SINGLE MOLECULE STUDIES OF A SHORT RNA

SINGLE MOLECULE STUDIES OF A SHORT RNA

A Dissertation Presented

by

PEKER MILAS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2014

Physics

SINGLE MOLECULE STUDIES OF A SHORT RNA

A Dissertation Presented

by

PEKER MILAS

Approved as to style and content by:

Anthony D. Dinsmore, Chair

Lori S. Goldner, Member

Jennifer L. Ross, Member

Ricardo B. Metz, Member

Rory Miskimen, Department Chair
Physics

DEDICATION

To mother, father, brothers; Ömer, Önder, and Hasan, and the owner of my heart Figen Güneş.

ACKNOWLEDGMENTS

I would like to thank my committee members, Prof. Anthony Dinsmore, Prof. Lori Goldner, Prof. Jennifer Ross, and Prof. Ricardo Metz for their helpful comments and input. Thanks to them, I learned to be more precise, and to pay attention to details especially in the process of writing.

I would also like to thank Goldner group members for their contributions to my work. Particularly, principal investigator Prof. Lori Goldner, group members and my collaborators Sheema Rahmanseresht, Benjamin Gamari, Louis Parrot, Kieran Ramos, and Nina Zehfroosh helped me a lot during my research which turned into my dissertation work.

I am very grateful to the helpful and patient collaborators in Ross Lab, Metz Lab, and Dinsmore Lab. Therefore, i would like to send my special thanks to Prof. Jennifer Ross, Juan Daniel Diaz-Valencia, Leslie Conway, Lynn Liu, Prof. Ricardo Metz, Abdülkadir Koçak, Gökhan Altınay, Prof. Anthony Dinsmore, and Jaime Hutchison.

During my graduate studies, i was honored to meet with Prof. Naci İnci who was my first Ph.D. advisor in Turkey and Prof. Carlo Dallapiccola who is the Graduate Program Director at Umass. To me, only the word "gentlemen" describes their personalities. Both were role models for me, not only on the technical side of my work but also in human relations.

Last but not the least, i would like to send my deepest thanks to my family and to my very special Figen Güneş. Without their moral support, and without their love, life would be so much harder. Likewise, i can hardly believe a five year research on an unknown could ever finish.

ABSTRACT

SINGLE MOLECULE STUDIES OF A SHORT RNA

SEPTEMBER 2014

PEKER MILAS, B.Sc., BOĞAZİÇİ UNIVERSITY

M.Sc., BOĞAZİÇİ UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Anthony D. Dinsmore

The material related with orientation of Cyanine dyes and their behavior at the ends of duplex RNA is also documented in [110]. Cyanine dyes are widely used to study the folding and structural transformations of nucleic acids using fluorescence resonance energy transfer (FRET). The extent to which FRET can be used to extract inter- and intra-molecular distances has been the subject of considerable debate in the literature; the contribution of dye and linker dynamics to the observed FRET signal is particularly troublesome. We used molecular dynamics (MD) simulations to study the dynamics of the indocarbocyanine dyes Cy3 and Cy5 attached variously to the $3'$ or $5'$ terminal bases of a 16 base-pair RNA duplex. We then used Monte Carlo modeling of dye photophysics to predict the results of single-molecule sensitive FRET measurements of these same molecules. Our results show that the average value of FRET depends on both the terminal base and on the linker position. In particular, $3'$ attached dyes typically explore a wide region of configuration space and the relative orientation factor, κ^2, has a distribution that approaches that of free-

rotators. This is in contrast to $5'$ attached dyes, which spend a significant fraction of their time in one or more configurations that are effectively stacked on the ends of the RNA duplex. The presence of distinct dye configurations for $5'$ attached dyes is consistent with observations made by others of multiple fluorescence lifetimes of Cy3 on nucleic acids. While FRET is frequently used as a molecular "ruler" to measure intramolecular distances, the unambiguous measurement of distances typically relies on the assumption that the rotational degrees of freedom of the dyes can be averaged out, and that the donor lifetime in the absence of the acceptor is a constant. We demonstrate that even for the relatively free $3'$ attached dyes, the correlation time of κ^2 is still too long to justify the use of a free-rotation approximation. We further explore the consequences of multiple donor lifetimes on the predicted value of FRET.

While providing detailed information about the individual members of a molecular ensemble, FRET technique is always limited by fluorophore brightness and stability. In the case of diffusing molecules, the experiment is further limited by the number of photons that can be collected during the time it takes for a molecule to diffuse across the detection volume. To maximize the number of photons it is common to either increase the detection volume at the expense of increased background, or increase the diffusion time by adding glycerol or sucrose to increase viscosity. As an alternative to current methods, here we demonstrated that water in oil nano-emulsions in perfluorinated compounds FC40 and FC77 can be used to confine biomolecules which results a dramatic increase in signal to noise ratios.

To efficiently use these attoliter volume (130 nm radius) aqueous containers, their characterization in terms of physical and chemical properties is necessary for both understanding the droplet environment and making better statements for single molecule experiment results within them. Characterization in terms of size is not an easy task because droplet system is a macroemulsion, so that droplets are kinetically stable. Therefore in addition to conventional dynamic light scattering (DLS) measurements,

we proposed a new method to measure droplet sizes using Mie scattering . The method was originally used in atmospheric physics and medical imaging and in our knowledge it was the first time of its use on an emulsion system.

We also measured the average pH within droplets by this new method using an absorptive dye (bromothymol blue). These measurements in combination with the emission measurements using another pH sensitive dye (fluorescein) showed that the droplet environment is acidic. Our initial attempts for changing the pH by adding a strong base to bulk aqueous samples prior to emulsification were succesful but the interpretation of result were not accurate because of working pH range of pH indicators.

Single molecule experiments within attoliter aqueous droplets provided supporting results for the low pH argument. We have seen that Cy3 only labelled duplex RNA ,in bulk experiments, was sensitive to ambient pH in the solution. This sensitivity appeared as a shift in the mean value of proximity histograms. The mean value of proximity ratio histograms at low pH (less than pH 7) was similar to the mean value of proximity ratio histograms we observed in droplet data. In addition, using photon counting histograms we identified a third brighter species in low pH solution data and in droplet data. All these materials related with single molecule works within droplets is also documented in [111].

In our knowledge, there is no analytical function given in the literature for the autocorrelation function of fluorescence correlation spectroscopy (FCS) data within droplets. Thus, we put effort for simulating the FCS of droplets. Although, this approach didn't provide us an analytical form, it was helpful for further understanding the internal environment of droplets. As a result of simulations, we observed that diffusion time within droplets long (on the order of a millisecond). In addition, we were able identify the source of different correlation times observed in FCS experimental data.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

FRET OF A SHORT RNA: EFFECT OF ATTACHMENT STRATEGY ON FREE ROTATION

The work in the current chapter is also documented in [110].

1.1 Background and Motivation

Fluorescence resonance energy transfer[49, 48] (FRET) is widely used in exploring the global structure or structural transformations of nucleic acids [25, 98]. Combined with single-molecule sensitive techniques, FRET provides a method for direct observation of the conformational changes of DNA and RNA [176, 65, 63, 184]. While observation of gross changes in distance between the donor and acceptor dyes is straightforward, the quantitative interpretation of FRET data to extract structure is generally complicated by dye photophysics[23] and linker dynamics.[154] As a minimum, quantitative interpretation of FRET requires that the configuration and dynamics of the dyes be understood or modeled.

Indocarbocyanine dyes Cy3 and Cy5 are commonly used to label nucleic acids for FRET. The location and orientation of these dyes attached to DNA[120, 78, 159] and DNA/RNA hybrids [79] have been the subject of some discussion recently, as has the behavior of closely related sulfoindocarbocyanines.[169, 148] Tethered to the 5′ end of a double-stranded A- or B-form helix, indocarbocyanines with structures and linkages as shown in Fig. 1.1 are known to spend a significant fraction of their time "stacked" on the double-stranded (ds) nucleic acid.[120, 78, 159] Steric hindrance and other intramolecular interactions are known to prevent the dyes from exploring all

orientational configurations[79, 97, 124]. Dye and dye-linker dynamics, as well as dye photophysics, can change significantly depending on the details of the local chemical and physical environment including, for example, the neighboring base and position of the dye linkage ($3'$ or $5'$ termini here). Indeed, MD simulations have already demonstrated differences in the free energy landscape of Cy3 terminally attached to different bases [159].

To explore the effects of linkage position and terminal base on FRET for dye-labeled RNA, we used a combined molecular dynamics (MD) - Monte Carlo (MC) approach. Molecular dynamics simulations with explicit solvent were run for dye-labeled dsRNA to extract the trajectory of the dyes with picosecond resolution over hundreds of nanoseconds. These trajectories were then used as the basis of a Markov chain Monte Carlo simulation of FRET that models the fluorescence from the donor and acceptor dyes. Using the MC simulation, we bridge the gap in timescales between molecular dynamics (MD) simulations ($< 1\mu$s) and typical measurements of FRET (time resolution $> 100\mu$s). To explore the effect of dye dynamics on FRET, we first took the dyes to have single fluorescent lifetimes and therefore single quantum yields;[157, 71] this assumption is commonly used in interpretation of FRET data. These results are compared with the more realistic case where the donor dye is permitted to have multiple quantum yields (and corresponding fluorescent lifetimes) that depend on the configuration of the dye on the RNA. For Cy3 and Cy5 dyes terminally attached to RNA as in Fig. 1.1, we find that FRET depends both on the dye linkage position, *i.e.*, $3'$ or $5'$ location, as well as the terminal base. For $5'$attachment, the dyes spend a significant fraction of their time in one or two states "stacked" on the double-stranded helix. For $3'$ attached dyes the situation is quite different; in most cases, the dyes explore a wide range of orientations about their tether. In all cases, FRET predicted from the Monte Carlo model is shifted significantly from what might be expected from freely-rotating dyes.

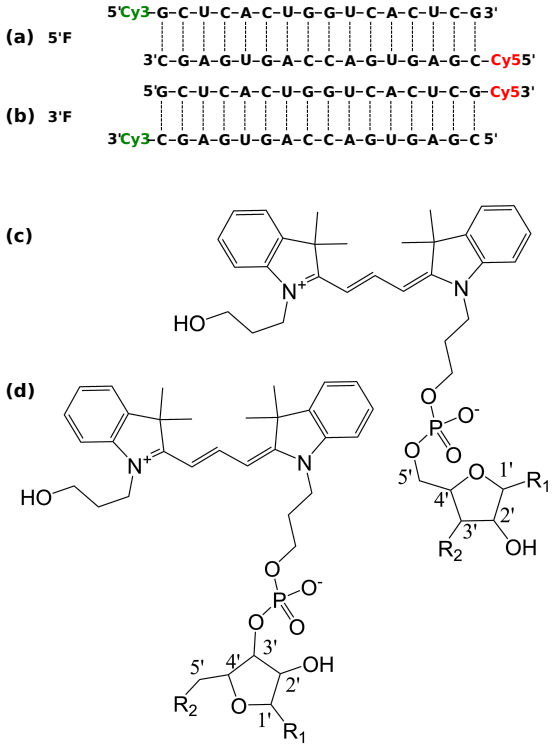

Figure 1.1. Duplexes, dyes, and linkers used in this study. (a) The 5′F duplex. 5′R is the same with dye positions swapped. (b) The 3′F duplex. 3′R is the same with dye positions swapped. (c) Dye attachment to the 5′ terminus. R1 is the terminal C or G and R2 is the phosphate of the next nucleoside. (d) Dye attachment to the 3′ terminus. R1 is the terminal C or G and R2 is the phosphate of the next nucleoside.

For either base-stacked or freely rotating dyes, the usual discussion of dye rotational freedom (or lack thereof) is obviated by the use of this MD/MC approach to modeling, since the dye positions are explicitly accounted for in the calculation of FRET. In the ideal dipole approximation, the efficiency of energy transfer from the donor to the acceptor dye is given by,

$$E = \frac{1}{1 + \left(\frac{R}{R_F}\right)^6}, \tag{1.1}$$

in which R is the distance between dyes and the Förster radius R_F is given by [25, 121],

$$R_F{}^6 = \frac{9c^4 J \eta_D \kappa^2}{8\pi n^4}. \tag{1.2}$$

In this expression, n is the solvent's refractive index, c is the speed of light, η_D is the quantum yield of the donor dye in the absence of the acceptor, and κ is an orientation factor defined below. The integral $J = \int f(\omega)\sigma(\omega)\omega^{-4} d\omega$ describes the spectral overlap of the donor emission and acceptor absorption: $f(\omega)$ is the fluorescence spectrum of the donor normalized such that its integral over all ω is one; $\sigma(\omega)$ is the molecular cross section of the acceptor; and ω is the angular frequency (not the wavenumber). For convenience, we also define the following quantity:

$$R_0 = \sqrt[6]{\frac{2}{3\kappa^2}} R_F \tag{1.3}$$

which is the commonly used value of R_F evaluated with $\langle \kappa^2 \rangle = 2/3$ for freely-rotating dyes.

To explore the effect of rotation alone, we initially held J and η_D constant, as is frequently assumed in FRET. To probe the effect of changes in the dyes' non-radiative processes that might accompany, for example, stacking on RNA, we added

to this model values of η_D that changed with the configuration of the dye on the RNA in the MD simulation. The quantum yield η_D can be expressed as:

$$\eta_D = \frac{k_{Dr}}{k_D} \tag{1.4}$$

where k_D, the decay rate of the donor excited state, is given by $k_D = \frac{1}{\tau_D} = k_{Dr} + k_{Dnr}$. Here τ_D is the excited state lifetime, k_{Dr} is the radiative decay rate of the excited state, and k_{Dnr} represents all other non-radiative decay processes. The orientation factor κ ranges from -2 to 2 and depends upon the relative orientations of the dye transition dipoles, whose unit vectors are denoted $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$\kappa = (\hat{\mu}_1 \cdot \hat{\mu}_2) - 3\left(\hat{\mu}_1 \cdot \hat{R}\right)\left(\hat{\mu}_2 \cdot \hat{R}\right). \tag{1.5}$$

Here $\hat{R}$ is the unit vector along $\vec{R}$, the displacement from the donor dye to the acceptor dye. Most often it is assumed that the dyes are freely rotating, so that the orientational average $\kappa^2 = 2/3$ can be used to estimate the rate of energy transfer. However, apparent orientational freedom is not by itself sufficient to justify the use of this average: correlations between R and κ^2 and long correlation times in either parameter (relative to τ_D, the lifetime of the donor) can modify FRET even when the relative orientation of the dyes has a distribution very similar to that of freely rotating dyes.[171, 58] These correlations are naturally included in the MD/MC model presented here.

Förster transfer described by Eq. 1.1 results from an ideal dipole approximation that is known to work well for cyanine dyes separated by at least 2 nm if they are freely-rotating, and at least 5 nm if they are statically oriented.[115] In this study all the dyes are moving to some extent, so the distance at which the ideal dipole approximation begins to lose validity lies somewhere between 2 nm and 5 nm. Given that dye separations are all roughly 5 nm in this work, the use of Eq. 1.1 is justified.

1.2 Simulation Methods

1.2.1 MD simulations of dye-labeled RNA

Molecular dynamics simulations were run to extract R and κ^2 trajectories needed for modeling FRET. Simulations were performed with Amber 11 and AmberTools 1.4 [20] using the FF99SB force field.[73] Each of the four different structures described in the caption of Fig. 1.1 were modeled.

The 16 base-pair RNA duplex was prepared using the Nucleic Acid Builder (NAB) package from AmberTools 1.4. Indocarbocyanine dyes Cy3 and Cy5, and the linkers for the dyes, are not among the default residues of the Amber package[20]. Therefore, we created models of the trans isomer for both dyes and their carbon linkers in Protein Data Bank (PDB) format and calculated their minimum energy conformation with Firefly 7.1.G [61]. This was performed using a 6-31G(d) basis set and density functional theory with the B3LYP1 functional. Restrained Electrostatic Potential (RESP) point charges were calculated[9, 24] for the Cornell *et al.* [27] force field using RED Tools vIII.3 [38]. These were used to parameterize the point charges for each atom in the new residues.

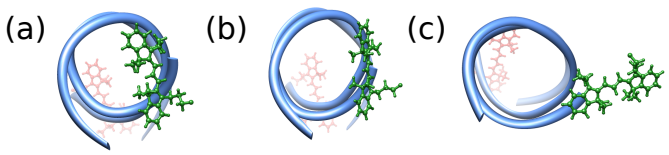

Figure 1.2. Typical snapshots of Cy3 attached to 5′ terminal G on the 5′F duplex:[127] (a) primary "base stacked" configuration with $r = 0.51$ nm, $\theta = 49.27°$, $R = 4.94$ nm and $\kappa^2 = 0.76$; (b) unstacked configuration with $r = 0.65$ nm, $\theta = 47.02°$, $R = 4.71$ nm and $\kappa^2 = 0.96$; (c) wandering dye with $r = 1.21$ nm, $\theta = -168.03°$, $R = 5.49$ nm and $\kappa^2 = 0.004$. Here r is the distance between the geometric centers of the terminal base-pair and the conjugated chain of the dye, θ is the twist angle between the dye and the nearest base-pair, and R and κ^2 are as defined in Eqs. 1.1 and 1.5.

The RNA, linkers, and dyes were combined into a single PDB structure. The dyes were initially oriented with the conjugated chain nearly parallel to the helical axis

of the RNA, extending outward from the end of the RNA. Using XLEAP, 22 Na^+ ions, corresponding to roughly 159 mM, were added to the system to neutralize the phosphates. The system was solvated in a truncated octahedral box with a 2 nm buffer of TIP3P water. Non-bonded interactions were calculated using Particle Mesh Ewald molecular dynamics (PMEMD) for electrostatics and a 1 nm cutoff for Van der Waals interactions. The Amber PMEMD.CUDA software was run on an nVidia GTX 480 graphical processing unit (GPU), which generated at most 10 ns per day.

Each duplex described in Fig. 1.1 was equilibrated for 20 ns prior to beginning production MD; a comparison of short- and long-term fluctuations in energy was used as an equilibration criterion, as described by Van Beek *et al.*[171] Following equilibration, production runs of 300 ns were done under conditions of constant pressure and temperature using a Langevin thermostat. Snapshots were saved every 1 ps. Typical configurations of Cy3 on 5′ terminal G from the 5′F duplex simulation are shown in Fig. 1.2. More configurations are presented in Movies 1-3 of the Supplement.

1.2.2 Monte Carlo modeling of FRET

To predict and model FRET, we used a Markov chain Monte Carlo model of a pair of interacting two-level dyes that utilized the MD trajectories for instantaneous values of the interdye distance R and the relative orientation represented by κ^2. The fluorescent and Förster processes accounted for in the Markov chain are shown in Fig. 1.3. The rate of donor excitation used here, $k_{exc} = 10^5\ s^{-1}$, is within the range of experimentally accessible values, determined by the laser intensity and the absorption cross section of the donor. The rate of excited state decay of the acceptor is given by $k_A = k_{Ar} + k_{Anr} = 1/\tau_A$ where k_{Ar} and k_{Anr} are the radiative and non-radiative decay rates of the acceptor, respectively. In the absence of the acceptor, k_D is known to be multi-valued for Cy3 terminally attached to nucleic acids, an effect which is attributed to a non-radiative decay rate k_{Dnr}, or to a donor quantum-yield η_D, that depends

on the dye's interaction with RNA.[70, 69] Triplet, isomerization, and charge transfer processes occur on timescales significantly longer than the total MC simulation time and so could not be accounted for here; in many cases it is possible to avoid the effect of these processes in experiments. We took J and the solvent refractive index n to have fixed values. The instantaneous rate of energy transfer, k_{FRET}, is given by,[48]

$$k_{FRET} = k_D \left(\frac{R_F}{R}\right)^6 = \frac{E}{1-E} k_D \tag{1.6}$$

where R_F is given in Eq. 1.2 and depends on κ^2 and η_D.

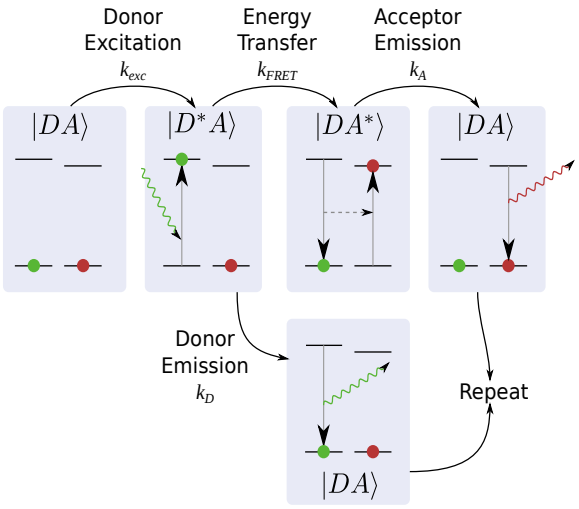

Figure 1.3. Markov chain diagram for the photophysical processes modeled here. Donor and acceptor relaxation can be radiative or non-radiative (not shown).

Simulations begin at $t = 0$ with the molecular configuration in the first frame of the MD trajectory and the dyes both in the ground state ($|DA\rangle$). The MC model then steps sequentially through the MD trajectory in increments of $\Delta t = 1$ ps. Sub-increments of $\Delta t = 100$ fs were also used to check for numerical accuracy; no significant differences were found.

The wait time, τ, for transitions from ground into the donor-excited state ($|D^*A\rangle$) is exponentially distributed assuming an excitation rate k_{exc}. From $|D^*A\rangle$ the system

has a probability, at each timestep, of either undergoing energy transfer (with rate k_{FRET}), decaying radiatively (with rate k_{Dr}), or decaying non-radiatively (with rate k_{Dnr}). If the system undergoes FRET, it ends up in $|DA^*\rangle$ from which state it can undergo radiative or non-radiative decay of the acceptor back to the ground state. The output of the MC simulation were photon emission times recorded on the radiative $|D^*A\rangle \rightarrow |DA\rangle$ and $|DA^*\rangle \rightarrow |DA\rangle$ transitions.

To determine when the system leaves an excited state, a Bernoulli variate ("true" or "false," for example) is drawn at each timestep with weight $p = \Delta t \ \sum_i k_i$, where the sum is carried out over potential destination states. If "true" is drawn, the system leaves the excited state and a destination state i will be drawn where $p_i = k_i / \sum_j k_j$.

When the end of the MD trajectory is reached, the simulation returns to the beginning of the trajectory. This periodic extension of the MD simulation results in an unphysical sampling in those cases where the system is excited upon reaching the end of the trajectory. However, with $\tau_D = 1$ ns,[70, 79, 34, 124] and 300 ns trajectories, this occurs no more than once for every 100,000 photons. Note that concurrent excitation of the dyes is unlikely in the weak excitation limit typical in most experiments, so this process is omitted from the model.

Throughout this work, the quantum yield of the acceptor is taken to be the same as the average quantum yield of the donor. This makes sense because the quantum yield of the acceptor does not modify Förster energy transfer, but it does appear in the experimentally relevant fluorescence detection-correction factor $\gamma = (\chi_A \eta_A)/(\chi_D \eta_D)$, where χ_A (χ_D) is the acceptor (donor) channel collection efficiency and η_A (η_D) is the acceptor (donor) quantum yield.[64] By equating the donor and acceptor quantum yields and omitting collection efficiency from the model we effectively set $\gamma = 1$, which makes the MC result for $\langle E \rangle$ directly comparable with γ-corrected FRET data.

For modeling FRET with a single-valued donor fluorescent lifetime, we take R_0 =5.8 nm. This value is within the range of 5.6 nm to 6.5 nm that has been reported in the litera-

ture for this dye pair.[181, 117, 135, 120] When multiple donor lifetimes are modeled, the value of R_0 changes as η_D changes during a trajectory. In this case, we take R_0 =5.8 nm at the population weighted average value of η_D.

