaction of dye in higher ion concentration, we performed another control experiment by monitoring fluorescence intensities of various DNA samples as function of NaCl concentration in the buffer solutions. (Figure 4.10) We found only DNA hairpin dual labeled with dye and quencher at the ends shows dramatic changes in fluorescence intensity. DNA hairpin and poly T molecules with only fluorescence dye at the end, do not show any significant changes. It supports the fact that fluorescence quenching results from dye-quencher interaction, not from salt effects on the dye properties, especially in the case of higher NaCl concentration. Although fluorescence quenching by guanine moiety in the DNA molecules are well known, the present control experiments show that the effect is not significant in our sample system, thus, the possibility can be excluded safely.

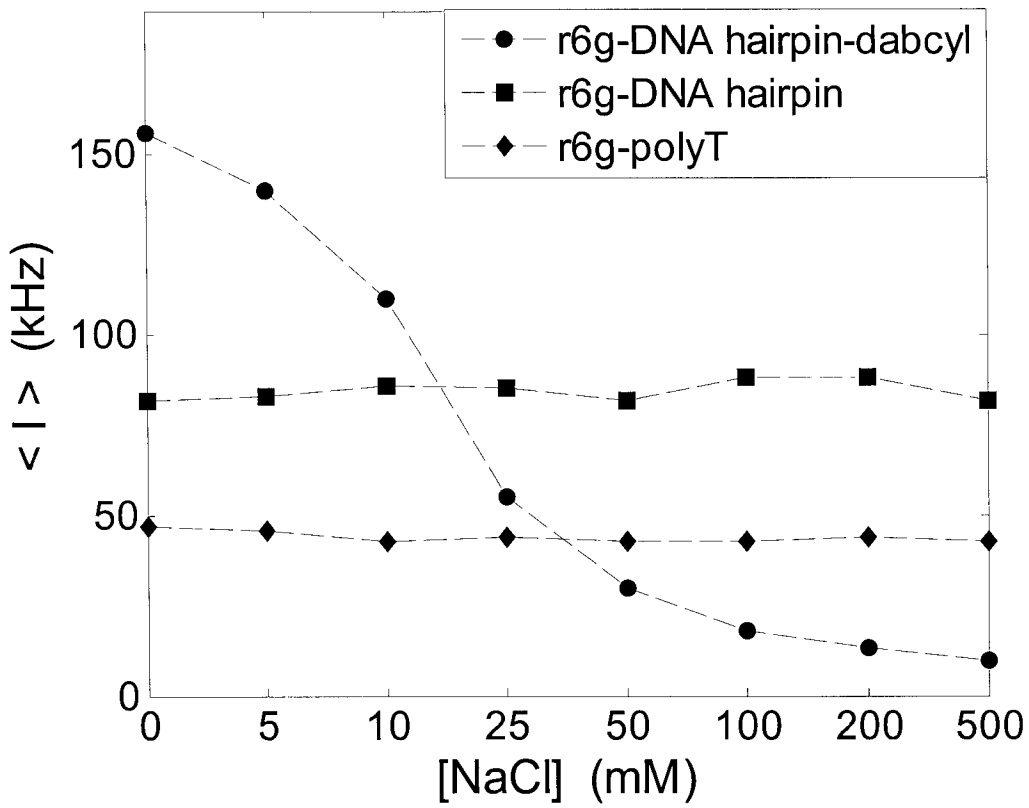

Figure 4.10 Fluorescence Intensity measurements as function of salt concentration in the buffer
Only dye-quencher dual labeled DNA hairpin shows dramatic changes in fluorescence intensity, while single dye labeled molecule (DNA hairpin and poly T) do not show the change.