1.3 Experimental materials and methods

Labeled RNA was purchased from Integrated DNA Technologies (IDT). Samples were prepared at a final concentration of 100 pM in HEPES-NaOH at pH 7.8 with 100 mM NaCl and 5 mM $MgCl_2$. Also included were 15 nM protocatechuate-3,4-dioxygenase (PCD) and 5 mM protocatechuic acid (PCA) as an enzymatic oxygen quenching system.[1] Single-molecule sensitive FRET measurements were accomplished using a confocal microscope with 514 nm excitation from an argon-krypton laser at 50 μW. Donor and acceptor channel photons were detected using homemade photon timing and laser control circuitry based on an FPGA[53] and two avalanche photodiodes (PerkinElmer SPCM-AQRH-15). All FRET data are fully corrected for background, crosstalk, and the detection collection factor γ. Further details regarding sample preparation, data acquistion and data analysis are in the Supplement.

1.4 Results and Discussion

1.4.1 MD simulations

Trajectories from the MD simulation of the 5′F duplex are shown in Fig. 1.4; trajectories for the 5′R, 3′F, and 3′R duplexes are found in the Supplemental Materials, Figs. 1-3. The distance between the geometric centers of the dye molecules, R, is shown in Fig. 1.4(a). The orientation factor κ^2, shown in Fig. 1.4(b), is calculated by approximating the transition dipoles to be parallel to the conjugated bond structure between the two indole-like moieties. The instantaneous value of energy transfer efficiency, shown in Fig. 1.4(c), is calculated from R and κ^2 using Eqs. 1.1, 1.3 and 1.2. To the right of each trajectory in Fig. 1.4 is a histogram of the plotted values.

Immediately evident from these trajectories is that there are changes in R and κ^2 at timescales both faster and slower than the average fluorescence lifetime of the donor, $\tau_D \approx 1$ ns.[70, 79, 34, 124] In some cases, structures fluctuating about a particular R or κ^2 persist for tens of nanoseconds; much shorter fluctuations are also evident. The existence of correlations at many timescales makes it difficult to extract a correlation time from a finite MD trajectory; autocorrelation decays for R and κ^2 can be found in the Supplementary Information. Cross correlations between R and κ^2 are also evident in some cases. For the 5′F duplex whose trajectory is shown in Fig. 1.4, R and κ^2 have a modest anticorrelation (Pearson's correlation coefficient) of -0.173 while these two parameters appear uncorrelated for the reverse 5′ configuration (Pearson coefficient of -0.008, Supplement Fig. 1). For the 3′F and 3′R duplexes, the Pearson coefficients are 0.037 and -0.133, respectively (Supplement Figs. 2, 3). It is worth noting that the correlation of R and κ^2 for the 5′F and 3′R duplexes appears to be due to the relatively static position of Cy5 on cytosine, whether 3′ or 5′ attached (discussed below, Fig. 1.5(a) and 1.6(a)). In these cases, changes in R and κ^2 are due primarily to excursions of Cy3, which accounts for the correlation between them.

While there are significant structural fluctuations that can be seen in the R and κ^2 trajectories of Fig. 1.4, the 5′ attached dyes spend a significant fraction of their time effectively stacked on or near the end of the RNA at a distance roughly twice the axial rise between base pairs. A two-dimensional histogram of dye configurations on the 5′ terminal base is shown in Fig. 1.5. The angle θ in Fig. 1.5 and 1.6 is given by $\arccos(\vec{Z}_d \cdot \vec{Z}_b)$, where $\vec{Z}_d$ is the direction along the conjugated carbon chain in the dye, pointing toward the free end of the dye. $\vec{Z}_b$ points along the adjacent base-pair axis, from the dye-attached nucleotide to the opposite nucleotide. It is defined by connecting the two C1′ atoms of the purine and complementary pyrimidine nucleotides. For dyes stacked on the end of the RNA, θ is therefore a twist angle. The distance r is defined to be between the geometric centers of the conjugated chain

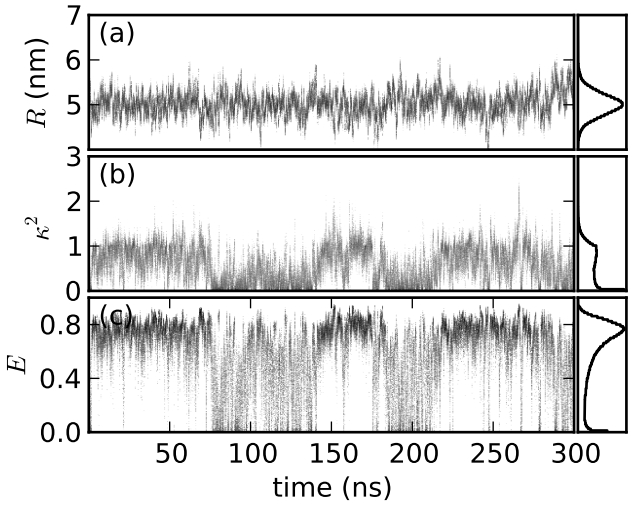

Figure 1.4. Trajectories and histograms for 5′F duplex. (a) The distance R between the geometric centers of the two dyes; a histogram of R values is on the right. (b) The orientation factor κ^2; a histogram of κ^2 values is on the right. (c) Instantaneous values of E; a histogram of E values is on the right. Here $\langle R \rangle = 5.01$ nm, $\langle \kappa^2 \rangle = 0.57$, and $\langle E \rangle = 0.580$. The Pearson coefficient between κ^2 and R is -0.173. To avoid plotting 300,000 points, each figure is a two-dimensional histogram 500 bins wide and 300 tall using grayscale where white corresponds to 0 occurrences and black corresponds to (a)$\geq$ 50 occurrences; (b)$\geq$ 100 occurrences; (c), black $\geq$ 30 occurrences. In all cases $\leq 0.1\%$ of the non-zero pixels are saturated.

and the base-pair. These definitions of the twist angle θ and distance r are similar to those in use elsewhere.[123]

The effect of stacking is particularly obvious for Cy5 attached to 5′ terminal C, Fig. 1.5(a), which spends the majority of its time with an average angular twist from the last base pair that is very close to the 30° helical twist of A-RNA (gray line in the figure). When attached to a 5′ terminal G, Fig. 1.5(b), the situation changes; there are now two locations that might be considered "stacked", one of which has θ very near zero degrees. Note that the peak near zero splits to avoid overlap of dye and base π-orbitals that occurs at 0°. Cy3 on 5′ terminal G, Fig. 1.5(d), also shows both peaks, although they show a larger distribution in both r and θ. For Cy3 on 5′ terminal C Fig. 1.5(c), a population is shifted to the zero degree peak, with a narrower distribution in r. Cy5 has only very rare excursions to large angles, while Cy3 exhibits occasional excursions to much higher angles and a wider range of distances. The boxed region in Fig. 1.5(c) represents 89% of the population. This is consistent with the work of Iqbal *et al.*[78], who found that FRET on a DNA/RNA hybrid is commensurate with dyes that spend 12% of their time freely rotating, and is otherwise found confined to a configuration that has lateral rotations within a 42° half-width half-maximum distribution. Our simulations suggest that Cy3 is responsible for most of the free rotation, and that the stacked configuration consists of several distinct states.

Although the stacked configurations near $\theta = 30°$ are still in evidence, comparison of Fig. 1.6 and 1.5 shows that cyanine dyes have substantially more configurational freedom when connected to the 3′ terminus. From Fig. 1.6(b), it is evident that Cy5 on 3′ terminal G explores a wide range of configurations in the course of the simulation. Cy3 shows a similarly large range of configurations on both 3′ terminal C (Fig. 1.6(c)) and 3′ terminal G (Fig. 1.6(d)). In all cases, the 3′ attachment appears to preclude the strong structural peaks on either side of $\theta = 0°$ that are favored for

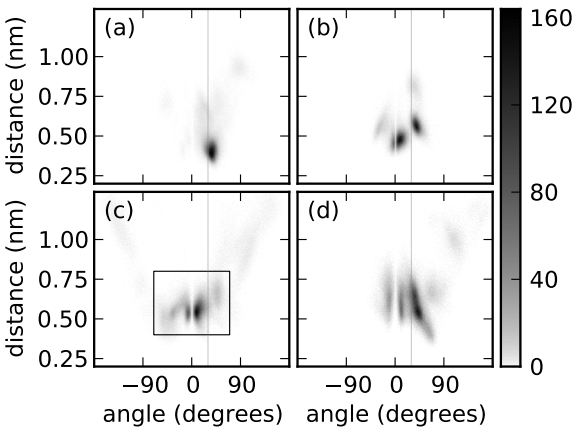

Figure 1.5. Two-dimensional histogram of distance r and angle θ for 5′ attached dyes. r and θ are defined in the text and describe the distance between and relative orientation of the dye and nearest base pair. (a) Cy5 on 5′ terminal C; (b) Cy5 on 5′ terminal G; (c) Cy3 on 5′ terminal C; (d) Cy3 on 5′ terminal G. (a) and (d) are taken from the 5′F duplex simulation. (b) and (c) are from the 5′R duplex simulation. The vertical line at 30° corresponds to the helical twist of A-RNA. The boxed region in (c) runs from −70° to 70° and 0.40 nm to 0.80 nm; it contains 89% of the population. Note that to bring out sparse populations, the color scale is proportional to the population raised to the power 0.6.

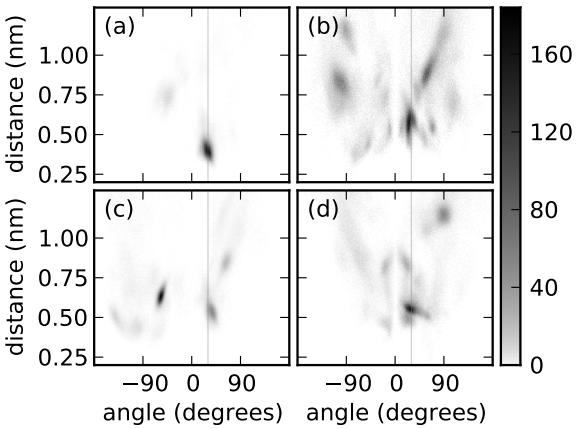

Figure 1.6. Two-dimensional histogram of distance r and angle θ for 3′ attached dyes. r and θ are defined in the text and describe the distance between and relative orientation of the dye and nearest base pair. (a) Cy5 on 3′ terminal C; (b) Cy5 on 3′ terminal G; (c) Cy3 on 3′ terminal C; (d) Cy3 on 3′ terminal G. (a) and (d) are taken from the 3′R duplex simulation. (b) and (c) are from the 3′F duplex simulation. The vertical line at 30° corresponds to the helical twist of A-RNA. Note that to bring out sparse populations, the color scale is proportional to the population raised to the power 0.6.

most of the 5′ attached dye structures. If so, then the loss of the peaks near $\theta = 0°$ may account for some of the additional rotational freedom of 3′ attached dyes. Of the 3′ configurations studied here, only Cy5 attached to a 3′ terminal C, Fig. 1.6(a), shows little evidence of configurational freedom.

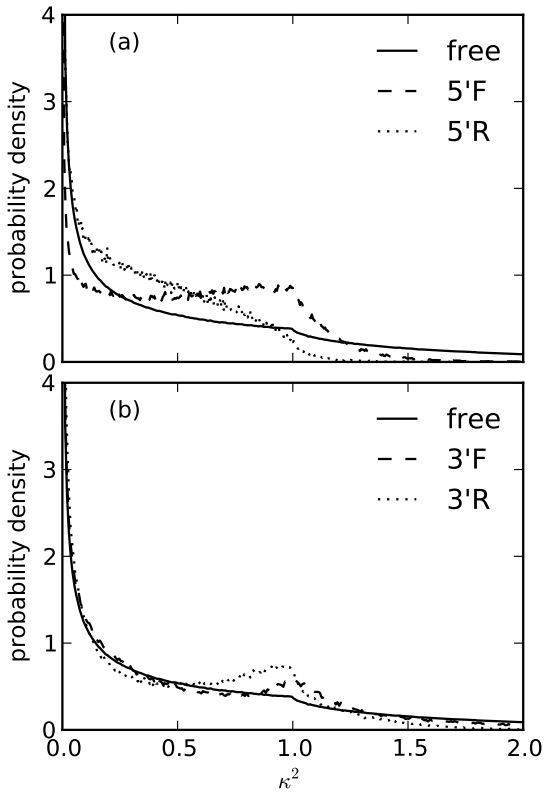

Figure 1.7. (a) Distribution of κ^2 for 5′F and 5′R duplexes for $0 < \kappa^2 < 2$. (b) Distribution of κ^2 for 3′F and 3′R duplexes. In both cases, the distribution of κ^2 for freely-rotating dyes is shown for comparison.

The greater rotational freedom of these 3′ attached dyes is reflected in their κ^2 distributions, which are shown in Fig. 1.7. With dyes attached to the 5′ end of the RNA, Fig. 1.7(a) shows significant deviations from the freely rotating distribution for all values of κ^2. The κ^2 distributions for the 3′ attached dyes are much more similar to the freely rotating distribution, differing only in the peak region around

$\kappa^2 = 1$, and in the tail at high κ^2. Despite the differences in the distributions, the average κ^2 values are similar for 3′ and 5′ attachment with $\langle\kappa^2\rangle = 0.57$ and 0.36 for the 5′F and 5′R duplexes respectively, and $\langle\kappa^2\rangle = 0.53$ and 0.43 for the 3′F and 3′R duplexes. All four systems yield $\langle\kappa^2\rangle$ values that are below the freely rotating value of 2/3. This is at least partly due to the fact that high values of κ^2 correspond to the case where the conjugated chain of the dyes would need to be aligned along the helical axis: most of these configurations are sterically forbidden. This accounts for both the slightly low value of $\langle\kappa^2\rangle$ and the smaller populations (compared with the free-rotation distribution) at $\kappa^2 > 1.5$.

1.4.2 Monte Carlo model and consequences for FRET

Because the MD simulations discussed above show fluctuations in R and κ^2 occurring at timescales that are both faster and slower than τ_D (Fig. 1.4 and Supplement Figures 1-3,8 and 9), and because there are correlations between R and κ^2 in the 5′F and 3′R duplexes, there is no simple approximation of FRET that can be used here. Instead, we used the Monte Carlo model discussed above to predict FRET histograms in the four systems of Fig. 1.1.

1.4.2.1 Consequences of rotation

To investigate the effects of rotational motion separate from changes in lifetime and quantum yield, we first made a simplifying approximation in which the non-radiative processes are assumed to be independent of dye configuration, so that k_D is single-valued and constant throughout the simulation. For this study we used a radiative lifetime for Cy3 of $\tau_{Dr} = 2.0$ ns as determined by Sanborn *et al.*[148] We further set $\eta_D = 0.5$, so that the $\tau_D = 1$ ns, which is an intermediate value for Cy3 on nucleic acids.[70, 79, 34, 124]. Below we consider other values of η_D. To facilitate comparison with γ-corrected data, the quantum-yield of the acceptor was set equal to that of the donor as discussed in the Methods section on Monte Carlo modeling of

FRET. The lifetime of the acceptor is irrelevant to the outcome of the model so for convenience we set $\tau_A = 1$ ns.

As described in the Methods section, the output of the MC simulation were donor and acceptor photon arrival times. For purposes of constructing a FRET histogram, photons generated by the simulation were binned into arrival time intervals (bins) with time width Δt chosen such that the mean number of photons per bin $\langle N_A + N_D \rangle = 100$. Here N_A (N_D) refers to the number of acceptor (donor) photons in a particular bin, and $\langle N_A + N_D \rangle = 100$ is typical for an experimental single molecule trajectory. From the binned list of photon counts we computed a corresponding list of ratios, $E = \frac{N_A}{N_A + N_D}$. This FRET trajectory was then used to construct the histogram shown in Fig. 1.8(b) and 1.8(d) for 5′ and 3′ attached dyes, respectively. For comparison, histograms of the instantaneous value of E from the corresponding MD trajectory are shown in Fig. 1.8(a) and 1.8(c). The MC simulations were run until a total of approximately 10^6 photons were generated, giving a run-to-run standard deviation in the value of $\langle E \rangle$ of 0.0005.

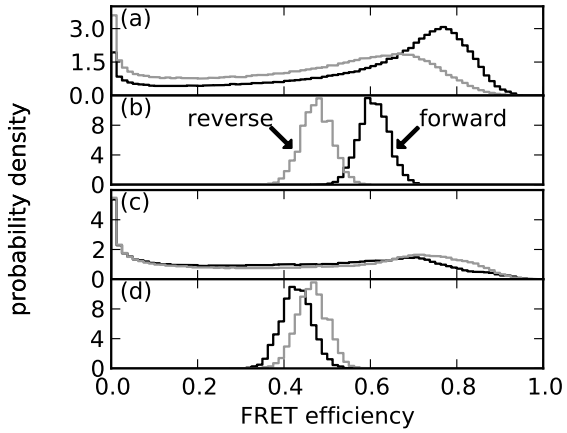

Figure 1.8. (a) histogram of instantaneous E from the MD trajectory of the 5′F (black) and 5′R (gray) duplexes; $\langle E_{inst} \rangle = 0.580$ and 0.447, respectively. (b) MC predicted distribution of E for the 5′F and 5′R duplexes; $\langle E \rangle = 0.638$ and 0.509, respectively. (c) histogram of instantaneous E from the MD trajectory of the 3′F and 3′R duplexes; $\langle E_{inst} \rangle = 0.402$ and 0.392, respectively. (d) MC predicted distribution of E for the 3′F and 3′R duplexes; $\langle E \rangle = 0.452$ and 0.445, respectively.

The donor quantum yield for Cy3 on DNA has been measured between 0.16 and 0.39,[148, 69, 104, 79] Since quantum yields are difficult to determine with accuracy, we take a modeling approach and consider the change in $\langle E \rangle$ with values of of η_D that are both higher and lower than those found in the literature. The results for $\eta_D = 0.5$ are reported above in Fig. 1.8. For $\eta = 0.1$, the values are shifted by only -3.4% to -4.8% and the results are given in Table 1. Note that for all entries in Table 1, $\tau_{Dr} = 2.0$ ns and $R_0 = 5.8$ nm. Changes in η_D are assumed to come from changes in k_{Dnr}.

In the case that the fluctuations in R and κ^2 are separately either faster or slower than the donor lifetime, analytical expressions derived by Gopich and Szabo[56, 57] can be used to estimate $\langle E \rangle$. Comparison of the MC predicted values for $\langle E \rangle$ with those of the analytic expressions gives insight into the effect of the observed correlations in R and κ on FRET; the various results are given in Table 1. These approximations require as input the distributions for R and/or κ^2; we obtain these distributions from our MD simulation. As should be expected, the limit of slow rotation and slow distance fluctuation ("slow-slow" in Table 1) reproduces the average of instantaneous FRET from the MD simulation alone (given in the captions of Fig. 1.4 and Supplement Figs. 1-3). It is also interesting to note that substitution of $R = \langle R \rangle$ and $\kappa^2 = \langle \kappa^2 \rangle$ into Eqs. 1.1 and 1.3, gives a result within 0.01 of the "fast-slow" approximation, a result which occurs in this approximation when the distance fluctuations are small on the scale of R_0. From Table 1 it is evident that the results for $\langle E \rangle$ from the MC model fall in a gap between approximations that assume fast and slow orientational motion. Note that the assumption of slow or fast changes in distance are relatively unimportant, especially for the 5′F and 5′R systems, which should not be surprising given the relatively small fluctuations in R for those molecules. Still, the assumption of fast fluctuations in R always gives higher FRET than the assumption of slow fluctuations, as has been discussed elsewhere.[151] It is also clear that the 3′R

duplexes give results for $\langle E \rangle$ that are very close to the slow-rotation limit, despite their relative rotational freedom.

			FRET Efficiency					
			orientation-distance				Monte Carlo	
duplex	$\langle R \rangle$ (nm)	$\langle \kappa^2 \rangle$	fast-fast	fast-slow	slow-fast	slow-slow	$\eta = 0.5$	$\eta = 0.1$
5'F	5.01	0.57	0.69	0.67	0.59	0.58	0.64	0.62
5'R	5.08	0.36	0.56	0.55	0.46	0.45	0.51	0.49
3'F	5.50	0.53	0.56	0.53	0.42	0.40	0.45	0.43
3'R	5.34	0.43	0.56	0.52	0.42	0.39	0.44	0.42

Table 1.1. Mean values of R and κ^2 from the MD simulations, along with $\langle E \rangle$ predicted using both analytic approximations[56, 57] and the MC model presented here. For all approximations and simulations, $\gamma = 1$ and $R_0 = 5.8$ nm.

While there is no simple analytic expression for estimating FRET when fluctuations occur on timescales both shorter and longer than τ_D, it is possible to approximate $\langle E \rangle$ using integrals over short sections of the MD trajectory,[56] as described in the supplement of Ref. [11]. We compared the results of the MC model to this integration, sampling 50,000 starting points and taking the integration out to 15 ns. We found no significant differences between the two methods: agreement in $\langle E \rangle$ was better than 0.001 in all cases and the difference could be attributed to sampling error.

1.4.2.2 Consequences of multiple fluorescence lifetimes

It is quite natural to include in these simulations changes in the quantum yield of the donor, η_D, that may be associated with changes in the configuration of the dye. Multiple fluorescent lifetimes are consistently found for Cy3 on DNA[78, 124]. These include a fast component, near 0.3 ns, that is close to that of Cy3 in solution and that is identified with dyes that are free to undergo excited state isomerizations [6]. Excited state isomerization introduces a non-radiative decay pathway that increases k_{Dnr} thereby lowering the fluorescence lifetime and quantum yield [6]. Longer lifetimes are associated with Cy3 that is more tightly confined, for example by interactions with the DNA, which reduces the isomerization rate.

Of the many measurements of Cy3 lifetime on nucleic acids, [70, 79, 34, 124] the most relevant here was for Cy3 attached to a 5′ terminal cytosine on an RNA/DNA duplex[79] which were consistent with three component fluorescence decays with τ_{D1} = 1.77 ns, τ_{D2} = 0.96 ns and τ_{D3} = 0.31 ns.[79] Most of the steady-state fluorescence signal was associated with τ_{D1} and τ_{D2} and only 5% of the intensity was attributed to τ_{D3}.[79] In the absence of lifetime measurements for Cy3 on the RNA modeled here, we use these values for τ_{D1},τ_{D2}, and τ_{D3} to explore the consequences of multiple lifetimes on FRET in the 5′R duplex, which also contains a Cy3 attached to a terminal cytosine.

While many different degrees of freedom might be used to assign configurational states associated with these three lifetimes, we started here with the simplest parameter that we could identify, namely the distance between the dye and the terminal base pair as defined previously for Fig. 1.5 and 1.6. For the 5′R duplex, the histograms of dye-to-base distance, shown in Fig. 1.9, are reasonably fit by a Gaussian mixture model[126] with three Gaussians, which we associate with three different configurational states of the dye, labeled I, II, and III. A maximum likelihood inference was used to assign the states shown in Fig. 1.9. Since the fluorescence lifetime of cyanine dyes increases with the decrease in isomerization that occurs in more rigid or confining environments,[148, 26, 18] we associate the narrowest distribution (I), which is closest to base-stacked, with the longest lifetime. The broadest distribution (III), which also has the largest average distance between the dye and base pair, is associated with the shortest lifetime. Consistent with the results of Iqbal *et al.*[79], the latter population is also by far the smallest. Complete Gaussian mixture fit results are given in the caption of Fig. 1.9.