4.6.9 Programming code for the PCH analysis (Matlab)

```
%PCH fitting for one-photon excitation, modified since 05/30/04
function fitting = pch

%Choose data file
[fname,pname] = uigetfile('*.*','Select data file');

%Choose channel to analyze
pch = input('channel to analyze : ');
% '0' if no pch fitting,
% '1' for first channel fitting,
% '2' for second channel fitting

%Choose analysis mode
epch = 1;
% '0' if only pch fitting,
% '1' for other information (i.e. count rate)

%%%%%%%%%%%%Reading data of the channel%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Establishing data matrix
Gtemp=[]; %PCH data
Gtemp2=[]; %count rate data
Gtemp3=[];
F=[];

%Opening data file for reading
fid=fopen(fname);

%Reading other part of the data file (FCS part)
for j=1:105
   str=fgetl(fid);
end

%Reading the PCH data
cmp=0;
g=1;
while cmp ~= 1
   str=fgetl(fid);
   cmp=strcmp(str,'');
   if cmp ~= 1
      Gt = str2num(str);
      if Gt(pch+1) < 1
         cmp =1;
         cmax = g-1;
      end
      Gtemp=[Gtemp;Gt] ;
      g=g+1;
   end
end

%Reading the rest of PCH data
for j=1:20-cmax
   str=fgetl(fid);
```

```
end

%Reading the Count-rate data
if epch == 1
  for j=1:3
   str=fgetl(fid);
  end
cmp=0;
while cmp ~= 1
   str=fgetl(fid);
   cmp=strcmp(str,'');
   if str == -1
      cmp = 1;
   end
   if cmp ~= 1
      Gtemp2=[Gtemp2;str2num(str)];
   end
end
end
status=fclose(fid);

%%%%%%%%%%%%%%%%%Experimenatal PCH%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Establishing PCH
pc = Gtemp(1:cmax,1); %the x axis of PCH, 'k'
his_1 = Gtemp(1:cmax,pch+1); %the y axis of PCH 'counts'
histot1=sum(his_1); %total number of data points,'M'
his1=his_1/histot1; %normalized 'counts' (probability, 'p')
his111=1-his1; %'q'
his11=sqrt(his1.*his111./histot1); %sigma = sqrt(M*p*q)

I=mean(Gtemp2); %Average fluorescence intensity
nc= cmax - 1; %maximum value of 'k'

%Showing the PCH
close all
if pch == 0
   figure(1)
   semilogy(pc,his1,'k.')
   return
end

%%%%%%%%%%%%%%%%%%%%%Control Panel%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Choose the PCH fitting mode
npch = input('PCH mode : ');
% '1' for one species fitting,
% '2' for two species fitting with FCS constraint,
% '3' for two species (no FCS constraint)

%Input initial guess values for the PCH fitting
e0=input('molecular brightness of species 1 : ');
e20=input('molecular brightness of species 2 : ');
N0=input('number of species 1 : ');
N20=input('number of species 1 : ');
F0=input('F value : ');

%Choose F property
```

```
Ffix = input('F value is fixed ?: ');
% '0' if F is not fixed,
% '1' if F is fixed to F0

bod=2.0; %fitting boundary factor

%%%%%%%%%%%%%%%%%%%%PCH Fitting%%%%%%%%%%%%%%%%%%%%%%%%%%%%
tau1=0:nc;
p0=[e0;F0;N0];
l_b=[e0/bod;F0/bod;N0/bod]; %lower boundary
u_b=[e0*bod;F0*bod;N0*bod]; %upper boundary
d=numel(p0); %degree of freedom
if Ffix == 1
   l_b=[e0/bod;F0*0.9999999999;N0/bod];
   u_b=[e0*bod;F0*1.0000000001;N0*bod];
   d=d-1;
end

figure(1)
subplot(2,1,1)
semilogy(pc,his1,'k.')
hold on

%%%%%%%%%%%%%%%%One species PCH Fitting%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if npch == 1
   [p,resnorm,residual,exitflag,output,lambda,jacobian] = lsqcurvefit(@fida,p0,pc,his1,l_b,u_b,[],nc);
   Fm=fida(p,tau1,nc);
   Ec=p(1);
   Nc=p(3);
   Fc=p(2);
%%%%%%%%%%%%%%%% Calculate geometric factor gamma %%%%%%%%%%%%%%%%%%%%%%%%
    if epch == 1
      Gtemp2=Gtemp2(:,pch);
      I=mean(Gtemp2); %average intensity, '<k>'
      I=I(1);
      gamma1=I(pch)*(1+Fc)/(Nc*Ec*1000/9);

%%%%%%%%%%%%%%%% Correction for triplet state formation%%%%%%%%%%%%%%%%
%Choose values from FCS analysis
      tau = input('timebin of PCH in us : ');
      taut = input('triplet relaxation time in us : ');
      T = input('Teq in FCS analysis : ');
      gammatri=1+2*T*taut/(tau*(1-T))*(1-taut*(1-exp(-tau/taut))/tau);
      N_0=Nc*gammatri;
      E_0=Ec/gammatri;

%%%%%%%%%%%%%%%%Calculation of G(0) based on moment analysis%%%%%%%%%%%%%%%%
      Gtempp=Gtemp2.*Gtemp2;  %k2
      I2=mean(Gtempp);  %<k2>
      Gtemp3=Gtempp-I2;
      I3=mean(Gtemp3); %<delta k2>
      gzero=I3-I/(I)
    end

%%%%%%%%%%Tw0-sepcies fitting with FCS Constraint%%%%%%%%%%%%%%%%%%%%%%%%
else if npch == 2
```

```
      B = input('Beq in FCS analysis : ');
      Nt = input('Ncc in FCS analysis : ');

      p0=[e0;F0;N0];
      l_b=[e0/bod;F0/bod;N0/bod];
      u_b=[e0*bod;F0*bod;N0*bod];
      d=numel(p0);
      if Ffix == 1
         l_b=[e0/bod;F0*0.9999999999;N0/bod];
         u_b=[e0*bod;F0*1.0000000001;N0*bod];
         d=d-1;
      end

      [p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@fida2,p0,pc,his1,l_b,u_b,[],B,Nt,nc);
      Fm=fida2(p,tau1,B,Nt,nc);
      ci = nlparci(p,residual,jacobian); %confidential interval
      E1=p(1);
      N1=p(3);
      F=p(2);
      N2=Nt-N1;
      K=N2/N1;
      E2=E1*(1-sqrt(B/K))/(1+sqrt(B*K));
      Q=E2/E1;
      BB=K*((1-Q)/(1+K*Q))^2;

%%%%%%%%%%%%%%%%%%%%%%Two-species fitting without FCS constraint%%%%%%%%%%%%%%%%
   else if npch == 3
    p0=[e0;e20;F0;N0;N20];
      l_b=[e0/bod;e20/bod;F0/bod2;N0/bod;N20/bod];
      u_b=[e0*bod;e20*bod;F0*bod2;N0*bod;N20*bod];
      d=numel(p0);
      if Ffix == 1
         l_b=[e0/bod;e20/bod;F0*0.9999999999;N0/bod;N20/bod];
         u_b=[e0*bod;e20*bod;F0*1.0000000001;N0*bod;N20*bod];
         d=d-1;
      end
      [p,resnorm,residual,exitflag,output,lambda,jacobian] =
lsqcurvefit(@fida3,p0,pc,his1,l_b,u_b,[],nc,Nt);
      Fm=fida3(p,tau1,nc,Nt);
      E1=p(1)
      E2=p(2)
      N1=p(4)
      N2=p(5)
      F=p(3)
      K=N2/N1
      Q=E2/E1
      BB=K*((1-Q)/(1+K*Q))^2
      end
   end
end
ci = nlparci(p,residual,jacobian); %confidential interval
resn1=residual./his11;
resn2=resn1.*resn1;
resn=sum(resn2);
resn=resn/(nc-d);
```

```
figure(1)
subplot(2,1,1)
semilogy(tau1,Fm,'r')
hold on
subplot(2,1,2)
plot(tau1,-resn1,'r:+')
hold on
end

%%%%%%%%%%%%%%Subroutines%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%% for One species PCH %%%%%%%%%%%%%%%%%%%%%%%%%
function FIDA = fida(p,taul,nc)
Q=6; % an arbitrary number
Nm=floor(4+Q*p(3)+5*sqrt(Q*p(3)));
Pr=[]; Prom=[]; PoiN=[]; Po=[]; Prm=[];

%Calculation of single particle function
fname='epsilon'; %precalculated probability function
eps0=0.01; %minimum E
epsint=log10(1.01);
Gtemp=[];

%Opening file for reading
fid=fopen(fname,'r');
for j=1:nc
   str=fgetl(fid);
   Gt = str2num(str);
   Gtemp=[Gtemp;Gt] ;
   P0=0;
   Po=[Po;P0];
end
index=floor((p(1)-eps0)/epsint);

% Linear interpolation
ggl=Gtemp(:,index);
ggu=Gtemp(:,index + 1);
lower=eps0+index*log10(1.01);
upper=eps0+(index+1)*log10(1.01);
ratio1=(upper-p(1))/(upper-lower);
ratio2=(p(1)-lower)/(upper-lower);
gg=ggl*ratio1+ggu*ratio2;
status=fclose(fid);

% correction of out of focus beam
Pro1k=p(1)*p(2)/(Q*2*sqrt(2));
gg(1)=gg(1)+Pro1k;
Pr=gg./((1+p(2))^2);
Po=[1; Po]; % no-particle probability function
Pr0=1-sum(Pr);
Pr=[Pr0;Pr] ; % one particle probability function

%Poisson function for N
N=0:Nm+1;
    Poin = poisspdf(N,Q*p(3));
    PoiN= Poin';