To include multiple lifetimes in the MC model, each step in the MD simulation was assigned one of the states I, II, or III as described above. For state I, $k_{D1} = 1/\tau_{D1}$; for state II, $k_{D2} = 1/\tau_{D2}$; and for state III, $k_{D3} = 1/\tau_{D3}$. Again assuming that the

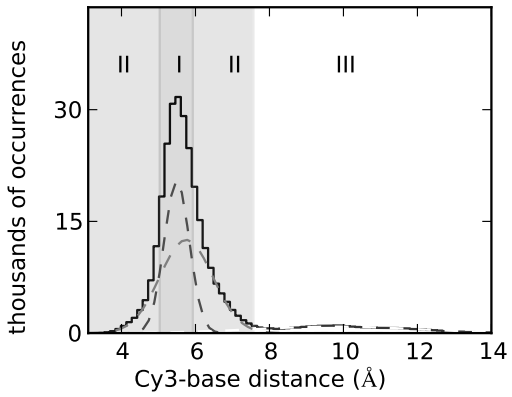

Figure 1.9. Histogram of the distance between Cy3 and the 5′ terminal base pair for the 5′R duplex. The parameters resulting from this fit are: stacked state I (dashed dark gray curve) mean $\langle r \rangle = 5.49$ nm, standard deviation $\sigma = \sqrt{\langle (r - \langle r \rangle)^2 \rangle} = 0.12$ nm, weight $w = 0.39$; stacked state II (dashed light gray curve) $\langle r \rangle = 5.75$ nm, $\sigma = 0.51$ nm, $w = 0.50$; unstacked state (dashed black curve) $\langle r \rangle = 9.79$ nm, $\sigma = 4.87$ nm, $w = 0.11$.

underlying radiative rate is given by $\tau_{Dr} = 2.0$ ns,[148] the quantum yields were calculated using Eq. 1.4 to be $\eta_{D1} = 0.89$, $\eta_{D2} = 0.48$, and $\eta_{D3} = 0.16$. Using these quantum yields, R_0 was also assigned appropriately for each state, subject to the constraint that $R_0 = 5.8$ nm at the mean quantum yield of $\langle \eta_D \rangle = 0.68$. That this value of $\langle \eta_D \rangle$ is higher than might be expected[148, 69, 104, 79] underscores the need for better determinations of quantum yield and radiative lifetime in future work, but it does not prevent us from exploring the general consequences of multiple lifetimes on FRET histograms.

With a value of $\langle \eta_D \rangle$=0.68, and using multiple lifetimes as discussed above, we find $\langle E \rangle = 0.5095$ from the MC model of the 5'R duplex. If we instead use the same value of η_D but assume only a single fluorescent lifetime given by $\tau_D = \eta_D / k_{Dr} = 1.36$ ns, we find $\langle E \rangle = 0.5139$. Unlike the effect of including rotational dynamics, which give a substantial shift in $\langle E \rangle$, the inclusion of three separate lifetimes makes only a small change (-0.0044) in the calculated value of FRET. In addition, because the changes in

dye configuration seen in the MD simulation are faster than the interphoton time, we would not expect to see any broadening due to the inclusion of multiple lifetimes. In all cases the resulting distributions are homogeneous with widths that are shot-noise limited.[55]

1.4.3 Comparison with data

Single-molecule sensitive solution FRET data for the 5′R duplex, corrected for background, crosstalk, and γ are shown in solid gray in Fig. 1.10. Data are acquired on molecules freely-diffusing in solution at a concentration of 100 pM. The details of data analysis are discussed in the Supplemental materials. To construct the FRET histogram, photon arrival times in each of two channels (donor and acceptor) were histogrammed into 5 ms bins, with bursts of photons indicating the presence of a labeled RNA molecule in the detection volume. Bins with more than 20 photons (sum of both channels) were used in calculations of FRET. This threshold resulted in an average of 32 photons per bin. The FRET efficiency distribution was fit to a mixture of two beta distributions. The component corresponding to the FRET peak is shown in Fig. 1.10 as a continuous solid black line. The lower peak corresponds to donor-only molecules and it not of interest here. From the FRET data we find $\langle E \rangle = 0.540 \pm 0.003$ from 1713 above-threshold bins, which is a bit higher than the prediction using $\eta_D = 0.68$ and $R_0 = 5.8$ nm, above. Data taken on the 5′F duplex was very similar, with only a small positive shift in FRET: $\langle E \rangle = 0.576 \pm 0.008$ from 335 above-threshold bins. The uncertainty in both cases is the standard error on the mean of the distribution; it does not include the larger contribution arising from the uncertainty in γ.

From this work it is clear that measurements of the τ_D's and τ_{Dr} or $\langle \eta_D \rangle$, specific to each molecule, are critical to the success of these models. For example, for the 5′R duplex, bringing the model into agreement with the data requires modifying η_D

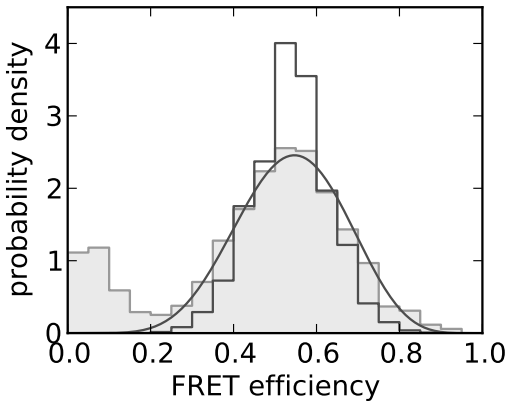

Figure 1.10. A comparison of data and model. The model used here includes the effect of multiple lifetimes associated with different configurational states of the dye, although the result is nearly indistinguishable from that of a single-lifetime model at the same $\langle \eta_D \rangle = 0.68$. FRET histogram from data is shown filled with gray. The smooth solid black line is a best-fit to the data of a beta function, giving $\langle E \rangle = 0.540 \pm 0.003$. FRET histogram from the model, with $R_0 = 5.9$ nm at $\langle \eta_D \rangle = 0.68$, is also shown in a solid black line. Data have been corrected for crosstalk, background and gamma as described in the text.

or R_0 or both. Since η_D is already considerably higher than might be expected, and would need to be further raised to bring the model into better agreement with the data, we ran the MC simulation at various values of R_0, and found that for $R_0 = 5.9$ nm, $\langle E \rangle = 0.536$ (with 27,000 bins in the simulation the statistical uncertainty is negligible), which brings the model into agreement with the data.

To model the FRET histogram of Fig. 1.10, we binned the simulated data with an average of 32 photons per bin. The resulting shot-noise limited histogram is shown as a dark gray outline in the figure. The width of the data histogram, with a standard deviation of 0.128 is slightly larger than what would be expected from shot noise,[56] for which the standard deviation is 0.099. Broadening in Cy3-Cy5 FRET histograms has been previously reported and attributed to long-lived states of the acceptor dye that are not accounted for in this model.[86]

CHAPTER 2

CHARACTERIZATION OF SUB-MICRON SIZE DROPLETS FOR SINGLE-MOLECULE STUDIES

2.1 Background and Motivation

Compartmentalization through the water-in-oil (W/O) emulsions is a relatively new method among single-molecule techniques [108, 15, 139, 39, 43, 180, 35, 137, 10, 163, 82, 29, 167, 94, 77, 54, 47, 7]. The method inherently removes the necessity of tethering the biomolecules to a solid surface i.e glass substrate, so it minimizes the possible substrate-target molecule interactions. It is also cost and labor effective in terms of the preparation times as well as the amount of materials used in a typical single-molecule experiment.

Although capillary action based applications (microfluidic devices and injection techniques) are frequently seen in literature [47, 108, 94, 77, 43, 180], other emulsification methods such as application of low power ultra-sound were also applied to single molecule studies [15, 139, 93, 137, 163, 54]. Former method provides large (typically larger than 1 μm) but highly uniform droplet sizes [4, 106, 163, 17], thanks to sophisticated device designs. The latter, provides much smaller (typically smaller than 1 μm) but very frequently a heavy-tailed distribution of sizes [37, 46, 107, 130] without requiring any additional effort on instrument design.

Regardless of the emulsification method, perfluorinated compounds (PFCs) are frequently used for continuous phase liquids. The initial design idea of PFCs was to use them as artificial blood substitutes because of their very high carriage capacity (up to 40% of volume) for non-polar gases [168, 83, 100, 101, 146, 90], such as oxygen.

In time, their inert nature to both hydrocarbon compounds and water as well as their generally very low surface tension [101, 140, 158] drew researchers' attention to use them in compartmentalization studies and even in drug delivery systems [95, 134, 158, 122].

The environment within sub-micron size droplets is also very interesting. First of all, aqueous droplets are constrained volumes for diffusing molecules. Assuming it is merely a spatial constraint, both rotational and translational diffusion near the droplet boundary expected be much slower [60, 99, 102, 62, 19, 142, 44] in the nanometer length scale. In addition, electrostatic interactions between free ions and a nearby interface are shown to alter interfacial tension in liquid-liquid case [22, 85, 152]. These interactions also produce concentration gradients for ions within the space between the interface and the bulk solvent [131, 91, 172].

Chemistry in droplet environment is even more complex. IR spectroscopy based measurements on small micelles (23 nm or smaller in diameter) show that water near the center of a micelle exhibits similar behavior to bulk water, whereas molecules near the boundary (within 1 nm distance) organized in a completely different way so that they are thought to be in a frozen state [182, 128, 129, 45, 113, 112, 103, 156]. This difference in water's hydrogen network, directly affects processes ranging from proton diffusion to protein folding in a micelle. Furthermore, reported redox and charge transfer reactions near the droplet boundary of larger systems (sub-micron diameter) can be found in food emulsions research [28, 150, 107, 75, 160, 68]. These observations were indirectly supported by other measurements showing the saturation of interface by hydroxyl ions which changes the ambient pH [105, 166].

Putting everything together, droplet environment has its own physical and chemical characteristics which can be completely different than bulk. To address these differences as well as to make better evaluation of experimental result characterization studies are necessary prior to their utilization. Several commonly used charac-

terization techniques for size determination are microscopy, dynamic light scattering (DLS), and nuclear magnetic resonance (NMR) [107], among which DLS and NMR are the ones can be used for the sizes about 200 nm. Tools for characterizing internal environment of droplets are few in number. Most of the research in literature are conducted either by IR spectroscopy techniques or NMR and focused mostly on the behavior near the interface. In our knowledge, particularly for the sizes about 200 nm, there is no provided work in the literature about chemistry within droplets.

In the present work, we propose a new and effective method for measuring the size and the pH within aqueous droplets simultaneously. The method was originally developed and used in the field of atmospheric physics for determination of particulate matter sizes in atmosphere [3, 84, 59, 40, 138]. Later on, it was adapted to medical imaging for understanding the tissue structures [149, 174, 114]. The only requirement of the method is to have sub-micron size droplets in emulsion which are considered as Mie scatterers in the visible wavelength range.

Results of our method were found to be in good agreement with the other conventional methods, such as DLS (in the case of size determination) and ratio-metric pH determination using fluorescence emission.

2.2 Materials and Methods

2.2.1 Droplet Making Method

Previously reported droplet making method [54] revised and re-optimized. For the stability purposes according to **Bancroft Rule** [107, 161], previously used hydrocarbon based surfactant Triton X-100 replaced with a PFC based surfactant [72] (obtained from RainDance Technologies) which is a block copolymer consisting of oligomeric perfluorinated polyethers (PFPE) and polyethylene-glycol (PEG) as shown in Fig. 2.1.

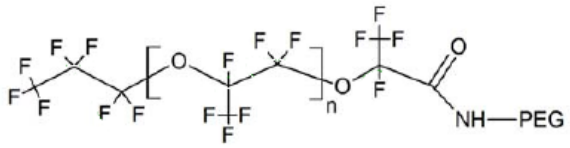

Figure 2.1. Structure of PFC based surfactant from Ref.[72].

To increase the initial dispersion efficiency, continuous phase PFC prior to ultrasonication, were first de-gassed using a Schlenk line system, then saturated with high purity N_2 as reported elsewhere [51, 52, 125]. To achieve kinetic stability as well as the reproducibility in method, emulsion parameters were tuned according to suggestions in literature. Among these, sonication time and salt concentration within droplets were adjusted for reducing the creaming rate [46, 107, 144, 118] by decreasing the average droplet sizes to approximately 200 nm. Flocculation on the other hand reduced by first choosing a high viscosity PFC (FC-40) for emulsification. Secondly, volume to volume ratios of PFC, surfactant, and water based dispersed phase material were fixed to 1000:1:10 and so the concentration of droplets fixed in the system for further reduction [164, 80, 107].

Finalized procedure for making kinetically stable droplets was as following; prior to droplet making 10^{-2}(v/v) surfactant was dissolved in FC-40 by adding 500 μl of stock surfactant solution (the liquid came from RainDance Technologies as it is) into 50 ml FC-40. This solution was then vortexed until there is no residue observed at the bottom of conical tube. De-gassing of the solution was achieved by the freeze-pump-thaw method as follows; surfactant solution was placed in a sealed Schlenk flask and frozen in liquid nitrogen. The flask was then opened to vacuum and pumped to 12 mtorr, re-sealed and thawed in a warm water bath. After sitting for 30 minutes the process was repeated up to five times. At the end of the de-gassing process, samples were saturated with high purity N_2. Here the term saturation refers sending the N_2

into the sealed flask while keeping the flask in a solid CO_2 and isopropanol bath and stirring it frequently. This procedure allows the diffusion of N_2 into de-gassed liquid with an increasing rate by cooling it down and so increasing the gas solubility within. Following the saturation, liquid within the flask was filtered by passing through 0.2 μm millipore filters. Filtered liquid transferred to 50 ml conicals, and conicals were sealed with parafilm. Finally, they stored in a desiccator which had a N_2 saturated atmosphere. Same procedure was applied to pure FC-40, before they were used for diluting the emulsions.

Emulsions prepared through the following steps; 10 μl 10^{-2}(v/v) de-gassed surfactant solution in FC40 was added into 190 μl de-gassed FC-40 (surfactant dilution to 10^{-3}(v/v)). 2 μl of dispersed phase was placed into this solution. Here, the dispersed phase always consisted of 20 mM Tris at pH 7.8 and 200 mM NaCl. To see the initial dispersion, combined system was vigorously shaken 10-20 seconds. After shaking, mixture was sonicated 2 minutes in an ultrasonic cleaner (Branson 1510). For consistency, throughout the sonication process, sample container (a 1.5 ml flat top Eppendorf tube) was placed very close to the center of the water tank and about 2 cm above the bottom of the tank which was supposed to be the focal point of the transducer [109]. The process provided 200 μl W/O emulsion in chosen PFC. A bright field microscope image of typical sample prepared by finalized droplet making procedure is shown in Fig. 2.2.

2.2.2 DLS Measurements

DLS measurements were conducted with a Brookhaven Instruments BI-200SM goniometer equipped with a PCI BI-9000AT digital correlator, a temperature controller, and a solid-state laser (model 25-LHP-928249, $\lambda = 633$ nm). The physical location of DLS instrument was The Silvio O. Conte National Center for Polymer Research building Room B 564 of University of Massachusetts Amherst. Samples

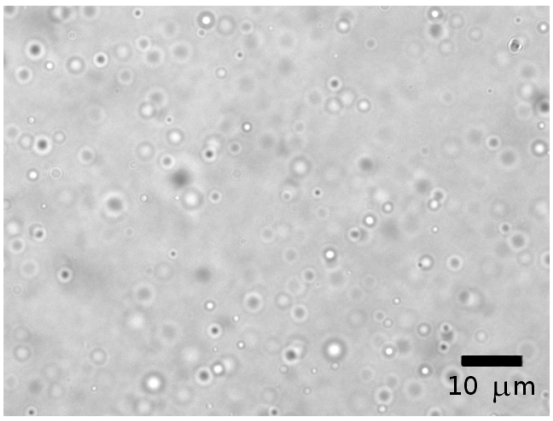

Figure 2.2. Bright field image of a typical droplet sample.

were prepared as explained in Sec. 2.2.1, by keeping FC-40 as continuous phase liquid. Following the emulsification, whole emulsion product was diluted 25 times by adding 200 μl emulsion product into 4.8 ml de-gassed FC-40. Diluted samples were placed into 5 ml nominal volume borosilicate glass culture tubes (Fisher Scientific, Cat. 14-961-25). Emulsion was mixed by gently turning the tube upside down several times. Total of 5 ml diluted and mixed emulsion was used as the actual sample for measurements.

Sample preparation and the measurements on the average took less than 5 minutes for each experiment. Measurements were carried out at room temperature and at the scattering angle of 90°. Correlation functions of the scattering data were analyzed through built-in software on DLS apparatus via non-negative least squares fit (NNLS) [143, 183, 175] and then used to determine diffusion coefficients (D). D can be converted into hydrodynamic radii using the Stokes-Einstein equation:

$$r = \frac{kT}{6\pi\eta D} \tag{2.1}$$

To observe shot-to-shot changes, all measurements were done on freshly prepared samples. Build-in software generated intensities normalized and a lognormal func-

tional form as shown in Eq. B.8 corresponding the probability density function was fitted to normalized intensities.

2.2.3 Absorption/Attenuation Measurements

Absorption/attenuation measurements were conducted with a GE Healthcare / Amersham Biosciences Ultrospec 3100 Pro UV/Visible Spectrophotometer, equipped with a Xenon lamp as its light source. Instrument had a working wavelength range of 200 to 900 nm with a bandwidth of 3 nm. Data acquisition was done within the wavelength range of 400 nm to 850 nm. The physical location of UV/Visible Spectrophotometer was Hasbrouck Lab, Room 310 of University of Massachusetts Amherst.

Samples for size measurements, were prepared as explained in Sec. 2.2.1. Continuous phase PFC was FC-40 in all experiments. Following the emulsification, dense droplet sample was diluted 10 times by adding 100 μl of the emulsion into 900 μl de-gassed FC-40. System was mixed gently by pipette action. Homogeneous looking mixture was placed into 1 ml nominal volume disposable UV cuvettes (Fisher Scientific, Semimicro; Methacrylate 14-955-128). The path length of the cuvettes was 1 cm. Experiments were all carried out at room temperature and repeated at least three times on each sample.

For pH measurements, bromothymol blue in powder form (BTB; Sigma-Aldrich, 114413-5G) was first dissolved in high purity dimethyl sulfoxide (DMSO; Sigma-Aldrich, 41641-250ML) to prepare 10 mM stock solution. Stock BTB solution is then diluted 100 times to 100 μM in buffer solution of 20 mM Tris at pH 7.8 and 200 mM NaCl. For increasing the solubility of BTB, 100 μM final solution had 5% (v/v) DMSO in it. This solution was used as dispersed phase. The rest of the sample preparation was same as explained in Sec. 2.2.1. Similar to size measurements, dense emulsification product was diluted in FC-40 10 times by adding 100 μl droplet solution

in to 900 μl pure de-gassed FC-40 and mixing by pipette action, prior to experiments. Then the experiments were done under same conditions with size measurements.

Droplet size determination was done by direct application of Mie scattering theory as it was described in elsewhere [59, 114, 14, 149, 30, 92]. A numeric functional form given in Eq. 2.2, was fitted to absorption data by a homemade software written in Python which uses **lmfit** package for non-linear least squares fitting with constraints.

$$A = lN \sum_i w_i Q_{ext_i} \pi r_i^2 \ , \ \sum_i w_i = 1. \tag{2.2}$$

in which A is the absorbance, l is instrument's path length, Q_{ext_i} is the unitless extinction efficiency for each particle radius r_i, N_0 is the particle number density, and w_i are number fractions for each r_i. Details of fitting procedure as well as the relationships between parameters are provided in Sec. B.1.1.

In the case of BTB loaded droplets, for discriminating the background Mie scattering part in attenuation signal Ångström formula [3, 84, 87] was used. Functional form for transmission of light through a bulk of Mie scatterers was given in literature [162, 40, 136, 147, 174]. As it is shown in Sec. B.1.2, we first converted the functional form given in literature, to a comparable form corresponding to the data acquired from UV/Vis absorption measurements. The new functional form describing the absorbance in Eq. 2.3 was then fitted to measured absorbance data by a similarly designed software that we used for droplet size determination. As in the case of size measurements through direct application of Mie scattering theory, details of the fitting procedure and the definitions for fit parameters are provided in Sec. B.1.2.

$$A = l \frac{10^{a_0 + a_1 \log_{10} \lambda + a_2 (\log_{10} \lambda)^2}}{\ln(10)} \tag{2.3}$$

For pH measurements, we first fit the Ångström formula to the acquired absorbance data from droplet samples within the wavelength range of 400-850 nm. The fit results allowed us to identify Mie scattering component in the absorption signal

from the total spectrum. This component in absorbance was used for determination of the droplet sizes by direct application of Mie scattering. Measured absorbance minus the Mie component gives the BTB absorption spectra. Then, BTB absorption component in the signal was used for pH measurements. The ratio of BTB absorptions at 619 nm and 500 nm ($A_{619/500}$) was used to determine pH within droplets. Details of direct application of Mie scattering and Ångström formula is provided in Sec. B.1

2.2.4 Fluorescent Emission Measurements

Fluorescence measurements were conducted with a PTI QuantaMaster 40 Spectrofluorometer, equipped with a 75 W Xenon lamp as its light source and a photomultiplier unit (Hamamatsu R928). Instrument had a working wavelength range of 200 to 900 nm.

Fluorescein in powder form (Life Technologies, F1300) was dissolved in high purity dimethyl sulfoxide (DMSO; Sigma- Aldrich, 41641-250ML) to prepare 1 mM stock solution. Stock Fluorescein solution was then diluted 100,000 times to 100 nM in buffer solution of 20 mM Tris at pH 7.8 and 200 mM NaCl. This solution was used as dispersed phase. The rest of the sample preparation was same as explained in Sec. 2.2.1. Following the emulsification, dense droplet samples diluted 10 times in FC-40.

Fluorescein loaded droplets were excited at their *isosbestic point* 460 nm. Here the term "isosbestic point" refers to a specific point in the absorption spectra of a fluorophore at which absorption efficiency of dye stays same regardless of the ambient pH. Fluorescent emission data were collected within the wavelength range of 475 nm to 650 nm. Background fluorescence data in measurements were acquired by using water (buffer) only in dispersed phase. Fluorescence data from dye loaded droplets were corrected according to this measured background. The ratio of Fluorescein

emissions at 514 nm and 550 nm ($E_{514/550}$) was used to determine pH within droplets. Reasoning behind the excitation of fluorophore at its isosbestic point as well as the choice of emission ratio $E_{514/550}$ are explained in detail in Sec. B.3.

2.2.5 Fluorophore Calibration Measurements

Application of ratio-metric pH measurements require calibration of dye response. In BTB case, six different buffer solutions of pH 13.6, 7.2 , 6.8, 6.3, 5.8, and 0.9 were prepared using NaOH and/or HCl. All solutions had 20 mM Tris (Sigma-Aldrich, T1378) and 200 mM NaCl within them. Stock BTB solution of 10 mM was diluted to 1 μM in each buffer solutions separately. Absorption measurements on these solutions were done using the same instrument given in 2.2.3. Ratio of absorption values $A_{619/500}$ were measured from absorption spectra scans of 400 nm to 850 nm.

In Fluorescein case, seven different buffer solutions of pH 8.0, 7.2, 6.3, 5.4, 4.6, 3.7, and 2.8 were prepared. Similar to BTB case, all solutions had 20 mM Tris (Sigma-Aldrich, T1378) and 200 mM NaCl within them. Stock Fluorescein solution of 1 mM was diluted to 10 nM in each buffer solutions separately. Fluorescence emission measurements on these solutions were done using the same instrument given in 2.2.4. Ratio of emission values $E_{514/550}$ were measured from emission spectra scans of 475 nm to 650 nm.