```

```
%Convolution for multi particle probability
N=1;
Pr1=Pr;
for N=1:Nm
   Prc=conv(Pr1,Pr);
   Pr_2=Prc(1:nc+1);
   Prm=[Prm Pr_2];
   Pr1=Prc;
end
Prm=[Po Pr Prm];
Prob1=Prm*PoiN;
FIDA = Prob1;
end

%%%%%%%%%%for Two species PCH with FCS constraint %%%%%%%%%%
function FIDA = fida2(p,taul,B,Nt,nc)
N2=Nt-p(3);
K=N2/p(3);
E2=p(1)*(1-sqrt(B/K))/(1+sqrt(B*K));
Q=6;
Nm=floor(4+Q*Nt+5*sqrt(Q*Nt));
Pr1=[]; Pr_1=[]; Prom=[]; PoiN1=[]; PoiN2=[]; Po=[]; Pch2=[];

fname='epsilon';
eps0=0.01;
epsint=log10(1.01);
Gtemp=[];
fid=fopen(fname,'r');
for j=1:nc
   str=fgetl(fid);
   Gt = str2num(str);
   Gtemp=[Gtemp;Gt] ;
   P0=0;
   Po=[Po;P0];
end
index=floor((p(1)-eps0)/epsint);
index2=floor((E2-eps0)/epsint);
ggl=Gtemp(:,index);
ggu=Gtemp(:,index + 1);
lower=eps0+index*log10(1.01);
upper=eps0+(index+1)*log10(1.01);
ratio1=(upper-p(1))/(upper-lower);
ratio2=(p(1)-lower)/(upper-lower);
gg=ggl*ratio1+ggu*ratio2;
ggl2=Gtemp(:,index2);
ggu2=Gtemp(:,index2 + 1);
lower2=eps0+index2*log10(1.01);
upper2=eps0+(index2+1)*log10(1.01);
ratio12=(upper2-E2)/(upper2-lower2);
ratio22=(E2-lower2)/(upper2-lower2);
gg2=ggl2*ratio12+ggu2*ratio22;
status=fclose(fid);
Pro1k=p(1)*p(2)/(Q*2*sqrt(2));
gg(1)=gg(1)+Pro1k;
Pr1=gg./((1+p(2))^2);
```

```
Po=[1; Po];
Pr0=1-sum(Pr1);
Pr1=[Pr0;Pr1] ;
Pro1k2=E2*p(2)/(Q*2*sqrt(2));
gg2(1)=gg2(1)+Pro1k2;
Pr_1=gg2./((1+p(2))^2);
Pr02=1-sum(Pr_1);
Pr_1=[Pr02;Pr_1] ;

   N=0:Nm+1;
   Poin1 = poisspdf(N,Q*p(3));
   Poin2 = poisspdf(N,Q*N2);
   PoiN1= Poin1';
   PoiN2= Poin2';

N=1;
Pr1t=Pr1; Prm=[];Pr2t=Pr_1; Prm2=[];
for N=1:Nm
   Prc=conv(Pr1t,Pr1);
   Prtemp=Prc(1:nc+1);
   Prm=[Prm Prtemp];
   Pr1t=Prc;
   Prc2=conv(Pr2t,Pr_1);
   Pr2temp=Prc2(1:nc+1);
   Prm2=[Prm2 Pr2temp];
   Pr2t=Prc2;
end
Prm=[Po Pr1 Prm];
Prob1=Prm*PoiN1;
Prm2=[Po Pr_1 Prm2];
Prob2=Prm2*PoiN2;

%Two species PCH
Pch=conv(Prob1,Prob2);
Pch2=Pch(1:nc+1);
FIDA = Pch2;
end

%%%%%%%%%%%%Two-species PCH fitting without FCS constraint%%%%%%%%%%%%
function FIDA = fida3(p,taul,nc,Nt)
Q=6;
Nm=floor(4+Q*Nt+5*sqrt(Q*Nt));
Pr1=[]; Pr_1=[]; Prom=[]; PoiN1=[]; PoiN2=[]; Po=[]; Pch2=[];

fname='epsilon';
eps0=0.01;
epsint=log10(1.01);
Gtemp=[];
fid=fopen(fname,'r');
for j=1:nc
   str=fgetl(fid);
   Gt = str2num(str);
   Gtemp=[Gtemp;Gt] ;
   P0=0;
   Po=[Po;P0];
end
```

```
index=floor((p(1)-eps0)/epsint);
index2=floor((p(2)-eps0)/epsint);
ggl=Gtemp(:,index);
ggu=Gtemp(:,index + 1);
lower=eps0+index*log10(1.01);
upper=eps0+(index+1)*log10(1.01);
ratio1=(upper-p(1))/(upper-lower);
ratio2=(p(1)-lower)/(upper-lower);
gg=ggl*ratio1+ggu*ratio2;
ggl2=Gtemp(:,index2);
ggu2=Gtemp(:,index2 + 1);
lower2=eps0+index2*log10(1.01);
upper2=eps0+(index2+1)*log10(1.01);
ratio12=(upper2-p(2))/(upper2-lower2);
ratio22=(p(2)-lower2)/(upper2-lower2);
gg2=ggl2*ratio12+ggu2*ratio22;
status=fclose(fid);
Pro1k=p(1)*p(3)/(Q*2*sqrt(2));
gg(1)=gg(1)+Pro1k;
Pr1=gg./((1+p(3))^2);
Po=[1; Po];
Pr0=1-sum(Pr1);
Pr1=[Pr0;Pr1] ;
Pro1k2=p(2)*p(3)/(Q*2*sqrt(2));
gg2(1)=gg2(1)+Pro1k2;
Pr_1=gg2./((1+p(3))^2);
Pr02=1-sum(Pr_1);
Pr_1=[Pr02;Pr_1] ;

   N=0:Nm+1;
   Poin1 = poisspdf(N,Q*p(4));
   Poin2 = poisspdf(N,Q*p(5));
   PoiN1= Poin1';
   PoiN2= Poin2';

N=1;
Pr1t=Pr1; Prm=[];Pr2t=Pr_1; Prm2=[];
for N=1:Nm
   Prc=conv(Pr1t,Pr1);
   Prtemp=Prc(1:nc+1);
   Prm=[Prm Prtemp];
   Pr1t=Prc;
   Prc2=conv(Pr2t,Pr_1);
   Pr2temp=Prc2(1:nc+1);
   Prm2=[Prm2 Pr2temp];
   Pr2t=Prc2;
end
Prm=[Po Pr1 Prm];
Prob1=Prm*PoiN1;
Prm2=[Po Pr_1 Prm2];
Prob2=Prm2*PoiN2;

Pch=conv(Prob1,Prob2);
Pch2=Pch(1:nc+1);
FIDA = Pch2;
end
```

4.7 REFERENCE

1. Cantor, C.R. and P.R. Schimmel, *Biophysical Chemistry: the Behavior of Biological Macromolecules*. 1980, Freeman: New York.
2. Tyagi, S. and F.R. Kramer, *Molecular Beacons: Probes that Fluoresce upon Hybridization.* Nature Biotechnology, 1996. **14**: p. 303-308.
3. Bonnet, G., et al., *Thermodynamic Basis of the Enhanced Specificity of Structured DNA Probes.* Proc. Natl. Acad. Sci. USA, 1999. **96**(11): p. 6171-6176.
4. Liu, X. and W. Tan, *a Fiber Optic Evanescent Wave DNA Biosensor Based on Novel Molecular Beacons.* Anal. Chem., 1999. **71**: p. 5054-5059.
5. Wang, X. and W.M. Nau, *Kinetics of End-to-End Collision in Short Single-Stranded Nucleic Acids.* J. Am.Chem. Soc., 2004. **126**: p. 808-813.
6. Ansari, A., S.V. kuznetsov, and Y. Shen, *Configurational Diffusion Down a Folding Funnel Describes the Dynamics of DNA Hairpins.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 7771-7776.
7. Bonnet, G., O. Krichevsky, and A. Libchaber, *Kinetics of Conformational Fluctuation in DNA Hairpin Loops.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 8602-8606.
8. Goddard, N.L., et al., *Sequence Dependent Rigidity of Single Stranded DNA.* Phys. Rev. Lett., 2000. **85**: p. 2400-2403.
9. Kuznetsov, S.V., et al., *A Semiflexible Polymer Model Applied to Loop Formation in DNA Hairpins.* Biophys. J., 2001. **81**: p. 2864-2875.
10. Senior, M.M., R.A. Jones, and K.J. Breslauer, *Influence of Loop Residue on the Relative Stabilities of DNA Hairpin Structures.* Proc. Natl. Acad. Sci. USA, 1988. **85**: p. 6242-6246.
11. Shen, Y., S.V. kuznetsov, and A. Ansari, *Loop Dependence of the Dynamics of DNA Hairpins.* J. Phys. Chem. B, 2001. **105**: p. 12202-12211.
12. Wallace, M.I., et al., *FRET Fluctuation Spectroscopy: Exploring the Conformational Dynamics of a DNA Hairpin Loop.* J. Phys. Chem. B, 2000. **104**: p. 11551-11555.
13. Wallace, M.I., et al., *Non-Arrhenius Kinetics for the Loop Closure of a DNA Hairpin.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 5584-5589.
14. Li, H., et al., *Measuring Single-Molecule Nucleic Acid Dynamics in Solution by Two-Color Filtered Ratiometric Fluorescence Correlation Spectroscopy.* Proc. Natl. Acad. Sci. USA, 2004. **101**: p. 14425-14430.
15. Jung, J. and A. Van Orden, *Folding and Unfolding Kinetics of DNA Hairpins in Flowing Solution by Multiparameter Fluorescence Correlation Spectroscopy.* J. Phys. Chem. B, 2005. **109**: p. 3648-3657.

16. Ansari, A. and S.V. Kuznetsov, *Is Hairpin Formation in Single-Stranded Polynucleotide Diffusion-Controlled?* J. Phys. Chem. B, 2005. **109**: p. 12982-12989.