Data in ratiometric pH measurements were considered to follow a sigmoid shape [67, 31, 8, 12]. Thus the functional form given in Eq. 2.4 was fitted to $A_{619/500}$ data for BTB and $E_{514/550}$ data for Fluorescein.

$$\frac{\varphi(\lambda_1)}{\varphi(\lambda_2)} = \frac{\beta}{1 + 10^{\gamma - pH}} + \alpha \tag{2.4}$$

Here, $\varphi(\lambda)$ is either absorption (for BTB) or emission (for Fluorescein) of dye solution at a specific wavelength λ. Detail explanations on fitting parameters of Eq. 2.4 (α,

β, and γ) are provided in Sec.B.2 and B.3. Calibration data is shown in Fig. B.1 for BTB and in Fig. B.3 for Fluorescein. Sigmoid function fit results are shown in Fig. B.2 for BTB and in Fig. B.4 for Fluorescein.

2.3 Results and Discussion

2.3.1 Characterization of Droplet Sizes

The size distributions of aqueous droplet samples in FC-40 was first measured by DLS. Intensities from different droplet populations were analyzed through a built-in software attached to the instrument. Analysis method and physical specifications of instrument were given in Sec. 2.2.2. Intensity data (corresponding to populations in the droplet system) from four different samples which prepared under same conditions but different times and fits of a lognormal form to these data are shown in Fig. 2.3.

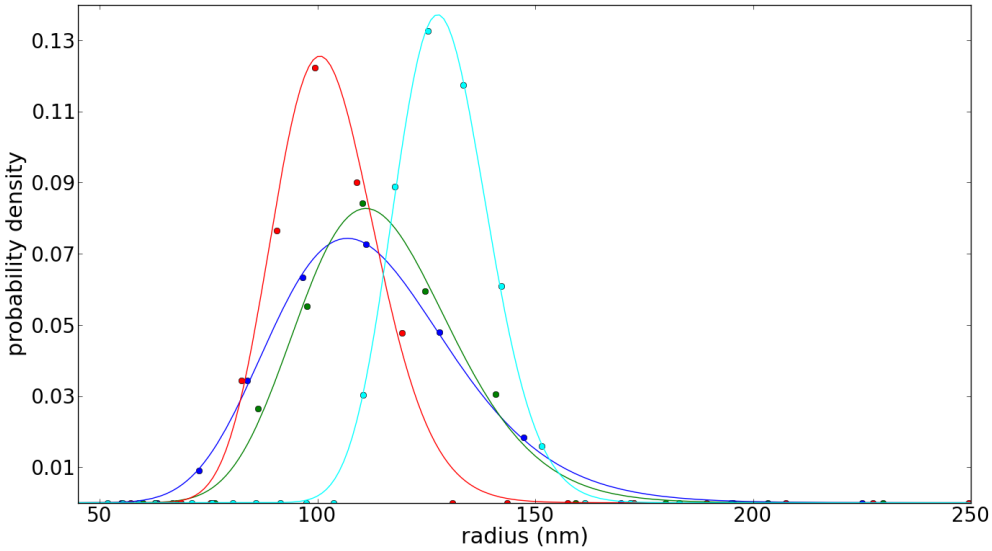

Figure 2.3. DLS results on four different droplet samples as population intensities and their fits to lognormal forms. Color code for the samples are; Sample 1 (blue), Sample 2 (green), Sample 3 (red), and Sample 4 (cyan). Actual data can be found in the place provided in Table. D.1.

Frequently, a population of larger size droplets (on the order of several microns in diameter) were observed in DLS experiments in addition to smaller sizes. Considering the timing in DLS measurements (within first 5 minutes after droplet preparation, Sec. 2.2.2), in the long term they moved to the top of the droplet solution (creamed). Moreover, because of their large sizes, their contribution to Mie scattering was expected to be minimal (Sec. B.1.1). Therefore, the Fig. 2.3 shows only the smaller size droplet populations in experiments.

As it is explained in Sec. B.1.1 a lognormal form can equivalently be described by two independent parameters. These parameters can be any independent couple of location parameter on a logarithmic scale (μ), scale parameter on a logarithmic scale (σ), Mean, Standard deviation (Std), Mode, and Median. Therefore all these parameters and the errors on them are provided in Table 2.1.

	Mean (nm)	Median (nm)	Mode (nm)	Std (nm)	μ	σ
Sample 1	108.8 (0.4)	106.9 (0.3)	103.1 (0.4)	20.8 (1.4)	4.67 (0.0)	0.19 (0.0)
Sample 2	112.8 (0.6)	111.4 (0.6)	108.5 (0.6)	18.2 (2.8)	4.71 (0.0)	0.16 (0.0)
Sample 3	101.2 (0.4)	100.5 (0.3)	99.2 (0.3)	11.9 (2.1)	4.61 (0.0)	0.12 (0.0)
Sample 4	128.0 (0.1)	127.6 (0.1)	126.7 (0.1)	10.7 (1.1)	4.85 (0.0)	0.08 (0.0)

Table 2.1. Distribution parameters with errors (in parenthesis) of DLS measurements in Fig. 2.3. Actual data can be found in the place provided in Table D.1.

Results of these measurements show that our droplet making method, even though it is a bulk emulsification technique so that it is prone to fluctuations, produces similar size droplets at each time. Shot-to-shot droplet size distribution in terms of mean of distribution means and mean of distribution standard deviations (with errors in parenthesis) are 112.7 (9.8) nm and 15.4 (4.2) nm correspondingly.

As an alternative to DLS experiments, we also utilized Mie scattering theory method as explained in Sec. B.1.1, through UV/Vis absorption data. Acquired absorbance data from two different samples prepared in same exact way but at different times, Mie scattering theory results as fits to data, and corresponding droplet size

estimations are shown in Fig. 2.4. Each absorbance data seen in Fig. 2.4, is average of two consecutive measurements of the same sample, so the errorbars for each data set were estimated by using *student's t-distribution* for the specified confidence interval of 90%.

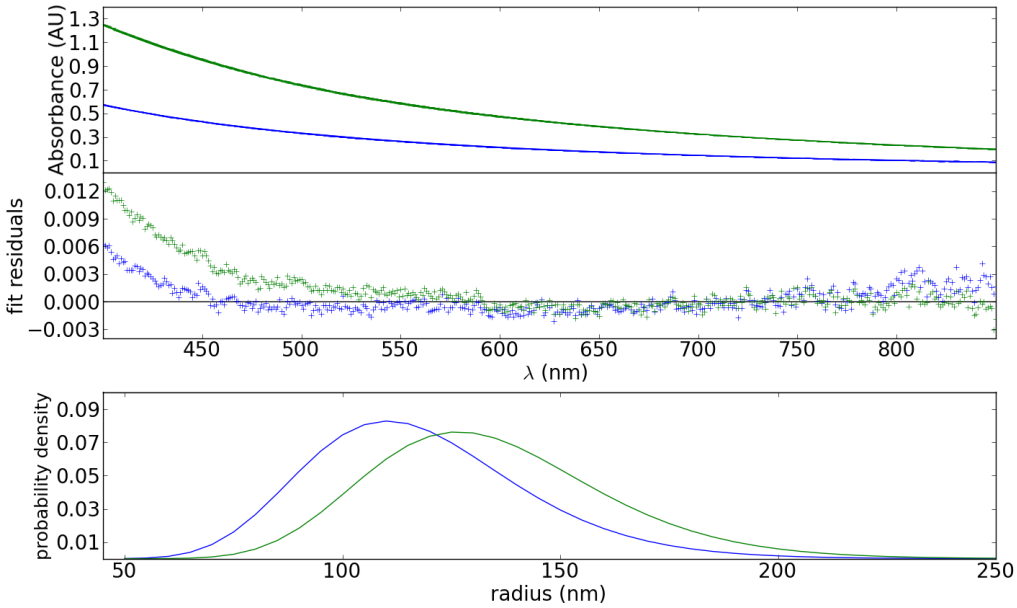

Figure 2.4. Top: Acquired absorbance data and corresponding Mie scattering fits. Bottom: Size measurements found by fitting. Color code for the samples are; Sample 1 (blue), and Sample 2 (green). Actual data can be found in the place provided in Table D.2.

A lognormal form was already inherent to Mie scattering fits. It was used in construction of numeric functional form for the absorption through the Eq. 2.2. Similar to DLS measurements, fitting parameters an the errors on parameters are provided in Table. 2.2. Here the shot-to-shot droplet size distribution in terms of mean of distribution means and mean of distribution standard deviations (with errors in parenthesis) are 126.4 (8.3) nm and 26.0 (1.0) nm correspondingly.

Comparison of size measurements by DLS and Mie scattering in terms of shot-to-shot averaged parameters is shown in Fig. 2.5. Here we can infer several important characteristics of the system from combined results of both measurements. First,

	N(x10^8/cm^3)	Mean (nm)	Median (nm)	Mode (nm)	Std (nm)	μ	σ
Sample 1	0.58 (0.07)	118.1 (9.7)	115.4 (7.4)	110.3 (8.8)	25.5 (29.4)	4.75 (0.06)	0.21 (0.24)
Sample 2	0.67 (0.03)	134.6 (4.8)	131.9 (3.7)	126.6 (4.4)	27.5 (15.4)	4.88 (0.03)	0.20 (0.11)

Table 2.2. Distribution parameters with errors (in parenthesis) of Mie scattering fits on absorption data in Fig. 2.4. Actual data can be found in the place provided in Table D.2.

our aqueous droplets are actually Mie scatterers so that their characteristic radii are on the order of 100 nm. Secondly, droplet making method successfully produced kinetically stable emulsions (no appreciable flocculation, neither creaming) and so the deviation between the direct application Mie theory which strictly requires spherical particles, and the DLS results are minimal. Finally, knowing that Mie scattering works in favor of forward scattering [14], and absorption instrument works in this particular configuration (DLS was done at 90°, see Sec. 2.2.2), size measurements through attenuation experiments are expected to be more sensitive.

Compared to DLS, size measurement through absorbance is slightly less sensitive to droplet concentration in the emulsion. In the case of kinetically stable emulsions, both flocculation and creaming are often the main concerns of stability. Thus, adjustment of droplet concentration in the emulsion might be required to promote stability in these systems. When this happens, despite healthy and reproducible measurements can be very hard to achieve by DLS, attenuation method can turn into a very useful alternative way to measure droplet sizes. Equivalently, attenuation method can be utilized to monitor dynamic changes in droplet size distribution parameters as a function of time, if there is any. Nevertheless, the downside of the method, given the scaling parameter ($x = (2\pi n_{media} r)/\lambda$) in Sec. B.1.1, is its sensitivity range in the sizes. To be more precise, while droplets of sub-micron radii can be modeled well, 1 μm or larger radii droplets are almost invisible to the method.

As a first attempt to monitor time dependent behavior of droplet size distribution parameters, Fig. 2.6 shows the results of size measurements from a single droplet

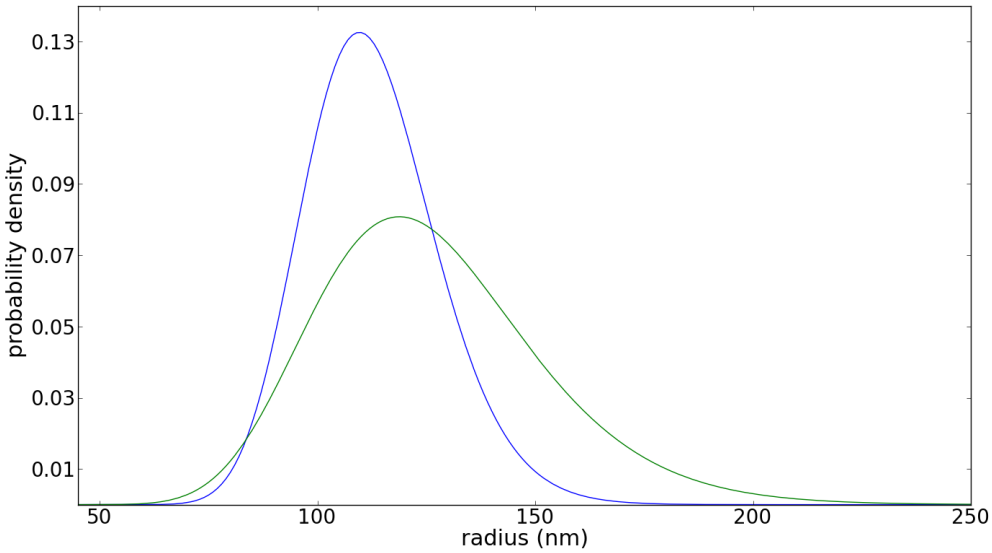

Figure 2.5. Comparison of size measurements using DLS (blue) and Mie scattering (green) in terms of shot-to-shot distributions for which mean and standard deviations were given in the text.

sample using direct Mie scattering theory application through absorbance as a function of time. Each absorbance data seen in Fig. 2.6, is average of two consecutive measurements of the same sample, so the errorbars for each data set were estimated by using *student's t-distribution* for the specified confidence interval of 90%. Relevant parameters for distributions are also given in Table. 2.3.

t (min)	N(x10^8/cm^3)	Mean (nm)	Median (nm)	Mode (nm)	Std (nm)	μ	σ
0	0.58 (0.07)	118.1 (9.7)	115.4 (7.4)	110.3 (8.8)	25.5 (29.4)	4.75 (0.06)	0.21 (0.24)
10	0.23 (0.05)	132.5 (14.8)	128.2 (11.4)	120.0 (12.1)	34.6 (36.6)	4.84 (0.09)	0.26 (0.26)
20	0.25 (0.02)	121.2 (6.4)	115.8 (4.8)	105.9 (5.3)	37.2 (14.2)	4.75 (0.04)	0.30 (0.11)
30	0.18 (0.08)	117.7 (27.3)	111.0 (20.2)	98.55 (21.9)	41.8 (54.5)	4.71 (0.18)	0.34 (0.42)

Table 2.3. Time dependent droplet sizes in FC-40. Actual data can be found in the place provided in Table D.3.

Data in Table. 2.3 show that no appreciable change observed in the droplet size distribution parameters within first 30 minutes. Only clear change was in droplet

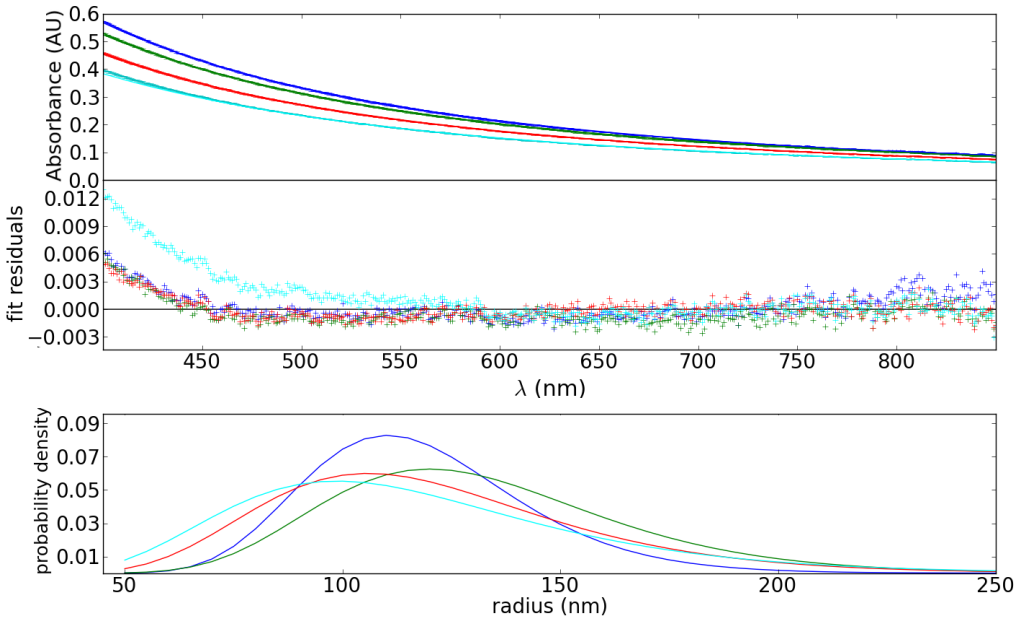

Figure 2.6. Droplet sizes as a function of elapsed time. Color code for the samples are; t=0 min (blue), 10 min (green), 20 min (red), 30 min (cyan). Actual data can be found in the place provided in Table D.3.

concentrations and they decreased in time. This change can be attributed to creaming in the emulsion system.

Lastly, for the completeness of the method, fit parameters of Ångström formula for the data sets in Fig. 2.4 and 2.6 are provided in Tables B.1 and B.2. We should note that these parameters are indirect reporters of the size and concentration and they can not be used directly for characterization.

2.3.2 Characterization of pH within Droplets

pH measurements within droplets were done by absorption and fluorescent emission experiments. In the absorption measurements BTB loaded droplets were prepared as explained in Sec. 2.2.3. Attenuation by Mie scattering and dye absorption components were separated by fitting Ångström formula to the acquired absorption data. Results of this procedure is shown in Fig. 2.7. The absorbance data seen in

Fig. 2.7, is average of three consecutive measurements of the same sample, so the errorbars for each data set were estimated by using *student's t-distribution* for the specified confidence interval of 90%.

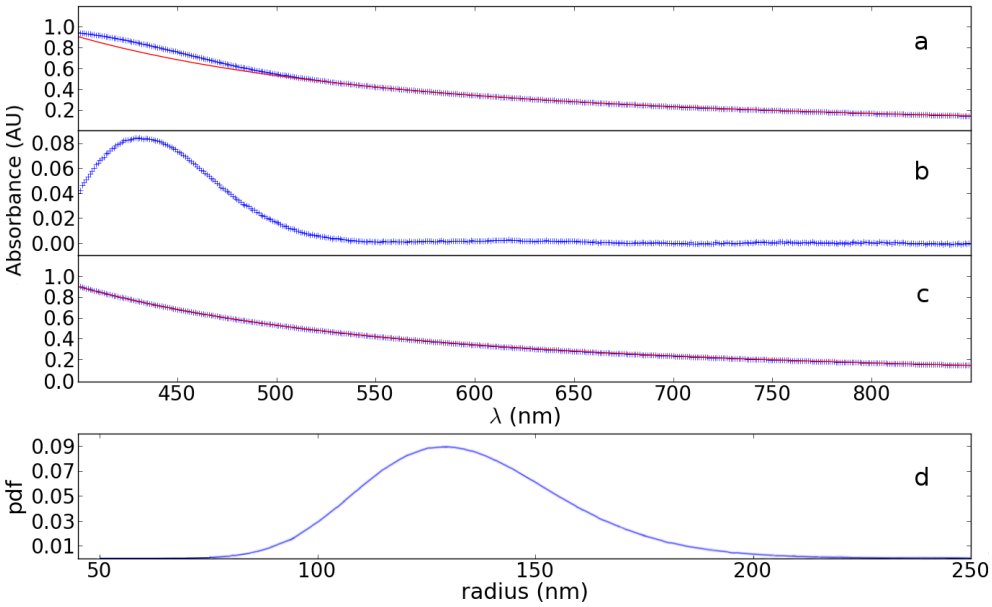

Figure 2.7. Absorption measurement on the BTB loaded droplets. **Panel a:** Total absorbance (blue '+') and fit of Ångström formula (red '-'). **Panel b:** Only BTB absorption after subtracting Mie scattering contribution. **Panel c:** Absorbance by Mie scattering (blue '+') and fit using direct Mie scattering theory (red '-'). **Panel d:** Droplet size distribution found by direct Mie scattering theory. Actual data can be found in the place provided in Table D.4.

Data after removal of Mie scattering contribution, shown in Fig. 2.7 b, gave us an $A_{619/500}$ of 0.14 (0.12). By comparing the $A_{619/500}$ value with the calibration curve provided in Sec. B.2, pH of the droplet interior were found to be 5.1. BTB is not sensitive to pH changes below pH 6 because of its pK_a which is 7.1. Therefore, we can only conclude that the pH within droplets, on the average, below 6. Here, we also calculated the droplet size distribution parameters by direct application of Mie theory on the Mie scattering attenuation part, shown in Fig. 2.7 c. These parameters as in Table. 2.3, 0.55 (0.01), 134.88 (1.00), 132.95 (0.77), 129.18 (0.943), 23.04 (3.78),

4.89 (0.01), 0.17 (0.03) for N, Mean, Median, Mode, Std, μ, and σ correspondingly. Droplet size measurement gave a similar size distribution parameters to the ones we previously measured (Fig. 2.6).

For the emission measurements, Fluorescein loaded droplets were prepared as explained in Sec. 2.2.4. To alter the pH within droplets a strong base (NaOH) was added to dispersed phase prior to droplet making. Acquired fluorescence emission as a function of NaOH concentration is given in Fig. 2.8. It is important note that bulk pH values (after NaOH addition) were not measured prior to droplet making.

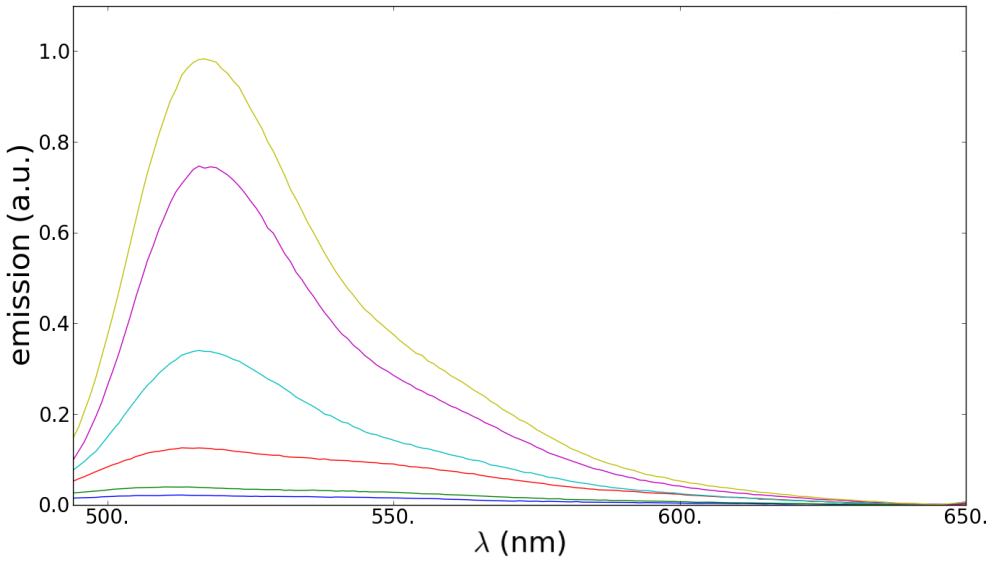

Figure 2.8. Fluorescein emission within droplets by addition of different amounts of strong base NaOH; 0 mM (blue), 1 mM (green), 5 mM (red), 20 mM (cyan), 50 mM (magenta), 100 mM (yellow). Actual data can be found in the place provided in Table D.5.

For each sample, emission ratios $E_{514/550}$ and corresponding pH values were calculated using the calibration curve and parameters given in Sec. B.3. Based on the results reported in Table. 2.4, we observed that even though pH of bulk dispersed phase was 7.8, measured pH within droplets were much lower, 5.3 (0.13). Secondly, measured pH for large amounts of NaOH in bulk Fluorescein solution prior to droplet

making are not correct, because Fluorescein looses its sensitivity around measured values. As a side note, we have also done 50 mM and 100 mM NaOH measurements using BTB. These data can be found in the place provided in Table D.4. Because fitting algorithm needed improvements for the analysis of these data, they will not be reported here. In particular, the residuals seen in Figs. 2.4 and 2.6 for the wavelengths ($\lambda \leq 450$ nm) caused the problems in fitting.