17. Ying, L., M.I. Wallace, and D. Klenerman, *Two-State Model of Conformational Fluctuation in a DNA Hairpin-Loop.* Chem. Phys. Lett., 2001. **334**: p. 145-150.

18. Chen, Y., et al., *The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy.* Biophys. J., 1999. **77**: p. 553-567.

19. Kask, P., et al., *Fluorescence Intensity Distribution Analysis and Its Application in Biomolecular Detection Technology.* Proc. Natl. Acad. Sci. USA, 1999. **96**(24): p. 13756-13761.

20. Perroud, T.D., et al., *Photon Counting Histogram for One-Photon Excitation.* Chemphyschem, 2003. **4**: p. 1121-1123.

21. Van Orden, A., K. Fogarty, and J. Jung, *Fluorescence Fluctuation Spectroscopy: A Coming of Age Story.* Appl. Spectrosc., 2004. **58**: p. 122A-137A.

22. Brinkmeier, M., et al., *Two-Beam Cross-Correlation: a Method to Characterize Transport Phenomena in Micrometer-Sized Structures.* Anal. Chem., 1999. **71**: p. 609-616.

23. Gosch, M., et al., *Hydrodynamic Flow Profiling in Microchannel Structures by Single Molecule Fluorescence Correlation Spectroscopy.* Anal. Chem., 2000. **72**: p. 3260-3265.

24. Huang, B., T.D. Perroud, and R.N. Zare, *Photon Counting Histogram: One-Photon Excitation.* Chemphyschem, 2004. **5**: p. 1523-1531.

25. Magde, D., W.W. Webb, and E.L. Elson, *Fluorescence Correlation Spectroscopy 3. Uniform translation and laminar flow.* Biopolymers, 1978. **17**: p. 361-376.

26. Dittrich, P.S. and P. Schwille, *Spatial Two-Photon Fluorescence Cross-Correlation Spectroscopy for Controlling Molecular Transport in Microfluidic Structures.* Anal. Chem., 2002. **74**: p. 4472-4479.

27. Kunst, B.H., A. Schots, and A.J.W.G. Visser, *Detection of Flowing Fluorescent Particles in a Microcapillary Using Fluorescence Correlation Spectroscopy.* Anal. Chem., 2002. **74**: p. 5350-5357.

28. LeCaptain, D.J. and A. Van Orden, *Two-Beam Fluorescence Cross-Correlation Spectroscopy in an Electrophoretic Mobility Shift Assay.* Anal. Chem., 2002. **74**: p. 1171-1176.

29. Fogarty, K. and A. Van Orden, *Two-Beam Fluorescence Cross-Correlation Spectroscopy for Simultaneous Analysis of Positive and Negative Ions in Continuous Flow Capillary Electrophoresis.* Anal. Chem., 2003. **75**: p. 6634-6641.

30. Elson, E.L. and D. Magde, *Fluorescence Correlation Spectroscopy 1. Conceptual Basis and Theory.* Biopolymers, 1974. **13**: p. 1-27.

31. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence Correlation Spectroscopy 2. An Experimental Realization.* Biopolymers, 1974. **13**: p. 29-61.

32. Palmer, A.G. and N.L. Thompson, *Theory of Sample Translation in Fluorescence Correlation Spectroscopy.* Biophys. J., 1987. **51**: p. 339-343.

33. Widerngren, J., U. Mets, and R. Rigler, *Fluorescence Correlation Spectroscopy of Triplet States in Solution: a Theoretical and Experimental Study.* J. Phys. Chem. B, 1995. **99**: p. 13368-13379.

34. Widerngren, J. and R. Rigler, *Fluorescence Correlation Spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces.* Cell. mol. Biol., 1998. **44**: p. 857-879.

35. Van Orden, A. and R.A. Keller, *Fluorescence Correlation Spectroscopy for Rapid Multicomponent Analysis in a Capillary Electrophoresis System.* Anal. Chem., 1998. **70**: p. 4463-4471.

36. Rigler, R. and E.L. Eigen, *Fluorescence Correlation Spectroscopy: Theory and Applications.* 2001, Berlin: Springer-Verlag.

37. Manning, G.S., *The Molecular Theory of Polyelctrolyte Solutions with Applications to the Electrostatic Properties of Polynucleotides.* Q. Rev. Biophys., 1978. **11**: p. 179-246.

38. Chen, Y., et al., *Molecular Brightness Characterization of EGFP In Vivo by Fluorescence Fluctuation Spectroscopy.* Biophys. J., 2002. **82**: p. 133-144.

39. Chen, Y., L. Wei, and J.D. Mueller, *Probing Protein Oligomerization in Living Cells with Fluorescence Fluctuation Spectroscopy.* Proc. Natl. Acad. Sci. USA, 2003. **100**(26): p. 15492-15497.

40. Chen, Y., et al., *Probing Ligand Protein Binding Equilibria with Fluorescence Flutuation Spectroscopy.* Biophys. J., 2000. **79**: p. 1074-1084.

41. Sorin, E.J., et al., *Insights into Nucleic Acid Conformational Dynamics from Massively Parallel Stochastic Simulations.* Biophys. J., 2003. **85**: p. 790-803.

42. Sorin, E.J., Y.M. Rhee, and V.S. Pande, *Does Water Play a Structural Role in the Folding of Small Nucleic Acids?* Biophys. J., 2005. **88**: p. 2516-2524.

43. Grunwell, J.R., et al., *Monitoring the Conformational Fluctuations of DNA Hairpins Using Single-Pair Fluorescence Resonance Energy Transfer.* J. Am.Chem. Soc., 2001. **123**: p. 4295-4303.

44. Liphardt, J., et al., *Reversible Unfolding of Single RNA Molecules b Mechanical Force.* Science, 2001. **292**: p. 733.

45. Chen, Y., et al., *Fluorescence Fluctuation Spectroscopy.* Methods, 1999. **19**: p. 234-252.

46. Mueller, J.D., Y. Chen, and E. Gratton, *Resolving Heterogeneity on the Single Molecular Level with the Photon-Counting Histogram.* Biophys. J., 2000. **78**: p. 474-486.

47. Hillesheim, L.N. and J.D. Muller, *The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy with Non-ideal Photdetectors.* Biophys. J., 2003. **85**: p. 1948-1958.

48. Perroud, T.D., B. Huang, and R.N. Zare, *Effect of Bin Time on the Photon Counting Histogram for One-Photon Excitation.* Chemphyschem, 2005. **6**: p. 905-912.

49. Palo, K., et al., *Fluorescence Intensity Multiple Distributions Analysis: Concurrent Determination of Diffusion Times and Molecular Brightness.* Biophys. J., 2000. **79**: p. 2858-2866.

Chapter 5

Conclusion

In this dissertation, we described our efforts to obtain a better understanding of the biomolecule conformational dynamics with time range of micro-seconds to milli-seconds. Conformational fluctuations of biomolecules (i.e. folding and unfolding) play a critical role in their functioning. Due to the complexities in their biological behavior, a non-invasive and unperturbed approach should have many advantages for the investigation of biomolecule conformational dynamics. We developed a new Fluorescence Fluctuation spectroscopy (FFS)-based technique by focusing two laser beams into sample solution and applied it to the study of DNA hairpin folding and unfolding reaction in the solution. This seemingly simple reaction has been investigated for more than thirty years with the various experimental tools as well as theoretical

models. However, many of basic issues still remain unresolved. As a new angle of approach toward this decade old folding problem, the kinetics of the DNA hairpin formation was investigated using our two-beam FFS technique. Alternative scenario from a more probable description of the reaction mechanism was discussed in this dissertation. The approach is to detect fluorescence from molecules as they flow sequentially between two spatially offset laser probe volumes. The fluorescence from each probe volume is imaged onto separate detectors, and the detected photocounts are analyzed simultaneously from three different vantage points: cross-correlation analysis of the two detection channels relative to each other, independent auto-correlation analysis of the two detection channels, and photon counting histogram (PCH) analysis. In this study, the relaxation time for opening and closing of a stem-loop DNA hairpin molecule is precisely determined under thermodynamic equilibrium conditions. The DNA hairpins are labeled with Rhodamine 6G (R6G) and 4-{[4-(dimethylamino) phenyl] axo} benzoic acid (dabcyl) at the 5' and 3' ends, respectively. When the hairpin adopts a closed conformation, the R6G fluorescence is quenched due to its proximity with dabcyl. The R6G fluorescence is recovered when the DNA adopts an open conformation. Equilibrium conformational transformations between open and closed states create fluctuations in the R6G fluorescence that occur on the same time scale as the transitions. The sample hairpin sequences were 5'-AACCC-$(T)_n$-GGGTT-3' where *n* is the number of deoxythymine

residues in the loops. For example, hairpin named as hp-T_{21} contains twenty one thymines and hp-T_{30} has thirty thymines on the loops, respectively.