NaOH (mM)	$E_{514/550}$	pH
0	1.38 (0.03)	5.30 (0.13)
1	1.39 (0.03)	5.32 (0.12)
5	1.43 (0.04)	5.38 (0.12)
20	2.35 (0.06)	6.47 (0.13)
50	2.54 (0.06)	6.81 (0.22)
100	2.57 (0.06)	6.88 (0.24)

Table 2.4. pH measurements within droplets prepared in FC-40. Actual data can be found in the place provided in Table D.5.

Given that the surfactant could have unreacted polymers which have carboxylic acid groups, we also tested the effect of surfactant on droplet pH. Our measurements indicated that the removal of surfactant increased the pH by 0.3 and made it 5.60 (0.10). Thus the effect of surfactant was minimal on droplet pH.

CHAPTER 3

SINGLE-MOLECULE STUDIES WITHIN SUB-MICRON SIZE DROPLETS

The work in the current chapter is also documented in [111].

3.1 Background and Motivation

Fluorescence resonance energy transfer (FRET) from dye-labeled molecules is widely used in molecular biophysics to understand folding, binding, and structural changes in proteins [119] and RNA. [96] The simplest and most frequently used single-fluorophore-sensitive FRET measurements involve molecules that are freely-diffusing in solution. In this case, a burst of fluorescent photons are recorded as a molecule crosses the detection volume of a confocal microscope. The signal-to-noise ratio in these experiments is determined by the brightness of the molecule, the background in the detection channel(s), and length of time spent in the detection volume. A larger detection volume increases the detection time at the expense of higher background. To reduce background, femtoliter detection volumes, corresponding roughly to the focal volume of an oil- or water-immersion high numerical aperture lens, are often used. For typical biomolecules in aqueous buffer this results in diffusion-limited detection-region dwell times of < 1 ms.

For spherical particles or attodroplets, the diffusivity is given by the Stokes-Einstein equation as $D = k_B T/(6\pi\eta r)$, where r is the hydrodynamic radius of the particle, k_B is the Boltzmann constant, T is the temperature and η is the dynamic viscosity. For Brownian motion, the relevant diffusion time τ will be proportional to

w^2/D where w is the waist diameter of the confocal detection volume. The diffusion time therefore scales with the radius of the particle and the viscosity of the medium.

Characterization of the droplets used in this study in terms of sizes were done both by absorption/attenuation and DLS measurements. As it can be seen in Fig. 2.5, both type of measurements produced similar results (in fact same within the errors) consistent with a lognormal distribution of droplet radii. Distribution parameters found to be $\langle r \rangle = 112.7$ nm corresponding to $\mu = 4.72$ and $\sigma = 0.14$ in DLS measurements, whereas absorption/attenuation measurements showed $\langle r \rangle = 126.4$ nm corresponding to $\mu = 4.82$ and $\sigma = 0.20$ (Fig. 2.5). As it was explained in Sec. 2.3.1 the mean radii were the mean values of measured distribution means for DLS and attenuation measurements. Fluorescence correlation spectroscopy (FCS) has the potential to serve as a third method for measuring droplet sizes. The size information of droplets can be found by the analysis of autcorrelation function from an FCS measurement. On the other and ,in our knowledge, there is no proposed analytical form for autocorrelation function (ACF) in literature. Therefore, FCS method requires simulation of the droplet system which consists. It is also important to note that simulations can provide information about the environment within droplets, in addition to droplet size determination.

Regardless of the size measurement method, most of the droplets in current work (95%) have a radius between 80 and 175 nm. Corresponding volumes are 2.1 aL - 22 aL with a most frequently observed value of 7.1 aL. We study FRET from doubly-labeled duplex RNA confined to these droplets at a nominal concentration of 10 nM; 7.1 aL droplets therefore contain 0.04 RNA molecules on average, while a droplet of radius 175 nm has an average of 0.13 molecules within. While the vast majority of droplets are empty, confinement is 100% efficient, meaning that the hydrophilic RNA molecules do not partition into the continuous (perfluorinated) phase. Only droplets containing molecules are observed in fluorescence measurements.

Droplet confinement has several advantages for single-molecule-sensitive measurement: in addition to being efficient, samples are relatively easy to prepare. The droplet interface is far more homogeneous than that provided by, for example, a glass substrate. However, of all the methods previously used to confine and isolate molecules for individual study, droplets are unique because there is no particle reservoir to draw on: each droplet is stochastically unique in its contents, and there is no chemical equilibrium between droplets. In fact this behavior is promoted in our studies by using a surfactant as explained in Sec. 2.2.1. This can be advantageous for high-throughput screening of disparate species or dynamics, but only if the environment in the droplet is well understood. For attoliter volume droplets, as we have shown in Chapter. 2, the pH conditions will be different than the initial bulk solution. There is also no guarantee that both pH and/or salt conditions will be the same even between droplets.

In all studies presented here, NaCl is present at 200 mM so that even the smallest droplets contain 10^5 salt molecules on average. The Debye length at 200 mM salt is approximately 0.7 nm, similar to the Bjerrum length in water. NaCl is expected to form double layers at an oil interface that give roughly the same overall density as in the bulk (no interfacial excess), [172] so we believe the salt concentration is probably not significantly altered by the droplet environment.

It is important to note that biomolecules confined in these droplets show no evidence of sticking at the perfluorinated walls, whether or not surfactant is in use. This was demonstrated explicitly for green fluorescence protein [163] and also for nucleic acids [81] using polarization anisotropy lifetime measurements in sub-femtoliter droplets. Confocal images of large (micron size) droplets also show no evidence of biomolecules sequestered at the water/perfluorinate interface (Fig. C.1).

3.2 Experimental Methods

3.2.1 Sample Preparation

The RNA 16 base oligo 5′-Cy3-C-G-A-G-U-G-A-C-C-A-G-U-G-A-G-C-3′ and its complement with and without a Cy5 at the 5′ terminus, were obtained from IDT. Cy3 and Cy5 are indocarbocyanine dyes supplied by Glen Research. Donor (Cy3) and acceptor (Cy5) labeled ribonucleotides were prepared in 20 mM Tris at pH 7.8 with 200 mM NaCl. In this buffer, RNA at 0.75 μM was heated to 90° C in 5 minutes and then incubated at 90° C for 60 minutes before cooling to 4° C over 60 minutes. For measurements on freely diffusing molecules, 15 nM protocatechuate-3,4-dioxygenase (PCD) and 5 mM protocatechuic acid (PCA) were mixed in 20 mM Tris with 200 mM NaCl and incubated for 10 minutes. PCA/PCD functions as an enzymatic oxygen quenching system.[1] The dsRNA samples were diluted in this buffer to a concentration of 100 pM or 200 pM with 1 mM methylviologen (MV).

3.2.2 Droplet Preparation

The dsRNA sample was prepared in emulsion as follows: 2 μL of dsRNA at 10 nM or 20 nM with 10 nM PCA, 100 nM PCD and 1 mM MV was added to a 200 μL of a continuous phase consisting of degassed perfluorinated oil (3M Fluorinert FC-40 or FC-77) with 10^{-3} v/v perfluorinated surfactant (RainDance).[72] After shaking, the mixture was sonicated for 2 minutes in an ultrasonic cleaner (Branson 1510), which formed the emulsion. FC-77 is primarily 2-(nonafluorobutyl)heptafluorofuran, with average molecular mass of 415, viscosity of 1.36 cP and refractive index of 1.27. FC-40 is primarily perfluorotributylamine, with average molecular mass 650, viscosity of 4.1 cP, and refractive index of 1.29. Note that in both cases the refractive index is near but lower than that of water ($n = 1.33$).

Approximately 50 μL of emulsion was withdrawn and placed between a coverslip and microscope slide separated by double-sided sticky tape, which was then sealed with silicone vacuum grease.

De-oxygenation of the perfluorinated oils was achieved by the freeze-pump-thaw method. Perfluorinated oils were placed in a sealed Schlenk flask and frozen in liquid nitrogen. The flask was then opened to vacuum and pumped to 12 mtorr, re-sealed and thawed in a warm water bath. After sitting for 30 minutes the process was repeated up to five times.

3.2.3 Photon-Counting Histogram Analysis

A photon-counting histogram (PCH) is a histogram of the number of photons per bin during a photon-counting experiment. The bin time is typically short; here we use 200 μs. Analysis of the PCH gives the average number of photons per bin and average number of molecules in the detection volume for multiple species. Species are distinguished only by their brightness, not by their diffusivity as in fluorescence fluctuation correlation spectroscopy (FCS). PCH is often used as a complement to FCS in the analysis of photon statistics in single-molecule-sensitive measurement. Analysis of PCHs was developed by Chen *et al.* [21] and later updated by Huang *et al.* [76]. Here we follow the method and nomenclature of the latter. This model assumes a three dimensional Gaussian detection volume, with two correction parameters that describe deviations from Gaussian. Fitting parameters therefore include the brightness ϵ and molecular concentration $\langle n \rangle$ for each species as well as beam-shape correction factors F_1 and F_2. More details and fit results are given in the Sec. C.3.

3.2.4 Burst Detection

Burst detection was accomplished using a simple Bayesian method based on photon inter-arrival times. All the photons (both channels) are used, and the method distinguishes between photons from fluorescent bursts and photons from background.

To determine if the ith photon originates from a burst, the arrival times of N photons on either side of the ith photon were examined. Here we use a "window" with $N = 5$ photons. Further details are provided in the Sec. C.2.

3.2.5 FRET Measurements

In fluorescence resonance energy transfer (FRET), an excited donor dye transfers its energy to a redder acceptor dye if the molecules are sufficiently close:

$$E = \frac{1}{1 + \left(\frac{R}{R_F}\right)^6}, \tag{3.1}$$

where E is the energy transfer efficiency, R is the distance between dyes and the Förster radius R_F is given by [25, 121]

$${R_F}^6 = \frac{9c^4 J \eta_D \kappa^2}{8\pi n^4}. \tag{3.2}$$

In this expression, n is the solvent's refractive index, c is the speed of light, η_D is the quantum yield of the donor dye in the absence of the acceptor, and κ is a factor that describes the relative orientation of the dyes; $\langle \kappa^2 \rangle = 2/3$ for freely rotating dyes. The symbol J describes the overlap of the donor emission and acceptor absorption spectra.[121] With the assumption of freely rotating dyes, FRET can be used to measure distance between disparate points in a molecule; more frequently it is used to qualitatively observe global changes in molecular structure or binding.

FRET was measured using a ratiometric technique. An Olympus IX50 microscope was modified for single-molecule confocal detection with a 50 μm pinhole. A UPlanSApo 60×, 1.2 NA water immersion objective was used for both fluorescence excitation and collection of emitted photons. The donor dye was excited with the 514 nm line of an Argon-Krypton laser with a nominal power (measured at the entrance to the scope) of 20 μW or 50 μW. Fluorescent photons were split into two

channels (donor, acceptor) and detected using single photon counting avalanche photodiodes (τ-SPAD by Picoquant). Photon timing information was recorded with 8 ns resolution using homemade instrumentation. [53] We make no attempt here to correct for background, crosstalk, or γ, the parameter that describes the relative quantum yields and collection efficiencies of the dyes and instrument. [56] Rather, we report on the proximity ratio

$$P = \frac{N_A}{N_A + N_D} \tag{3.3}$$

where N_A and N_D are the number of photons in the acceptor and donor channels, respectively, in a given time interval. The proximity ratio $P = \langle E \rangle$ only in the ideal case where $\gamma = 1$ and there is no cross-talk, direct acceptor excitation or background. [56, 25] Defining $N_t = N_A + N_D$, proximity histograms are formed using all bins containing $N_t > N_{th}$. Here $N_{th} = 25$, 50, or 75 photons. The shot-noise limited variance of a peak in this distribution is given by [56]

$$\sigma_s{}^2 = \frac{\langle P \rangle \left(1 - \langle P \rangle\right)}{\langle N_t \rangle} \tag{3.4}$$

3.2.6 Fluorescence Correlation Spectroscopy

FCS data were acquired on donor-only labeled RNA under the same conditions used for FRET. For a single diffusive dye with no kinetic terms due, *e.g.* to a triplet or isomer, the ACF can be approximated as:[165, 132, 141]

$$G_D(\tau) = A \; \frac{1}{1 + \frac{\tau}{\tau_D}} \; \frac{1}{(1 + \frac{1}{\omega^2}\frac{\tau}{\tau_D})^{\frac{1}{2}}}, \tag{3.5}$$

where τ_D is the diffusion time, $A = 1/M$ where M is the average number of molecules in the detection region, and $\omega = 10.2$ is a factor that describes the ellipticity of the detection volume and which was separately measured for our instrument. If a

single kinetic term on a timescale faster than the diffusion time is present, this G_D is modified by a factor that is given by:[178, 179]

$$g(\tau) = \frac{1}{1-F}\left(1 - F + F\exp(\frac{-\tau}{\tau_k})\right) \tag{3.6}$$

where F is the fractional amplitude of the kinetic term and τ_k is the timescale of the kinetic process. For data with multiple species with two distinct diffusion times labeled with the same dye, data are fit using some or all of the terms in the following expression:[153]

$$G(\tau) = \left(\sum_{j=1,2} G_{D_j}(\tau_j)\right) \prod_{i=T,I} \frac{1}{1-F_i}\left(1 - F_i + F_i\exp(\frac{-\tau}{\tau_i})\right). \tag{3.7}$$

Here I and T refer to isomer and triplet timescales, respectively.

Here, specific to FCS within diffusing droplets, there is no analytical function for describing the acquired ACF (Sec. C.5). This led us to use approximate functional forms for fitting the experimental data. One these forms was two component model corresponding to two translational correlation times as shown in Eq. 3.7. The other functional form was a single translational diffusion model with a chemical reaction kind of term in front in which the scaling of the time was a stretched exponential type. This functional form is shown in Eq. 3.8. Slightly more detailed explanation about the use of this functional form is provided in Sec. C.5.

$$\begin{aligned} G(\tau) &= A\{1 + B\exp[-(\frac{\tau}{\tau_B})^{\alpha}]\} \times \frac{1}{1+\frac{\tau}{\tau_D}} \frac{1}{(1+\frac{1}{\omega^2}\frac{\tau}{\tau_D})^{\frac{1}{2}}} \\ &\times \prod_{i=T,I} \frac{1}{1-F_i}\left(1 - F_i + F_i\exp(\frac{-\tau}{\tau_i})\right) \end{aligned} \tag{3.8}$$

In a typical FCS experiment, we recorded arrival times for each photon released by fluorescently labeled molecules. ACF from recorded photon stream is calculated using the equation:

$$G(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle\langle I(t)\rangle} \tag{3.9}$$

Here I denotes fluorescent intensity.

To extract particle radii from diffusion time using the Stokes-Einstein relation, the diameter of the detection volume was measured using tetramethylrhodamine (TMR) and found to be approximately 260 nm. Similarly aspect ratio of the detection volume was measured as 10.2 using TMR.

3.3 Simulation Methods

Core part of our simulation code was a Monte-Carlo algorithm designed to generate spatial positions of molecules trapped within droplets. A software clock generates time stamps of 100 μs width (Δt). At $t = 0$ a droplet is generated within the simulation box of sizes $L_x = L_y = 7800$ nm and $L_z = 53040$ nm. The detection volume width and aspect ratio values were found as explained in Sec. 3.2.6.

System boundaries in diffusion simulations usually designed to be periodic so that when a particle leaves the simulation box it immediately re-enters the box from the opposite end [36, 32]. Use of a periodic box reduces the computational cost and speeds up the code. But, artificial periodicities in spatial coordinates can also introduce artefacts to simulations [2]. Because we wanted to minimize artificial periodicities and we had an already fast enough simulation code, we didn't apply periodic boundary conditions in our simulations. Instead, simulation code works in such a way that when a droplet leaves the simulation box, another one was created in a randomly chosen position within the box. The issue with this approach was the possibility of creating a droplet within the excitation volume. To overcome this problem, we used

a large enough box which is about 18,000 larger than the excitation volume. In other words, the probability of having a droplet created in the excitation volume in this simulation scheme is less than 10^{-4}.

In addition, by the fit results of droplet data given in Table 2.2, typical droplet number concentration in our experiments was roughly 4×10^7 $1/\text{cm}^3$. On top of that, the probability of having an observable droplet meaning that a droplet with at least one fluorophore in it, was 2×10^{-2} (Sec. 3.1) for most frequently observed droplet sizes. Thus, the average droplet concentration in simulation box is expected to be roughly 2.4×10^{-2}. In other words having on the average, one observable droplet in simulation box was not unphysical.

At each droplet creation, the size of the droplet was assigned based on the measured lognormal distribution (either from DLS, or absorption/attenuation measurement results) of sizes. The assigned droplet size as well as the material properties of continuous phase oil were used for finding the translational diffusivity of the droplet. This was followed by assigning the number of fluorescent molecules within the given size droplet based on the bulk solution concentration in experiments. This was done by drawing a random Poisson variate with the mean calculated by the mean number of particles per given droplet size. The translational diffusivity for the molecules within droplets were assumed to be same as the diffusivity found from bulk experiments as explained in Sec. 3.2.6.

At each clock tick, positions of both droplet and the molecule within droplet were updated. Step lengths were chosen from one Gaussian and two uniform random variates corresponding to radial, polar and azimuthal directions. Mean of the Gaussian random variates were set to be 0 and standard deviations were assigned using the root-mean-squared displacement found from corresponding diffusivities and software clock increment Δt. Interactions between droplet boundary and the molecules within droplets were assumed to be perfectly elastic collisions.

Excitation laser beam was supposed to be Gaussian as in Eq. C.5. At each clock tick, based on the position of fluorescent molecule(s), excitation intensity was calculated and compared with a uniform random variate to specify whether the molecule(s) was excited or not. If the molecule was excited, release of a photon was decided by drawing a Poisson variate of mean photon release rate 10^6 photons/second and Δt. Here we tested a range of photon release rates from 5×10^4 to 2×10^6 and we observed no change in ACF (data not shown).

Each simulation ran until collecting 2 million photons. Timestamps were recorded as 32-bit integers with a wrapper at 10^7 clock ticks. Because the largest 32-bit integer is supposed to be around 2×10^9, we expected no overflows in the code. Implementation of random variates were done by using Gnu Scientific Library (GSL) and the code was written in C. Rotational diffusion was not implemented in the simulation code neither the photo-physical properties of dyes were taken into account. Similar to the analysis of acquired photon records in experimental data, ACFs from photon timestamps were calculated by using a homemade software [53]. Fitting of analytical ACFs to the data were done by another homemade software written in Python using the *lmfit* package which was designed for non-linear fitting problems with constraints. Eq. 3.7 using only two components (corresponding to two diffusing species with different diffusivities) without chemical reaction terms for dye photo-physics (triplet formation and isomerization) was used as approximate analytical form of the ACF.

3.4 Results

Raw data from freely diffusing and droplet confined 16 base-pair RNA duplexes labeled with Cy3 (donor) and Cy5 (acceptor) at their $5'$ termini are shown in Fig. 3.1. Photons in the donor and acceptor channels are binned in 5 ms intervals and plotted in green and red, respectively. The donor channel is plotted upside-down, with the number of photons per bin on the right axis, for clarity. The panels on the left are 30

second portions of data files that are between 25 and 75 minutes long. The smaller panels on the right are an 0.5 s expansion of the data colored black in the left panel. The peaks in Fig. 3.1(a) correspond to freely diffusing molecules crossing the detection volume. In this case, the diffusion time across the volume is much less than 5 ms, so the peaks typically consist of only one, or at most two, above-background bins. For molecules confined to aqueous droplets in FC-77, Fig. 3.1(b), or FC-40, Fig. 3.1(c), the diffusion time is clearly longer and more photons are collected.

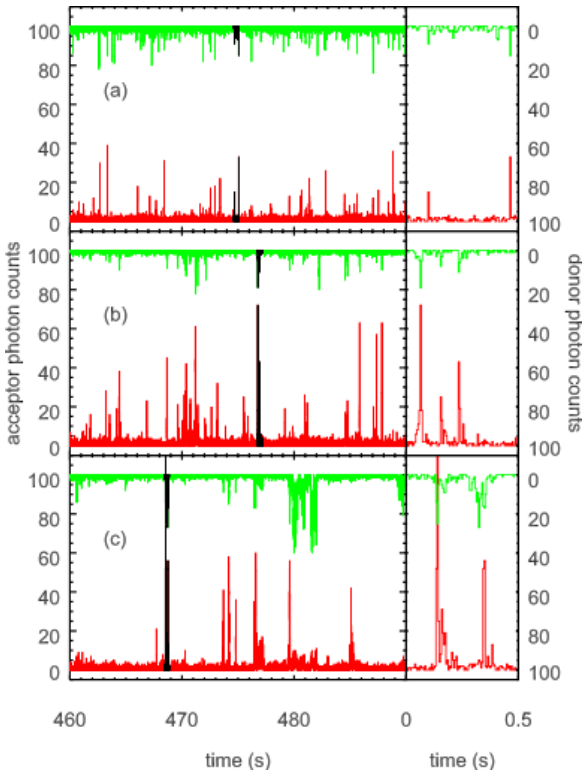

Figure 3.1. A comparison of (a) fluorescence from doubly-labeled freely diffusing RNA molecules with (b) FRET from doubly-labeled RNA confined to droplets diffusing in FC-77 and (c) FRET from doubly-labeled RNA confined to droplets diffusing in FC-40. In each case, the donor-channel is plotted upside-down, with the associated axis label on the right.

A comparison of the photon statistics for molecules confined to droplets with those diffusing freely in solution is given in the photon-counting histograms[21, 76] (PCHs) of Fig. 3.2. Here the bin size was decreased to be 200 μs. All data in Fig. 3.2 were

taken on the same day under identical conditions with an excitation power of 50 μW; only samples were changed. For freely diffusing molecules at pH 7 and above, the data are fit well by two species (two diffusing species with different diffusivities), one with roughly twice the brightness and less than 10% the population of the other. Even with our relatively large choice of bin size, these two species might be associated with different isomers of Cy3 evident in FCS, below. The situation is similar at pH 7.8.

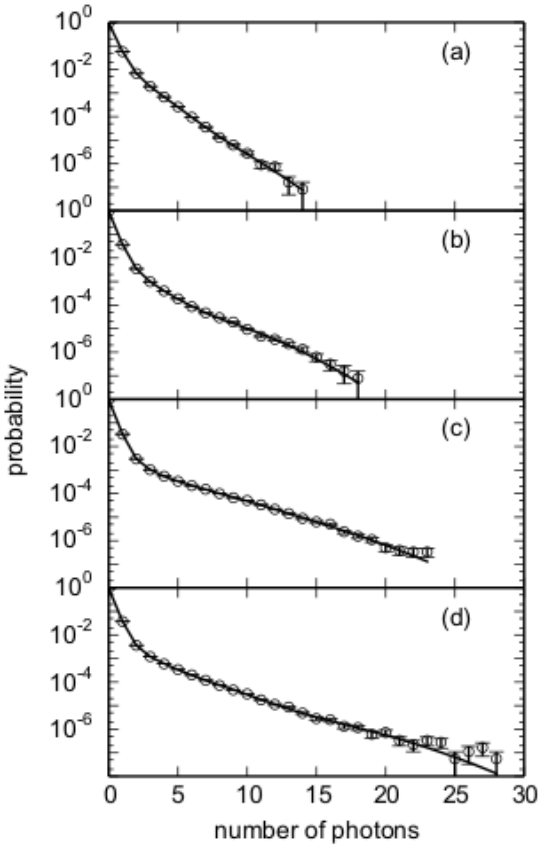

Figure 3.2. Comparison of photon-counting histograms (PCH) of donor-only-labeled RNA in solution with PCH of the same RNA confined to droplets. In all cases, the bin size is 200 μs. (a) RNA freely-diffusing in solution at pH 7 and a two species fit with $\chi^2 = 0.7$; (b) RNA freely-diffusing in solution at pH 4 and a three species fit with $\chi^2 = 1.3$; (c) RNA confined to droplets diffusing in FC-77 and a three species fit with $\chi^2 = 1.1$; and (d) RNA confined to droplets diffusing in FC-40 and a three species fit with $\chi^2 = 1.4$. Here χ^2 is calculated per degree of freedom in the fit. Complete fit results and the resultant fit parameter values can be found in Table C.1.