Auto-correlation analysis in Fluorescence Correlation Spectroscopy (FCS), the most prominent example of FFS, is proven to be an efficient technique for characterizing many types of molecular processes non-invasively, under conditions of thermodynamic equilibrium. It analyzes the spontaneous fluctuations in the fluorescence signal arising from a microscopic subvolume of a fluorescent sample. The measurements are carried out *in situ* and in real time. Macroscopic perturbation of the system equilibrium to establish a start time for the reaction is not required. We observed that difficulties with the analysis arise, however, when multiple processes are occurring simultaneously in the same system, especially when the relaxation times of the different processes occur on similar time scales. In the case when studying the conformational fluctuations of freely diffusing biomolecules like studies presented here, the resulting auto-correlation function thus contains fluctuation contributions characterized by both a diffusional relaxation time, τ_D, and a relaxation time for conformational fluctuations, τ_R. The usual way of decoupling the different relaxation times is to design a control molecule that gives an independent measure of τ_D, which then constrains the correlation function for the molecule of interest, thereby allowing the determination of τ_R. There are several disadvantages to this approach, including the ever-present possibility of false assumptions about the nature of

the control molecule. Furthermore, if the ultimate goal is to perform these types of measurements in the cellular environment, it would be much more desirable to extract all the relevant information about the dynamic processes of the molecule of interest from a single set of fluorescence intensity measurements obtained from the sample of interest alone. With this in mind, the presented approach was developed based on multi-parameter analysis technique. By nature, the analysis utilizes consistent data from single set of measurement with identical sample, under concurrent experiment condition. Cross-correlation analysis of fluorescence from the two detection volumes reveals the transport properties of the molecules, as well as the average molecular occupancy of the two volumes. Autocorrelation analysis of the fluorescence from the individual detection volumes then reveals the relaxation rate of the fluctuations. Finally, the photon counting histogram analysis, which distinguishes molecular species by differences of their molecular brightness, reveals the equilibrium distribution of the different conformers. This multi-parameter approach allows all the relevant dynamics information to be extracted from a single set of fluorescence intensity fluctuation measurements, eliminating the need for separate measurements on control molecules.

Through the studies presented in this dissertation, we found our two-beam FFS approach is indeed a powerful tool for the study of biomolecule conformational dynamics. Based on all the observation from two-beam FFS experiments, we proposed a *three-state*

mechanism for the folding and unfolding fluctuation of DNA hairpin structure. It consists of fast fluctuation between open conformation and intermediate states of DNA hairpins and considerably stable form of closed hairpin state. Our measurement determined the time scale of fast fluctuation as ~50μs and ~100μs for the hp-T_{21} and hp-T_{30}, respectively, with high precision. The reaction rate of the whole folding and unfolding process should be slower than ~2ms which is beyond the time limit of the FCS analysis. The intermediate state can be recognized as an ensemble of collapsed conformations with mismatched base pair(s). It should be noted that our observation is the first experimental report that shows the formation of stable intermediate in the process of DNA hairpin formation. Many theoretical models based on semi-flexible polymer chain theory, have recognized that there are short-lived, unstable intermediates resent in the transition state ensemble. Recently, large-scale, parallel, molecular dynamics simulation studies predicted the existence of stable intermediates. These simulations observed a nonspecific collapsed mechanism by which the hairpin folds analogous to the hydrophobic collapse in proteins. The individual conformations observed in the collapsed state represent an ensemble of misfolded traps with base-paring interaction and hydrogen bonding interactions between residues in the stem. Interestingly, there were some sets of single molecule experiments recently reporting slower reaction of hairpin structure: 1) Single molecule FRET measurement to monitor the conformational fluctuations of DNA hairpins immobilized on a glass surface, and 2) Mechanical force experiments with single

RNA hairpin to monitor the end-to-end distance and the reaction rate from the distance hops. The observations of much longer time scale in these experiments are often attributed to some artificial effects under the experimental conditions (surface interaction, presence of applied force, etc.). However, we suggested an alternative interpretation for the observations, which is consistent to the *three-state mechanism* for the folding and unfolding process as described above.

To elucidate the detailed mechanism of DNA or RNA hairpin formation, much more work should be done. Systemic investigation for the equilibrium distributions of the DNA hairpin conformers may answer many of the unresolved questions. We are extending the approach to address the dynamic behavior of DNA hairpins under various experimental conditions, such as various temperature, salt concentration, and solution viscosity. The kinetics of DNA hairpins with various loop lengths, base-pair size, and loop sequence are currently under investigation in our group. With more information from these further studies, we will obtain a more complete picture of the puzzle involving DNA hairpin conformational dynamics. In addition, based on our knowledge regarding the characteristics of these methods, we come to understand more complicated biological systems in the future. Folding in proteins and RNA are important future targets. Another future task would be applying this approach to the study of molecular processes in the cellular environment. We envisioned that our two-beam FFS approach with scanning

sample stage has great potential and versatility for these studies under equilibrium condition, where non-invasive and fast measurement is critical.

I complete this dissertation with a hope that the work presented here can make a valuable contribution toward the understanding of dynamic processes in biological systems.