Below pH 7 a brighter species emerges and three populations are required for a good fit to the PCH. In solution at pH 4 Fig. 3.2(b), this new species is roughly eight times brighter than the dimmest species and comprises roughly 4% of the population (Table C.1). For droplets the situation is similar, Fig. 3.2(c) and (d); the new species is 5 to 6 times brighter than the dimmest species, and comprises at most 5.5% of the population. Differences between FC-77 and FC-40 are mostly insignificant; a complete set of PCH fitting parameters with uncertainties is given in the Sec. C.3. For nearly all species, the droplet data have higher brightness than the solution data; this could be a result of the longer dwell time of droplets in the detection volume and our choice of bin time.

While it would not explain the bright species in solution at low pH, one possible source of an apparent brighter species in droplets might be that some of the droplets contain two RNA molecules. From stoichiometry alone, less than 1% of the very largest droplets should contain more than one molecule. An analysis of the number of photons per burst appeared to confirm that none of the data sets discussed here had droplets containing more than one molecule. For example, eliminating data from the longest bursts, which on average should correspond to the largest droplets, does not significantly affect the data analysis. There is however a weak correlation between the approximate rate in a burst and its length, with a Pearson's coefficient of 0.279 for FC-40 and 0.134 for FC-77. This might be explained if we assume that droplets that take longer to traverse the detection volume also tend to travel through the center of the volume, where the excitation and collection probabilities are higher. The increase in brightness at low pH and in droplets is also noticeable in the burst analysis: The average number of photons per burst for freely-diffusing molecules at pH 7 is 6.9, and at pH 4 is 10.3. For molecules confined to droplets diffusing in FC-77 the average number of photons per burst increases to 36.8. Since FC-40 is the most viscous oil used, bursts last longer yet and contain an average of 51.8 photons. Note that while

no threshold has been applied to the burst detection algorithm, there is a dependence of burst size on the choice of window in the algorithm: notably, larger windows tend to result in longer and larger bursts. With that in mind, it is worth noting that the average length of a burst is virtually unchanged between pH 4 and pH 7, from 1.03 to 1.06 ms. For diffusing droplets, the average burst lasts 3.7 ms in FC-77 and 4.3 ms in FC-40. However, diffusion time through the detection volume is better discussed in the context of FCS, below.

Proximity ratio histograms for single RNA molecules confined to attodroplets in FC-40 are shown in Fig. 3.3. The solid lines are the result of a best fit to three beta probability distribution functions (PDFs); only two peaks are shown. The third peak, corresponding to donor-only signal at a proximity ratio of about 0.15, has been removed (Sec. C.4). The three panels in Fig. 3.3 correspond to different thresholds, with $N_{th} > 25$ photons required in (a), $N_{th} > 50$ in (b), and $N_{th} > 75$ in (c). The resulting average number of photons per bin was greater than 55 for panel (a), approximately 90 in panel (b), and greater than 115 in panel (c) for each of the two peaks. The total number of bins under both peaks is 2800 for (a), 1160 for (b) and 620 for (c). As expected, the width of peaks decrease as the average count rate increases, and for the peak at the lower proximity ratio, this width is 0.059 in panel (c), only slightly larger than shot-noise limited value of 0.041 given by Eq. 3.4. Proximity histograms for attodroplets in FC-77 are similar, but with a more obvious splitting evident between the two peaks (Fig 3.5 and C.2). In all cases, the correlation between burst length and proximity ratio is negligible, with Pearson's coefficients falling in a range between -0.1 and +0.1.

A proximity histogram for freely-diffusing RNA in solution at pH 7.0, taken under conditions identical to those of the droplet-confined RNA, are shown in Fig. 3.4(a). Data at pH 7.8 are similar. This histogram is substantially different from that observed in droplets, both in mean value and in shape of the distribution. Here the bin

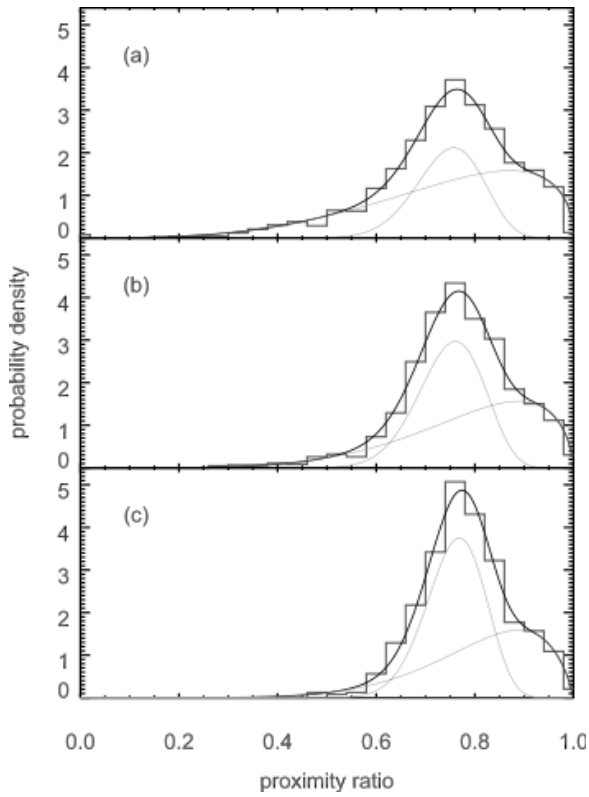

Figure 3.3. Proximity histograms from single RNA molecules confined to freely-diffusing aqueous droplets in FC-40. Photon bin time is 2 ms. The three panels represent the same data but with different thresholds for inclusion in the histogram: (a) $N_{th} > 25$, (b) $N_{th} > 50$, and (c)$N_{th} > 75$. The data are fit with beta PDFs; fit parameters are given in Table C.3.

time is again 2 ms, the threshold is fixed at $N_{th} = 25$, the resulting average number of photons per bin is 29. As above, the data were fit with three beta PDFs and the donor-only peak near $\langle P \rangle = 0.15$ was removed. Efforts to use a single beta PDF or Gaussian to fit the FRET peak resulted in distinctly poorer fits. There are more than 600 photons under each FRET peak in Fig. 3.4(a), but it is not possible to substantially increase the threshold from $N_{th} = 25$; there are only eight bins with $N_{th} \geq 50$.

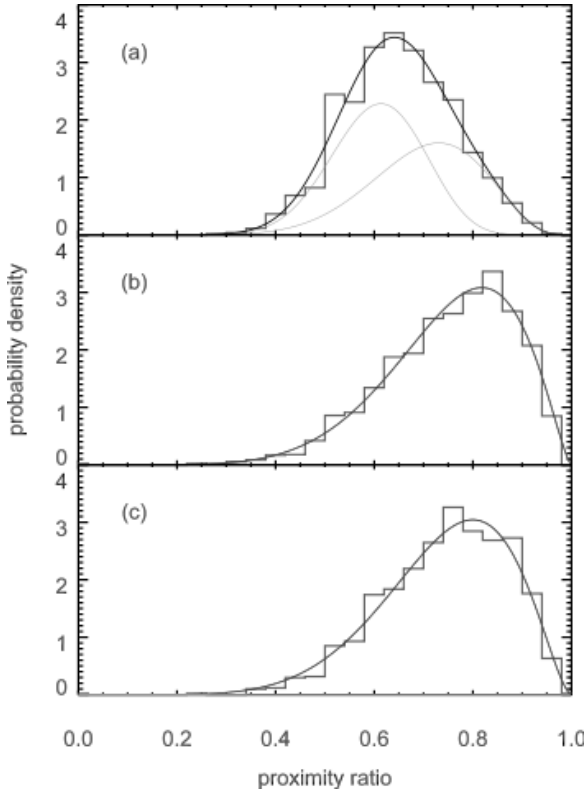

Figure 3.4. Proximity histograms from freely diffusing RNA at (a) pH 7.0, (b) pH 6 and (c) pH 4. Photon bin time is 2 ms, and the threshold for inclusion in the histogram is set at $N_{th} > 25$ in all three cases. The data are fit with beta PDFs, fit parameters are given in Table C.4.

In comparing Fig. 3.3 and Fig. 3.4(a), the most immediate and obvious change is a shift in the average value of the proximity ratio. The cause of this shift was not immediately apparent. The index difference between FC-40 and water is small, only

0.04, and would not be expected to significantly affect the dye lifetime or R_F. We have no evidence that the crosstalk or γ changes in the attodroplet samples, although the background is somewhat smaller for attodroplets because they are more dilute than molecules and the perfluorinates seem to contribute less background than water in any case. FRET for this system was also insensitive to salt concentration between 100 mM and 800 mM (data not shown).

On the other hand, a decrease in pH causes a shift in FRET that is demonstrated in Figs. 3.4(b) and (c), which are taken at pH 6 and pH 4, respectively. Similar data were acquired at pH 3 and 5; between pH 3 and pH 6 the proximity data in solution are unchanged.

The other obvious difference between Fig. 3.3 and Fig. 3.4 are the shapes of the distribution. A direct comparison of low pH data in droplets and solution is given in Fig. 3.5. The low pH solution data falls in a similar range as the droplet data, but the attodroplet distribution shows a clear heterogeneity that the solution data does not. However, the broadness of the low-pH solution distribution would seem to hint at dynamics on a time scale longer than the interphoton-arrival time but shorter than the bin time. Unfortunately, efforts to substantially decrease the bin time of the solution data also decrease the average photon number, which just broadens the distribution again. Efforts to narrow the distribution by increasing N_{th} also failed: for some of the low pH data in solution, the increase in brightness made it possible to increase the threshold in N_{th} to 50 instead of 25. This gave a value for $\langle N_t \rangle$ that was similar to that of the droplet data with a threshold of 25, leaving 290 bins under the FRET peak. No significant change in the shape of the FRET peak was evident. It was not possible to resolve distinct peaks in the solution data at low pH.

FCS data were acquired to address this point by investigating the underlying photo-physical origin of the broadening in the proximity histogram. These measurements also served to directly demonstrate the improved dynamic range that droplet-

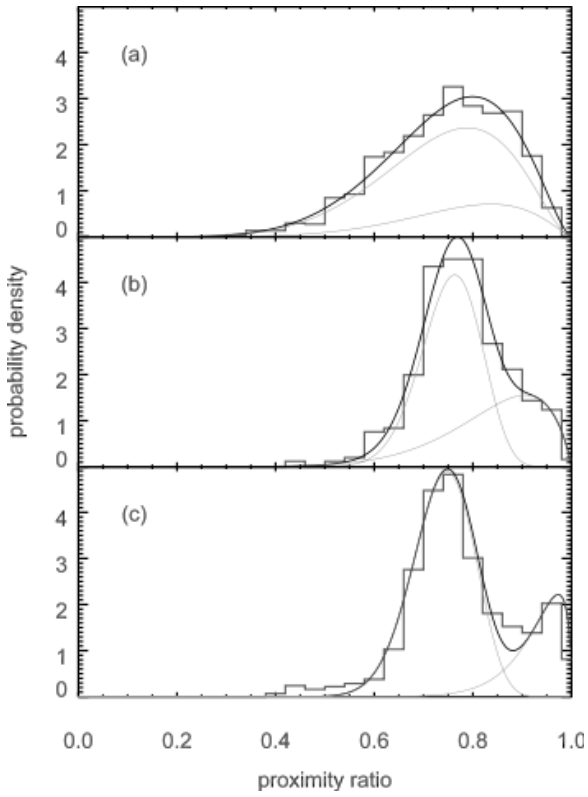

Figure 3.5. Proximity histograms from (a) freely diffusing RNA at pH 4, (b)droplet-confined RNA with FC-40 as the continuous phase and (c) with FC-77 as the continuous phase. Fitting parameters are given in Table C.5. In all three cases, two beta PDFs are used to fit the FRET peak(s), although in (a) a single beta PDF is sufficient for a good fit, as demonstrated in Fig. 3.4. Photon bin time is 2 ms, and the threshold for inclusion in the histogram is set at (a) $N_{th} > 25$ and (b, c) $N_{th} > 75$.

confinement affords. FCS of RNA in attodroplets and RNA in solution are shown in Fig. 3.6(a), with residuals for each fit in panels (b)-(f). For the purposes of FCS, the RNA is labeled only with the donor. Attodroplets are denoted by diamonds in FC-77 and by triangles in FC-40. Solution data at pH 7 is denoted by squares, pH 6 by '×', and pH 3 by '+'. ACFs at pH 4 and 5 fall between those at pH 3 and 6, with an amplitude that increases with decreasing pH. The fit parameters for the data of Fig. 3.6 are given in Table 1. In solution at pH 7, the FCS behaves as expected for Cy3 on RNA,[177] with a single diffusion time of 224 μs, a fast triplet (2 μs) and slower isomer (50 μs). The ACF at pH 7.8 is similar. If the pH is lowered to 6, the ACF changes dramatically. Kinetic terms still describe a triplet and an isomer. However, the data are best fit by a longer diffusive term. This might be naively attributed to large aggregates forming at low pH, but this seems unlikely given that the concentration of RNA here is 200 pM. Furthermore, there is no evidence for two diffusive species (*e.g.* a monomer and an aggregate) from the FCS data, while the PCH indicates an additional bright species at low pH that occupies only a small percent of the population. Rather, we suspect that this apparent increase in diffusion time is an artifact caused by a kinetic term for transformations between the bright and dim species that occurs on a time scale too long to be modeled using Eq. 3.6; that is, the correlation time for conversion is of the order or somewhat longer than the diffusion time. In this case there is no good analytical model for the FCS function in terms of the dynamics, Eq. 3.7 in particular is not valid, and it would perhaps not be unusual for dynamical behavior to mimic a longer diffusion time. Such an interconversion might also explain the very broad low pH solution FRET data. We are working towards a model for FCS to test this assertion.

The situation in droplets is quite different. Despite there is no known analytical functional form ACF in droplets, by fitting to a two component model (Sec. C.5) we observed two distinct diffusive timescales; attempts to fit with multiple kinetic

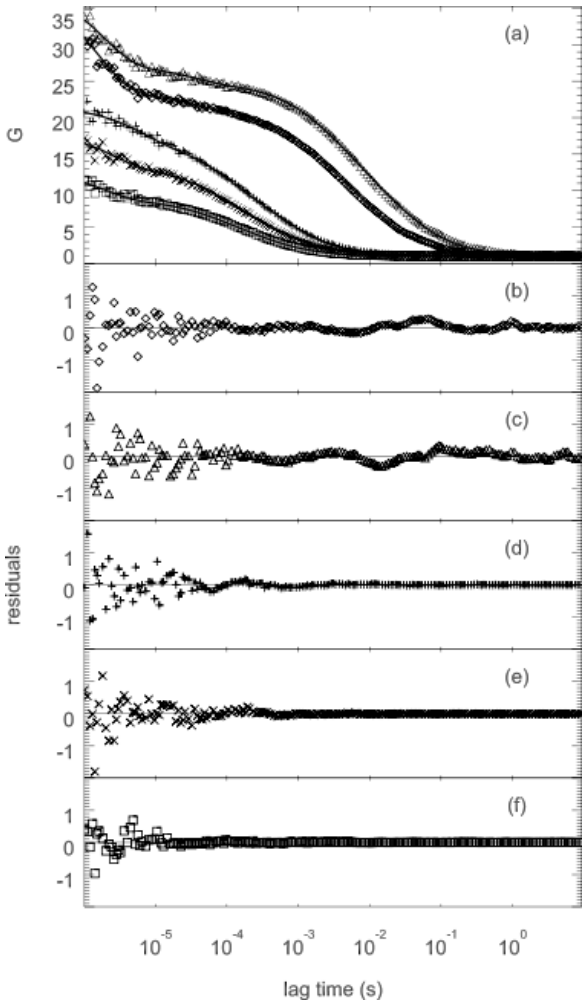

Figure 3.6. FCS data, fits, and residuals for donor-only labeled RNA. Diamonds denote attodroplets in FC-77; triangles denote attodroplets in FC-40, squares denote solution data at pH 7, × denotes solution data taken at pH 6.0, and + denotes solution data at pH 3.0. FCS data and fits are shown in panel (a); residuals are below in panels (b)-(e). The larger residuals for droplet data seen in (b) and (c) result from the difficulty of sampling a sufficient number of the slowest-moving droplets, which contribute substantially to the deviations at long correlation times.

sample	A_1	τ_D (μs)	F_T	τ_{k_T} (μs)	F_I	τ_{k_I} (μs)
pH 3	11.57 (0.49)	483 (23)	0.22 (0.04)	5.93 (2.44)	0.27 (0.04)	83 (28)
pH 6	8.64 (0.51)	366 (24)	0.29 (0.05)	2.53 (0.95)	0.29 (0.04)	76 (23)
pH 7	6.45 (0.07)	224 (3)	0.34 (0.04)	1.96 (0.34)	0.17 (0.01)	50 (6)

Table 3.1. Fit parameters for the solution data of Fig. 3.6 for a function of the form of Eq. 3.7. Uncertainties are given in parentheses.

sample	A_2	τ_{D_2} (ms)	A_1	τ_{D_1} (μs)	F_T	τ_{k_T} (μs)	F_I	τ_{k_I} (μs)
FC 40	10.84 (0.99)	29.83 (2.64)	13.29 (0.9)	3093 (413)	0.42 (0.04)	1.37 (0.27)	0.10 (0.01)	34.7 (11.2)
FC 77	12.22 (0.75)	9.61 (0.50)	4.83 (0.61)	930 (322)	0.49 (0.04)	1.13 (0.19)	0.11 (0.02)	30.00 (0.53)

Table 3.2. Fit parameters for the droplet data of Fig. 3.6 for a function of the form of Eq. 3.7. Uncertainties are given in parentheses.

sample	A	τ_D (ms)	B	τ_B (μs)	α	F_T	τ_{k_T} (μs)	F_I	τ_{k_I} (μs)
FC 40	14.75 (0.72)	24.16 (1.63)	0.67 (0.1)	3354 (423)	0.74 (0.08)	0.43 (0.05)	1.24 (0.3)	0.09 (0.02)	23.6 (9.9)
FC 77	13.59 (0.44)	8.88 (0.34)	0.31 (0.08)	784 (213)	0.66 (0.20)	0.51 (0.06)	0.92 (0.21)	0.12 (0.03)	14.8 (6.7)

Table 3.3. Fit parameters for the droplet data of Fig. 3.6 for a function of the form of Eq. 3.8 (model was not shown in the figure). Uncertainties are given in parentheses.

timescales gave substantially worse fit results (Table 3.2). The longer diffusive time unambiguously corresponds to the diffusion time of the droplets through the detection volume. Other relevant time scales include the diffusion time of the molecule across the droplet and diffusion time of the molecule across the detection volume. For droplets very much larger than the detection volume, this faster time should approach the same diffusion time seen for RNA in solution. For droplets much smaller than the detection volume, the only relevant time scale will be that of droplet diffusion. Here we are in an intermediate regime where this smaller diffusion time appears to be somewhere between the two limiting timescales. No evidence for a slow interconversion, other than the triplet and isomer, is evident in the droplet data. Static populations with different brightness are not evident in FCS but show up clearly in the PCH and are the likely cause of heterogeneity in the FRET data.

Here the FCS simulations provide a better description of the system. Our simulations show that the longer correlation time in ACF corresponds to the diffusion time of the most frequently observed size droplets. When the droplets are diffusing in FC-40, parameters describing size distribution (assuming lognormal) are $\mu = 4.88$, $\sigma = 0.20$, and $Mode = 126.5$ nm. In addition, the acquired ACF from same simulation data is found to be almost completely overlapping with the experimental ACF, when the diffusion time of molecule within a droplet is long (on the order of a mil-

lisecond). If it is considered as an effect of only the viscosity, this long diffusion time corresponds to a dynamic viscosity of 16 cP within droplets. Comparison of simulated and experimental ACFs are shown in Fig. 3.7.

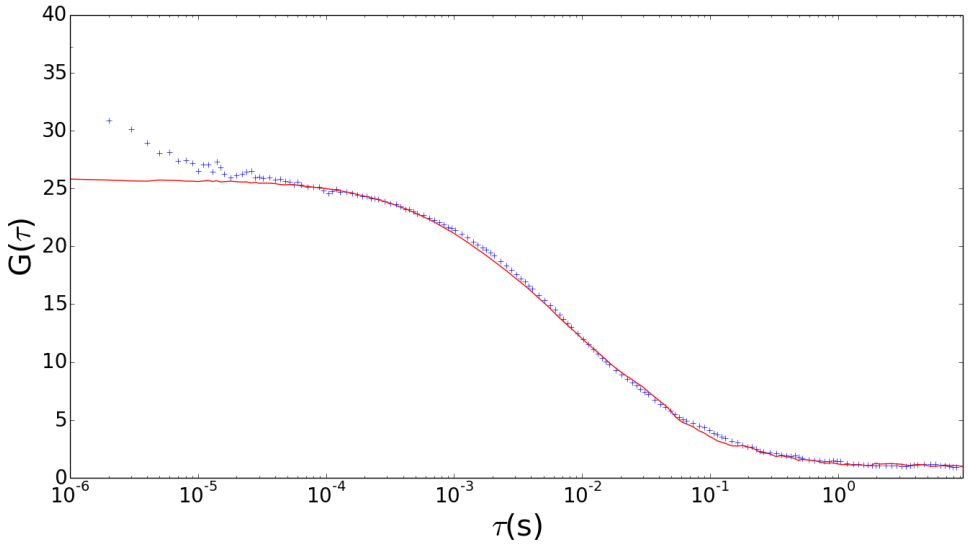

Figure 3.7. Comparison of simulated (red solid line) and experimental (blue +) ACFs. Simulation assumes droplets are in FC-40. Size distribution parameters are $\mu = 4.88$ and $\sigma = 0.20$. Bulk RNA concentration was set to be 10 nM. Actual data for simulation and for experiment can be found in the place provided in Table. D.10.

Measurements of pH in droplets as we have shown in Table. 2.4 confirm that the attodroplet environment is quite acidic. Although, only reported values were for droplets in FC-40, similar measurements for droplets in FC-77 showed similar results (data not shown). Moreover, the effect of enzymatic oxygen quenching system and/or reducer/oxidizer (ROXS) system on droplet pH was not more than 0.1. Similar effect was observed between the cases of having a surfactant in droplet making procedure or not.

We also showed that just to push the droplet pH to near neutral values we needed to add 50 mM or more NaOH into bulk RNA solution prior to droplet making. Never-

theless, creating a highly basic environment in bulk solution provokes dye cleavage at the end of RNA. Therefore, we used maximum 20 mM NaOH in bulk RNA solution prior to FRET measurements which was supposed to increase droplet pH to 6.47. Consistent with the results of solution FRET data, proximity ratio histograms taken with between 0 and 20 mM NaOH (Fig. C.3) showed no systematic change from the data of Fig. 3.3, although the relative height of the two peaks identified in Fig. 3.3 did change from data set to data set. This change did not appear to be related to pH; while the high FRET peak was nearly always smaller than the low FRET peak, the relative amplitudes changed between data sets even for samples prepared in the same way.

3.5 Discussion

Neither RNA nor the indocarbocyanine dyes used here are expected to undergo large structural changes at low pH. However, while most ribonucleotides have a pKa below 4, the exception is CMP with a pKa of 4.5; the presence of the phosphate backbone can also increase the pKa.[13] It seems likely that the slow dynamics apparent at low pH are due to a changing interaction between the dye and the RNA. For example, if protonation of the terminal C causes a fraying of the end of the RNA, the cyanine dyes could be intercalating into the single strand. This is consistent with observations that Cy3 is brighter on single-stranded oligonucleotides, and that the presence of Cy3 can lower the melting temperature of RNA; effects that have been attributed to intercalation.[133] Intermittent cyanine dye intercalation, made possible by RNA fraying due to protonation of the cytosine, might explain the broad FRET peak, the bright species evident at low pH, and the FCS results in solution at low pH.

The situation in droplets is more interesting. The bright species is present, and the average value for the proximity ratio is similar to that observed in solution at

low pH this is consistent with the previous pH measurements in Sec. 2.3. However, the shape of the proximity distribution is qualitatively different for droplets and solution data. If the broadening or splitting of the FRET peaks is due to different configurations of the dye on the RNA, then in droplets the transition between states appears static on the timescale of the FCS measurements. Observed long diffusion times in droplets through FCS modelling may be a clue for the mechanism within droplets which provides static like heterogeneity here.

There is a small possibility that the difference in the shapes of the proximity histograms of Fig. 3.5 might simply arise from the significantly greater number of photons collected from FRET data in droplets, which leads to narrower FRET peaks because of reduced shot noise. If this was the case, then it should be possible to see a similar shape in both sets of histograms by thresholding in such a way that the shot noise is similar for solution and droplet data. The limitations of FRET with freely-diffusing molecules make this a difficult comparison, but we did make the attempt; the disparities in the proximity distributions persisted.

One difference between the droplet and bulk environment that cannot be ignored is that the droplet contents are determined stochastically at the time of droplet formation. This may help further the apparent static heterogeneity of RNA in droplets. Drop-to-drop changes in NaCl at 200 mM is probably not a significant player in this regard, since there are an average of 1.2 million salt molecules in a typical 10 aL droplet. Furthermore, changes in salt concentration from 100 mM up to 800 mM in solution made no noticeable difference in FRET (data not shown). However, small changes in the content of hydronium or hydroxyl ions can make an enormous change in the pH of the droplet. Consider that in a 10 aL droplet, pure water would have only 0.6 hydronium ions, on average. Buffer at 20 mM will have an average of 120,000 molecules in the same droplet, which would seem to be sufficient to maintain pH. However, the effect of the surface cannot be ignored. While the surfactant used here

is uncharged, up to 5% of the surfactant components are unreacted,[72] potentially giving millions of acid groups on the surface of a typical droplet. However, as discussed above and shown previously,[54] it is quite possible to make attodroplets in these oils without surfactant; in that case, the pH increases only marginally and is still similar to measured values in Table. 2.4. As the last contributor to stochastic nature of droplet contents, diffusion time for individual molecules within droplets can change as a function of droplet radius. Having said that, there is a possibility that interactions with the boundary might affect dynamics in the droplet. Negative charge accumulation and the binding of hydroxyl ions to PFC/water interface [105] can result in reorganization of water within droplets and so might explain the slower diffusion of small RNA oligos in the droplet.

CHAPTER 4

CONCLUSIONS

With notable exceptions,[120, 78, 124] almost all analysis of FRET data assumes dyes that are freely rotating on a timescale that is fast compared to τ_D. Indeed it is now common to extract distances or distance distributions from FRET data in this limit. Here we have modeled a common system for which steric hindrance and "sticking" of the dyes on RNA makes the slow-rotation approximation a better choice. For fast free-rotation, it is well known that fast distance fluctuations about $R = \langle R \rangle$ result in an increase in FRET over the case of static R. [151] Here we see a case where slow-rotation and steric hindrance gives a reduction in the FRET efficiency over what might be expected from static values of $R = \langle R \rangle$ and $\kappa^2 = \langle \kappa^2 \rangle$.

The failure of the free-rotation approximation for 5' attached dyes is perhaps not surprising given that the κ^2 distribution from the MD simulation is quite different than what would be expected for free rotation. However, for 3' attached dyes we find considerable configurational freedom that mimics the distribution of κ^2 for freely-rotating dipoles. This is not sufficient to justify the use of the freely-rotating approximation because the coherence time for free rotation is longer than τ_D in all cases explored here. None of the approximate forms are as accurate as the MC simulation, which accounts naturally for the large range of coherence times for both R and κ^2.

The details of dye configuration and dynamics and the predicted FRET distributions were also shown to depend on the terminal base. The work of Iqbal *et al.*,[78] showed dramatic oscillation in FRET as a function of duplex length due to base stack-

ing; the nucleic acids used in their work had Cy3 and Cy5 attached to a 5' terminal cytosine. Of those modeled here, this is the configuration for which we would expect the greatest effect of stacking and the largest correlation between R and κ^2, as only one structure dominates for cyanine dyes on 5' terminal C, Fig. 1.5(a) and (c). We have not modeled cyanines on A or U, but note that Spiriti *et al.*[159] predict that stacking interactions are weakest for Cy3 attached to T on B-DNA. They further point out that this might be attributed to the methyl group on thymine, which is absent in RNA as T is replaced by U.

In summary, in Chapter. 1, we have shown that FRET from cyanine dyes attached to RNA depends on both the terminal base[159] and the details of how the dyes are linked to the RNA. In three of four cases studied, 3' attached dyes do not achieve a stable stacked configuration; these dyes explore a much larger region of configuration space than 5' attached dyes. The configurational freedom is particularly evident for dyes attached to a 3' terminal G. An exception is Cy5 on a 3' terminal C, which is found almost entirely in the stacked geometry. The configurational freedom of Cy3 at the 3' terminus is sufficient to yield a distribution of κ^2 similar to that expected for freely rotating dyes. Nonetheless, the use of the free-rotation approximation is unjustified for any attachment of cyanine dyes studied here; this is due to the presence of rotational coherences that are both slower and faster than τ_D, and correlations between κ^2 and R. From comparison with the MC simulations, we see that the best analytical model for these duplexes assumes slow relative rotation of the dyes and fast fluctuations in relative distance.

In Chapter. 2, we have presented a simple and efficient emulsification procedure for making sub-micron size aqueous droplets. Characterization studies in terms of droplet sizes on the emulsion system were once done by conventional DLS technique. A new, cost and labor effective method, which utilizes Mie scattering theory, for same measurements has been presented. Results of new method has been shown to

be in agreement with the DLS measurements. Although further proof of its sensitivity postponed to FCS measurement results through modelling in Chapter. 3, new method was expected to be more precise than DLS thanks to the Mie scattering. It has also been shown that versatile nature of the new method makes it superior to DLS, because it allows the size measurements as well as pH measurements within droplets in a single experiment simultaneously.

In addition, pH within sub-micron size droplets has been shown much lower than initial bulk solution prior to droplet making. Attempts for controlling the pH within droplets by introducing strong base into initial bulk solution only increased it to near neutral values. Comparison of our results with the literature suggests that the pH drop is mostly driven by the interface. Therefore, we have concluded that the implications of our results cover a wide range of emulsion studies whenever sub-micron size W/O emulsions are used.

In Chapter. 3, we have further proved that the droplet environment is acidic and translational diffusion across the volume is substantially slower than would be expected in from the bulk properties of water. These unique properties of attoliter droplets has shown to be driven by the physical properties of the oil/water interface, particularly the droplet radius. Although these were considered to be the source of static heterogeneity in FRET, the actual reason for heterogeneity in small RNA system is currently unknown.

In picoliter droplets commonly used today, pH changes and viscosity differences are minimal because the surface/volume ratio is three orders of magnitude smaller than we use here, giving relatively more buffer per surface area. We have shown that to control pH in these very small droplets requires the addition of a considerable excess of a strong base, which may not be compatible with the biomolecules in the bulk phase. It seems likely that a surfactant capable of directly buffering the surface will need to be employed to successfully control pH in droplets. Alternatively, a

polyprotic buffer system (multiple pK_a values) has also been considered to have the potential for resisting the interface driven pH changes.

The high signal-to-noise and the extended dynamic range of FCS afforded by attodroplet confinement is at once promising and challenging. We have demonstrated that attodroplet confinement significantly extends the dynamic range of solution FCS measurements, and increases the signal-to-noise of FRET measurements, not only by slowing the diffusion of molecules through a detection volume but also slowing them down within the droplets too by observed long diffusion times. We have also demonstrated that attodroplet confinement appears to reveal heterogeneities that might be hidden in solution data.

Even though, the low pH was manifested itself as a problem for FRET of the small RNA duplex we used here, attoliter droplets should be considered a very strong and robust alternative to currently available single-molecule techniques. To minimum, long diffusion times of droplets a well as the diffusion times within droplets in addition to stochastic nature of the droplet environment offer an experimental system virtually equivalent to surface attachment techniques but providing much more detail information on state space of the molecules within. More importantly, they provide these informations without any need of a linker which is potentially a source of perturbations.

We have noted that the use of single-molecule fluorescence gives us a new tool to investigate the attodroplet interior, and by doing so elucidate the nature of the perfluorinate/water boundary and provide new routes to control the chemical and physical environment in these very small reactors. In that sense, modelling FCS was crucial in the lack of analytical models. Furthermore, the versatile structure of simulation code allows us to alter system parameters as we like, so we are currently able to calculate synthetic correlation functions for various system geometries as well as various forms of point spread functions.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR FRET OF A SHORT RNA: EFFECT OF ATTACHMENT STRATEGY ON FREE ROTATION

A.1 MD trajectories

Trajectories of R, κ^2, and E for the 5′R, 3′F, and 3′R duplexes are shown in Figs. A.1, A.2 and A.3 respectively. To avoid plotting 300,000 points, each trajectory is a two-dimensional histogram 500 bins along the abscissa and 300 along the ordinate. The distance between the geometrical centers of the dye molecules, R, is shown in panel (a) of each figure. The orientation factor κ^2 is given by 5 in the text and shown in panel (b) of each figure. The transition dipoles used to calculate κ^2 are approximated to be parallel to the conjugated bond structure between the two indole-like moieties. E is the instantaneous value of energy transfer efficiency, calculated from Eqs. 1 to 3 in the text using the MD trajectories for R and κ^2 and taking $R_0 = 5.8$ nm.[181, 117, 135, 120].

A.2 Additional considerations in modeling

A.2.1 Ergodicity of the MD simulation

The requirement for ergodicity of the MD simulation presents challenges. Our MD simulation was limited by computational speed to compute at best tens of nanoseconds per day. This is typically more than two orders of magnitude shorter than the average time between photons in a typical single-fluorophore sensitive measurement, and much shorter than conformational switching events that are typically studied by single-molecule techniques. Even the use of search algorithms [170, 145, 89, 159] designed

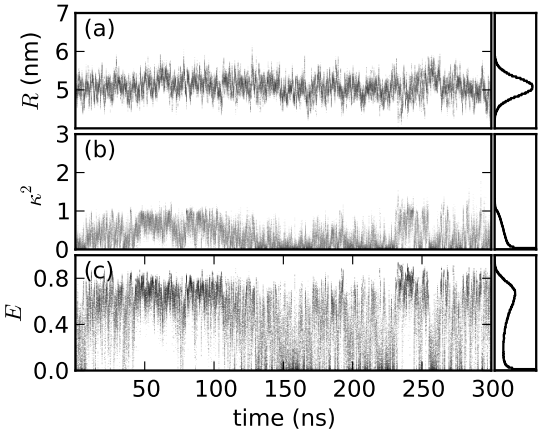

Figure A.1. Trajectories and histograms for 5′R duplex. (a) The distance R between the center of mass of the two dyes; a histogram of R values is on the right. (b) The orientation factor κ^2; a histogram of κ^2 values is on the right. (c) Instantaneous values of E; a histogram of E values is on the right. Here $\langle R \rangle = 5.08$ nm, $\langle \kappa^2 \rangle = 0.36$, and $\langle E \rangle = 0.447$. The Pearson coefficient between κ^2 and R is -0.008. To avoid plotting 300,000 points, each figure is a two-dimensional histogram 500 bins wide and 300 tall using grayscale where white corresponds to 0 occurrences and black corresponds to (a) ≥ 50 occurrences; (b)≥ 100 occurrences; (c)≥ 30 occurrences. In all cases $\leq 2.4\%$ of the non-zero pixels are saturated.

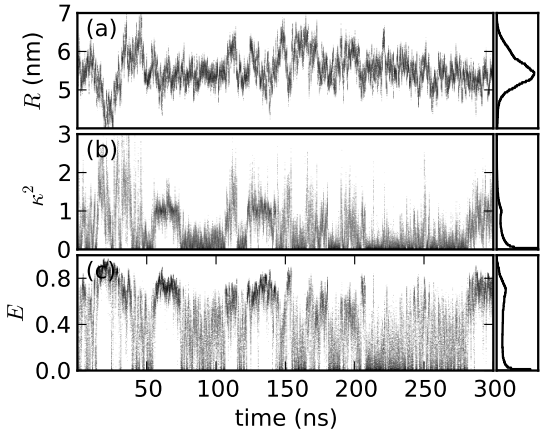

Figure A.2. Trajectories and histograms for 3′F duplex. (a) The distance R between the center of mass of the two dyes; a histogram of R values is on the right. (b) The orientation factor κ^2; a histogram of κ^2 values is on the right. (c) Instantaneous values of E; a histogram of E values is on the right. Here $\langle R \rangle = 5.50$ nm, $\langle \kappa^2 \rangle = 0.53$, and $\langle E \rangle = 0.402$. The Pearson coefficient between κ^2 and R is 0.037. To avoid plotting 300,000 points, each figure is a two-dimensional histogram 500 bins wide and 300 tall using grayscale where white corresponds to 0 occurrences and black corresponds to (a)$\geq$ 50 occurrences; (b)$\geq$ 100 occurrences; (c)$\geq$ 30 occurrences. In all cases $\leq 0.3\%$ of the non-zero pixels are saturated.

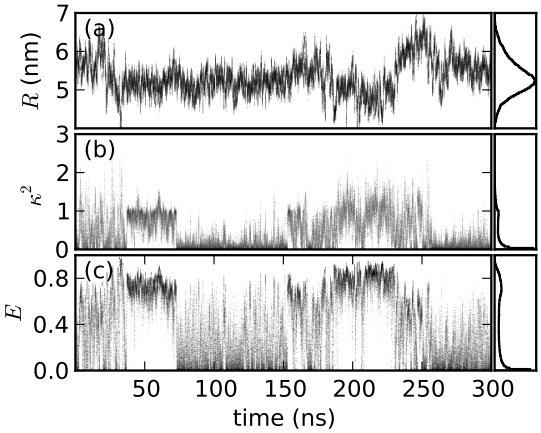

Figure A.3. Trajectories and histograms for 3′R duplex. (a) The distance R between the center of mass of the two dyes; a histogram of R values is on the right. (b) The orientation factor κ^2; a histogram of κ^2 values is on the right. (c) Instantaneous values of E; a histogram of E values is on the right. Here $\langle R \rangle = 5.34$ nm, $\langle \kappa^2 \rangle = 0.43$, and $\langle E \rangle = 0.392$. The Pearson coefficient between κ^2 and R is -0.133. To avoid plotting 300,000 points, each figure is a two-dimensional histogram 500 bins wide and 300 tall using grayscale where white corresponds to 0 occurrences and black corresponds to (a)$\geq$ 50 occurrences; (b)$\geq$ 100 occurrences;(c), black $\geq$ 30 occurrences. In all cases $\leq 0.4\%$ of the non-zero pixels are saturated.

to sample wider regions of phase space do not guarantee ergodicity without *a priori* knowledge of accessible states.

In the simulations discussed here, there are obvious fluctuations in R and κ^2 that occur on the timescale of tens of nanoseconds. It is not surprising then, that large variations in the predicted value of FRET occur for runs less than 100 ns. However, within the limitations of our simulation, we do find that the predicted FRET distribution generated by MC methods does not change substantially after 240 ns. This is demonstrated in Figs. A.4 through A.7.

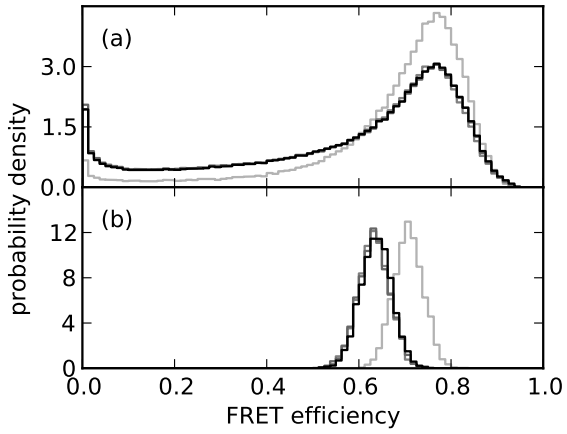

Figure A.4. FRET efficiency histograms of (a) the instantaneous and (b) MC predicted E distribution for the 5′F duplex after 80 ns (light gray), 160 ns (gray), 240 ns (dark gray), and 300 ns (black).

A.2.2 Correlation of the fluctuations in R and κ^2

Autocorrelation functions of R and κ^2 are given in Figs. A.8 and A.9, respectively. From these correlation functions it is clear that there are both long and short components to the correlation times. As expected from the results in the text, Fig.A.8 shows correlation times for R are shorter than those of Fig. A.9 for κ^2. Because most of these correlation functions have significant artifacts associated with the finite length of the MD trajectory, we make no attempt at fitting these data.

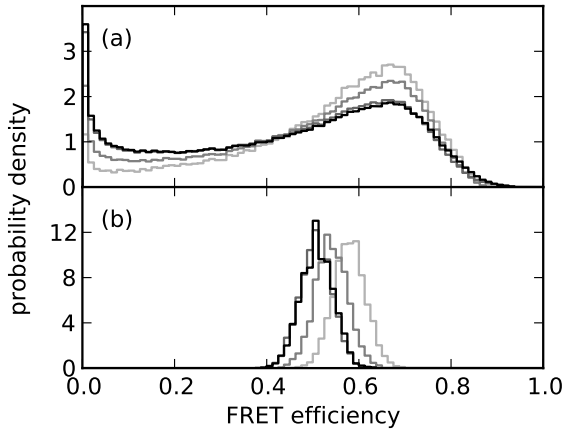

Figure A.5. FRET efficiency histograms of (a) the instantaneous and (b) MC predicted E distribution for the 5′R duplex after 80 ns (light gray), 160 ns (gray), 240 ns (dark gray), and 300 ns (black).

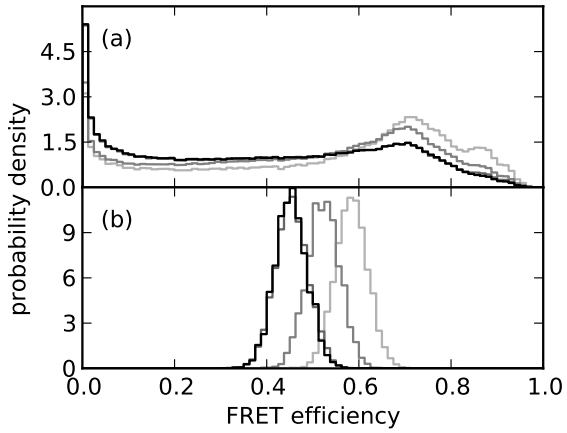

Figure A.6. FRET efficiency histograms of (a) the instantaneous and (b) MC predicted E distribution for the 3′F duplex after 80 ns (light gray), 160 ns (gray), 240 ns (dark gray), and 300 ns (black).

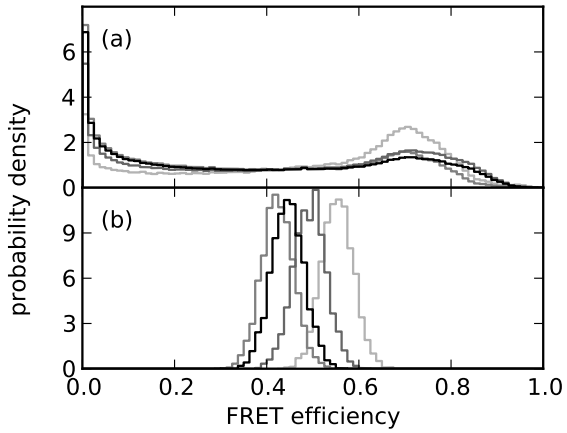

Figure A.7. FRET efficiency histograms of (a) the instantaneous and (b) MC predicted E distribution for the 3′R duplex after 80 ns (light gray), 160 ns (gray), 240 ns (dark gray), and 300 ns (black).

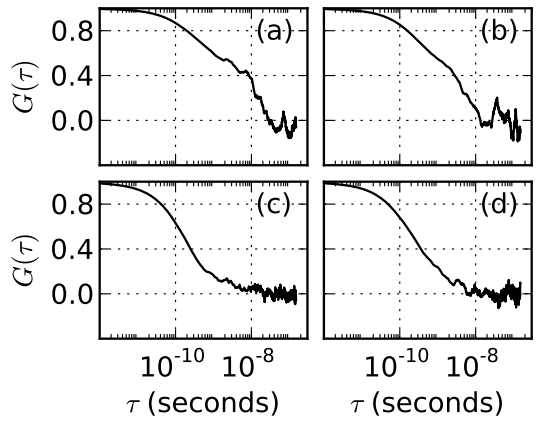

Figure A.8. Autocorrelation function of the 300 ns trajectory for R for the (a) 3′R, (b) 3′F, (c) 5′R, (d) 5′F duplexes.

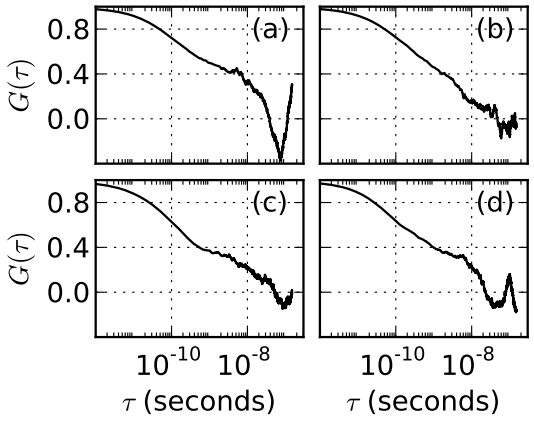

Figure A.9. Autocorrelation function of the 300 ns trajectory for κ^2 for the (a) 3′R, (b) 3′F, (c) 5′R, (d) 5′F duplexes.

A.2.3 Data Acquisition and Analysis

Donor (Cy3) and acceptor (Cy5) labeled ribonucleotides were prepared in 20 mM HEPES-NaOH at pH 7.8 with 100 mM NaCl and 5 mM $MgCl_2$. In this buffer, RNA at 2.5 μM was heated to 90 °C in 5 minutes and then incubated at 90 °C for 60 minutes before cooling to 4 °C over 60 minutes. The resulting dsRNA samples were then diluted to a concentration of 100 pM with 15 nM protocatechuate-3,4-dioxygenase (PCD) and 5 mM protocatechuic acid (PCA) as an enzymatic oxygen quenching system.[1]

Single-molecule sensitive FRET measurements were accomplished using an Olympus IX50 inverted microscope modified for confocal microscopy and a UPlanSApo 60X 1.2 NA water immersion lens. The 514 nm line of an all-lines air-cooled argon-krypton laser was used at an excitation power of 50 μW. Fluorescent photons were detected in two channels (donor, acceptor) using PerkinElmer SPCM-AQRH-15 avalanche photodiodes. Photon timing information was recorded with 8 ns timing resolution using homemade circuitry. Optics and circuitry are described in detail elsewhere.[53]

FRET histograms were corrected for background, crosstalk, and the differences in quantum yield and/or collection efficiency of the acceptor and donor channels

(contained in the detection-correction factor, γ). To construct FRET histograms, photon arrival times in each of two channels (donor and acceptor) were histogrammed into 5 ms bins. Bins with more than 20 photons (sum of both channels), corresponding to a labeled RNA molecule crossing the detection volume, were used in calculations of FRET. Background rates in the donor and acceptor channels were determined by taking the average number of counts per bin for all bins below the threshold of 20 counts; this overestimated background by as much as 8% when compared with a background determination based on statistical inference, but the difference in final FRET values is inconsequential to the results here. The crosstalk, $\chi = 0.18$, from the donor to the acceptor channel was determined from data taken on RNA molecules prepared as described above, but missing Cy5. The detection-correction factor, γ, was determined for each data set by considering the average number of photons per bin in the FRET peak, and the average number of photons per bin in a donor-only sample. The uncertainty is an estimate arrived at by considering different values of bin and threshold. γ is then given by:

$$\gamma = \chi - \frac{\langle N_a^{FRET} \rangle - \langle N_a^{Donly} \rangle}{\langle N_d^{FRET} \rangle - \langle N_d^{Donly} \rangle} \tag{A.1}$$

The FRET efficiency was computed for each bin using

$$E = \frac{N_a - \chi N_d}{N_a - \chi N_d + \gamma N_d} \tag{A.2}$$

Here N refers to the number of photons in each above-threshold bin, corrected for background. The subscripts a and d refer to photons in the acceptor or donor channel respectively. The superscripts FRET and Donly refer to photons associated with the FRET peak for FRET data and the single donor peak for donor-only data. The resulting FRET values are then histogrammed and shown in Fig. 10 in the text.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHARACTERIZATION OF SUB-MICRON SIZE DROPLETS FOR SINGLE-MOLECULE STUDIES

B.1 Light Attenuation by Mie Scattering and Modelling

B.1.1 Direct Application of Mie Theory

The actual formulas for Mie scattering are well known [14]. The solution for finding the Mie scattering cross-section of a given size spherical particle involves an incident plane wave and an outgoing spherical scattered wave. Because of the spherical symmetry of the system, the incident wave is expanded as an infinite series of vector spherical harmonics. Using the boundary conditions for electric and magnetic fields and after considerable mathematical manipulation, scattered fields can be determined. Procedure can be followed by the calculation of differential and total cross sections for spherical particle. Key parameters for these calculations are the Mie coefficients a_n and b_n to compute the amplitudes of the scattered field. These parameters are given by [14];

$$\begin{aligned} a_n &= \frac{\mu m^2 j_n(mx)\partial_x[xj_n(x)] - \mu_1 j_n(x)\partial_x[mxj_n(mx)]}{\mu m^2 j_n(mx)\partial_x[xh_n^{(1)}(x)] - \mu_1 h_n^{(1)}(x)\partial_x[mxj_n(mx)]}; \\ b_n &= \frac{\mu_1 j_n(mx)\partial_x[xj_n(x)] - \mu j_n(x)\partial_x[mxj_n(mx)]}{\mu_1 j_n(mx)\partial_x[xh_n^{(1)}(x)] - \mu h_n^{(1)}(x)\partial_x[mxj_n(mx)]}; \end{aligned} \tag{B.1}$$

where m is the relative refractive index of the sphere, the j_n's are spherical Bessel functions of the first kind, the h_n's are spherical Hankel functions, μ and μ_1 are the

magnetic permeability of the sphere and surrounding medium, respectively. For the present case $\mu = \mu_1$, and hence they cancel. The quantity $x = (2\pi n_{media} r)/\lambda$ is called the size parameter and ∂_x indicate derivatives with respect to x.

Using the Mie coefficients a_n and b_n, extinction efficiency for a spherical scatterer, and the reduced attenuation spectra of a bulk homogeneous uniform size sample can be approximately expressed by the following equations:

$$\begin{aligned} Q_{ext} &= \frac{2}{x^2}\sum_{n=1}^{\infty}(2n+1)Re(a_n + b_n), \\ A_{red} &= N_0 Q_{ext}\pi r^2 \end{aligned} \tag{B.2}$$

where A_{red} is the reduced attenuation (absorbance) of light by scattering so that $A = lA_{red}$ in which l is the path length of instrument, Q_{ext} is the unitless extinction efficiency, N_0 is the particle number density and r is the particle radius. The framework can be extended to approximate a multi-sized scattering particle medium by summing the scattering contributions over all particle sizes and adding an approximate normalized size distribution factor w_i that compensates for the number of particles at each given size, in which case Eq. B.2 becomes [149, 174];

$$A_{red} = N\sum_i w_i Q_{ext\,i}\pi r_i^2, \quad \sum_i w_i = 1 \tag{B.3}$$

The Eq. B.3 is a direct application of Mie scattering calculations to a real system. Because Q_{ext} values are actually infinite sums, they are often calculated up to a proposed value of n_{max} [14] given by;

$$n_{max} = x + 4x^{1/3} + 2 \tag{B.4}$$

Consequently, above method doesn't provide an analytical expression for describing the light attenuation. To model the experimental data, one needs to first construct

a grid of Q_{ext} values as a function of scatterer sizes and applied wavelengths. Once Q_{ext} values are calculated, using some *a priori* information about the particle size distribution, light attenuation data can numerically be reconstructed. As a last ingredient to the method, A_{red} can be converted to actual absorption by multiplying it with the measurement path-length in the experiment. Then, the attenuation in its final form as shown in Eq. B.5 can be used for finding the parameters of that distribution.

$$A = lN \sum_i w_i Q_{ext} \pi r_i^2, \quad \sum_i w_i = 1 \tag{B.5}$$

The size distribution for the droplets in an emulsion can be practically any kind of heavy-tailed distribution. This is a family of probability distributions which have heavier tails than the exponential distribution. The reason of having a heavy-tailed droplet sizes was explained in literature by a very simple model of two competing precesses. First is the one drives the disintegration in system, which can be called homogenization process. The second is the opposite of homogenization, which can be called coalescence process. A simple differential for describing the size distribution in emulsion is given by [37];

$$\frac{dn}{dx} = a\frac{n}{x}(X - x) \tag{B.6}$$

where n is the number of droplets with the size x, a is the net rate constant for homogenization process, X is the characteristic or most frequent size for the droplets, and N is the total number of droplets in the system. Using the constraints on the size distribution of droplets $\int_0^\infty f(x)dx = 1$, and positive semi-definite nature of $f(x)$, solution to Eq. B.6 was given by;

$$\frac{n}{N} = \frac{a^{aX+1}}{\Gamma(aX+1)} x^{aX} e^{-ax} \tag{B.7}$$

As it is seen in Eq. B.7, particle sizes in an emulsion can follow Γ distribution. On the other hand, for all practical purposes, droplet size distribution within emulsions are considered as lognormal distribution [107, 130, 116].

Therefore, for reconstructing the attenuation data in Eq. B.5, w_i values are chosen from a lognormal distribution. Lognormal distribution can be expressed by two parameters μ and σ which are respectively the mean and standard deviation of the variable's (x) natural logarithm. Therefore the functional form for the probability density function (PDF) of lognormal distribution is given by;

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}}, \ for\ x > 0 \tag{B.8}$$

PDF in Eq. B.8 can be equivalently described any two independent parameters of the distribution. For all practical purposes, we used the parameters median (2-quantile or the radius value separating the higher half of the PDF of sizes from lower half), and mode (the radius value that appears most often in PDF) to describe it. These two parameters and their relations with parameters μ and σ are given as following;

$$\begin{aligned} Median = e^{\mu} \quad , \quad Mode &= e^{\mu-\sigma^2} \\ \mu = \ln(Median) \quad , \quad \sigma &= \frac{1}{2}\ln(\frac{Median}{Mode}) \end{aligned} \tag{B.9}$$

The application of above procedure to data as shown in Sec. 2.3.1 was done as following. First, Q_{ext} values were calculated for a droplet size range of 10 nm to 1000 nm (by 1 nm increments) and a wavelength range of 400 nm to 850 nm (by 1 nm increments). This was done with a software (Mie Scattering Calculator, developed by Scott Prahl) which can be found in the public domain at `http://omlc.ogi.edu/calc/mie_calc.html`. A Python code fragment for calculating the Q_{ext} values as

well as a short "HowTo" document is provided in Sec. D.2. Storage locations of calculated Q_{ext} values for droplets in FC40, FC77, and polystyrene beads in water are given in Table D.6. Based on initial guess of parameters $Median$ and $Mode$, weights w_i for the size distribution was assigned using Eq. B.8. In addition, an initial guess for the number density of droplets N_0 was made. Finally using the Eq. B.5, a numeric absorption function was constructed. Fitting of the numeric function was done using constrained Levenberg-Marquardt non-linear least squares method over the parameters N_0, $Median$ and $Mode$, meaning that at every step of iteration over the fitting parameters, numeric absorption function was reconstructed using updated weights.

B.1.2 Attenuation Power Law Fit: Ångström Formula

Rather than applying Mie scattering theory directly, an alternative and more empirical approach can be used for light attenuation through bulk of Mie scatterers. The method first proposed by Ångström for quantifying the particulate matter size in atmosphere [3, 84]. It basically relates the transmission (T) to the concentration and sizes of scatterers inside the medium under the assumption of spherical scatterers with a broad range of particle sizes. T from a collection of Mie scatterers is given by,

$$T = \frac{I}{I_0} = e^{-l\tau}, \tag{B.10}$$

in which τ is the Optical thickness or Optical depth. Optical thickness is given by,

$$\tau = \beta\lambda^{\alpha}. \tag{B.11}$$

where λ is the wavelength of light, β is called the turbidity coefficient, and α is the Ångström exponent. Eq. B.11 in log-log scale is;

$$\log_{10}\tau = \log_{10}\beta + \alpha \log_{10}\lambda, \tag{B.12}$$

Therefore Eq. B.12 is linear with log of wavelength ($\log\lambda$). In a later work, King and Byrne suggested that Ångström formula needs to have an additional second order term in log-log scale for correcting the finite size distributions of Mie scatterers [87, 162]. Corrected form of the equation is;

$$\log_{10}\tau = a_0 + a_1 \log_{10}\lambda + a_2(\log_{10}\lambda)^2. \tag{B.13}$$

Although, Eq. B.13 is derived by considering the transmission through a series of Mie scatterers, one can adapt it for absorption measurements. Relation between absorption (A) and transmission (T) for a typical UV/Visible absorption measurement is give by;

$$A = -\log_{10} T, \tag{B.14}$$

Using Eqs. B.10 and B.13, Eq. B.14 can be rewritten as;

$$A = l\frac{10^{a_0 + a_1 \log_{10}\lambda + a_2(\log_{10}\lambda)^2}}{\ln(10)} \tag{B.15}$$

The Ångström Formula in the the form of Eq. B.15 is a functional form that attenuation data is supposed to follow. On the other hand, Eq. B.5 as the direct application of Mie scattering theory is a numeric form. It is easier to work with former in the fits, so for BTB loaded droplets we preferred to identify the Mie scattering part of the absorption/attenuation by using Ångström Formula.

Only assumption in the fitting procedure was the summability of attenuations by real absorption of BTB and by the Mie scattering of the droplets so that $A_{tot} = A_{Mie} + A_{BTB}$. For simplicity, BTB absorption was expected to contribute the total signal below a fixed wavelength λ_0. Thus the rest of the data which corresponds

$\lambda \geq \lambda_0$ was used in fits to Ångström formula. Best results were found when $\lambda_0 =$ 650 nm. In the absence of BTB, i.e. absorption from empty droplets, whole data set was used for fits. Although, there is no easy way of interpreting their meaning in terms of the size parameters, fit parameter values for the empty droplet experiments are given in Tables. B.1 and B.2. Actual data used in fits were given in Figs. 2.4 and 2.6 correspondingly.

	a_0	a_1	a_2
Sample 1	4.87 (0.13)	-1.25 (0.09)	-0.22 (0.02)
Sample 2	4.84 (0.14)	-1.00 (0.07)	-0.26 (0.02)

Table B.1. Fitting coefficients for Ångström formula to the data shown in Fig. 2.4. Actual data can be found in the place provided in Table D.2.

t (min)	a_0	a_1	a_2
0	4.87 (0.13)	-1.25 (0.09)	-0.22 (0.02)
10	4.41 (0.35)	-1.00 (0.21)	-0.25 (0.05)
20	4.34 (0.22)	-1.00 (0.13)	-0.25 (0.03)
30	4.25 (0.51)	-1.00 (0.31)	-0.25 (0.06)

Table B.2. Fitting coefficients for Ångström formula to the data shown in Fig. 2.6. Actual data can be found in the place provided in Table D.3.

B.2 Ratio-metric pH Measurements : Absorption

Aqueous solution of a pH sensitive fluorophore contains both of its conjugate acid and base forms. Thus it is considered as an equilibrium reaction between two species.

$$HA + H_2O \rightleftharpoons H_3O^+ + A^- \tag{B.16}$$

for the reversible reaction in Eq. B.16, equilibrium relation is given by;

$$K_a = \frac{[H_3O^+][A^-]}{[HA]} \tag{B.17}$$

Total fluorophore concentration and the concentrations of each forms are;

$$[HA]_{tot} = [A^-] + [HA],\ [A^-] = f_{A^-}[HA]_{tot},\ and\ [HA] = f_{HA}[HA]_{tot} \tag{B.18}$$

in which f_{HA} and f_{A^-} are the molar fractions of conjugate acid and base forms respectively. These fractions can be written in terms of $[H_3O^+]$ and K_a.

$$f_{A^-} = \frac{K_a}{K_a + [H_3O^+]}\ \ ,\ f_{HA} = \frac{[H_3O^+]}{K_a + [H_3O^+]} \tag{B.19}$$

Using the Eqs. B.18 and B.19, absorption from a mixture both forms (assuming unit path length) at a specific wavelength (λ) can be written as;

$$\begin{aligned} A(\lambda) &= \epsilon_{HA}(\lambda)[HA] + \epsilon_{A^-}(\lambda)[A^-] \\ &= \epsilon_{HA}(\lambda)\frac{[HA]_{tot}[H_3O^+]}{K_a + [H_3O^+]} + \epsilon_{A^-}(\lambda)\frac{[HA]_{tot}[K_a]}{K_a + [H_3O^+]} \\ &= \frac{[HA]_{tot}}{K_a + [H_3O^+]}(\epsilon_{HA}(\lambda)[H_3O^+] + \epsilon_{A^-}K_a) \end{aligned} \tag{B.20}$$

in which $\epsilon_{HA}(\lambda)$ and $\epsilon_{A^-}(\lambda)$ are extinction coefficients or equivalently absorption efficiencies of acidic and basic forms respectively at the specific wavelength (λ).

Clearly the Eq. B.20 is a function of total fluorophore concentration ($[HA]_{tot}$). This is not a desirable dependency particularly when the fluorophore concentration can not be controlled. To overcome the problem a ratio-metric method can be used so that rather than measuring absorption at single wavelength, two different wavelengths (λ_1 and λ_2) are used. The ratio of these two measurements is not concentration dependent. The functional form of the absorption ratios is given by;

$$\frac{A(\lambda_1)}{A(\lambda_2)} = \frac{\epsilon_{HA}(\lambda_1)[H_3O^+] + \epsilon_{A^-}(\lambda_1)K_a}{\epsilon_{HA}(\lambda_2)[H_3O^+] + \epsilon_{A^-}(\lambda_2)K_a} \tag{B.21}$$

By defining a news set of parameters and rewriting pH and pK_a, Eq. B.21 can be put into a simpler form.

$$\frac{A(\lambda_1)}{A(\lambda_2)} = \frac{\beta}{1 + 10^{\gamma - pH}} + \alpha \tag{B.22}$$

in which, $\alpha = \frac{\epsilon_{HA}(\lambda_1)}{\epsilon_{HA}(\lambda_2)}$, $\beta = \frac{\epsilon_{A^-}(\lambda_1)}{\epsilon_{A^-}(\lambda_2)}$, $\gamma = \log(\frac{\epsilon_{HA}(\lambda_2)}{\epsilon_{A^-}(\lambda_2)}) + pK_a$, $[H_3O^+] = 10^{-pH}$, and $K_a = 10^{-pK_a}$. If λ_2 is chosen to be isosbestic point of the fluorophore meaning that both acidic and basic forms have same absorption efficiencies, $\epsilon_{HA}(\lambda_{isosbestic}) = \epsilon_{A^-}(\lambda_{isosbestic})$, then fitting parameter γ is actually -pK$_a$. Once the fitting parameters are found, either of Eq. B.21 or B.22 can equivalently be used for pH sensing by feeding any of them with a measured absorption ratio [67, 31, 8, 12].

To apply ratio-metric method to BTB, we first did a series of calibration measurements. This was done by dissolving 1 μM dye in a buffer solution of 20 mM Tris and 200 mM NaCl at different pH. Absorption measurements were done using a UV/Vis spectrophotometer described in 2.2.3. Results of absorption measurements are shown in Fig. B.1.

Absorption maxima for the basic form of BTB was reported to be in the range of 613 nm to 619 nm [33, 88]. Also the isosbestic point of BTB was measured to be in the range of 489 nm to 508 nm [74]. Based on these information, we chose to use absorption values of BTB at 619 nm and 500 nm for pH sensing. Thus Eq. B.22 was fitted to measured absorption ratios A$_{619/500}$ found through the calibration experiments. Resultant sigmoid curve of fitting is shown in Fig. B.2.

Fitting parameters were found to be; $\alpha = 0.07$ (0.03), $\beta = 8.49$ (0.51), and $\gamma = 7.09$ (0.03). Here because of the choice of isosbestic point, pK$_a$ of the BTB was found to be 7.09 (0.03) which is in agreement with the reported literature value of 7.1 [88].

B.3 Ratiometric pH Measurements : Emission

The method presented in Sec. B.3 can equivalently be applied to emission spectra of a pH sensitive fluorophore as long as the excitation wavelength is set to be isosbestic point of the dye. By doing that, one can guarantee that the absorption efficiencies of

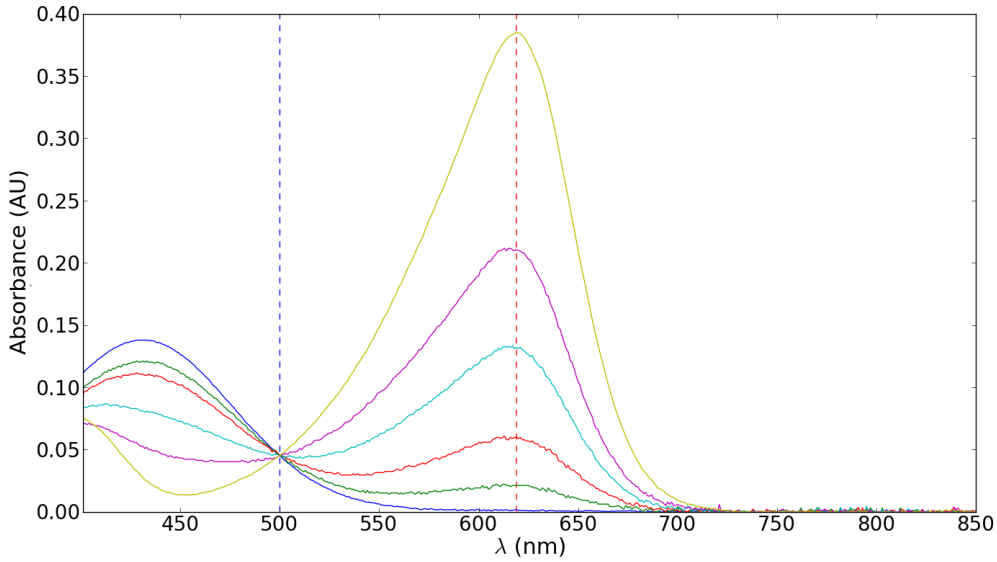

Figure B.1. BTB absorption in different pH buffers for calibration; 0.9 (blue), 5.8 (green), 6.3 (red), 6.8 (cyan), 7.2 (magenta), and 13.6 (yellow). Dashed lines indicates the wavelengths used in ratio-metric calculations; 500 nm (blue) and 619 nm (red). Actual data can be found in the place provided in Table D.7.

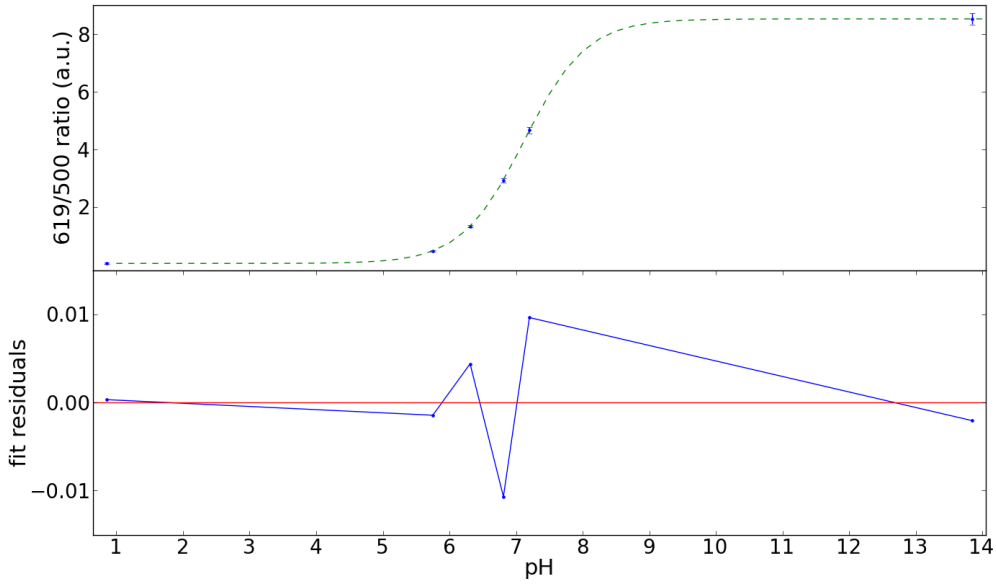

Figure B.2. $A_{619/500}$ ratios for BTB at different pHs and correspondign fit of sigmoid.

acidic and basic forms of the fluorophore are same. Thus the change in the emission ratio can only be a function of pH of the dye [66].

In emission experiments, we used Fluorescein as the pH sensitive dye. Similar to BTB, Fluorescein has an isosbestic point in its absorption spectra [31] at 460 nm and this was chosen to be the excitation wavelength. On the other hand, it doesn't have a isosbestic point in its emission spectra. Thus the fitting parameter γ in Eq. B.22, should not be understood as the -pK_a value of the dye.

Secondly, Fluorescein has three reported pK_a values of pK_{a1} = 2.1, pK_{a2} = 4.3, and pK_{a3} = 6.4 [155, 31]. In other words, depending on the ambient pH of the solvent the dissociation of the dye can produce its cation, neutral form, monoanion and dianion. Among which, dianion has the largest emission efficiency (equivalently relative quantum yield as a result of excitation at isosbestic point) of approximately 0.9. Next largest one is the anion of quantum yield approximately 0.4 [155]. Given that its neutral form also has a similar quantum yield to its monoanion, upon excitation at absorption isosbestic point, its emission spectra is not expected to change when the pH is lower than 5.3. This is why, Fluorescein is commonly considered as an indicator at near-neutral pH values meaning that it is an efficient indicator between approximately pH 5.3 and 7.4 [66].

It is also important to point out that Fluorescein doesn't have two distinct peaks corresponding to its monoanion and dianion forms, in its emission spectra. Rather, its monoanion has an emission maxima at 510 nm and a shoulder around 550 nm, and its dianion has an emission maxima at 515 nm [31]. Therefore we preferred to use the emission ratios of 514 nm and 550 nm for pH sensing.

Similar to calibration of BTB for Fluorescein, we first did a series of calibration measurements. This was done by dissolving 10 nM dye in a buffer solution of 20 mM Tris and 200 mM NaCl at different pH. Emission measurements were done using a

UV/Vis spectrophotometer described in 2.2.4. Results of emission measurements are shown in Fig. B.3.

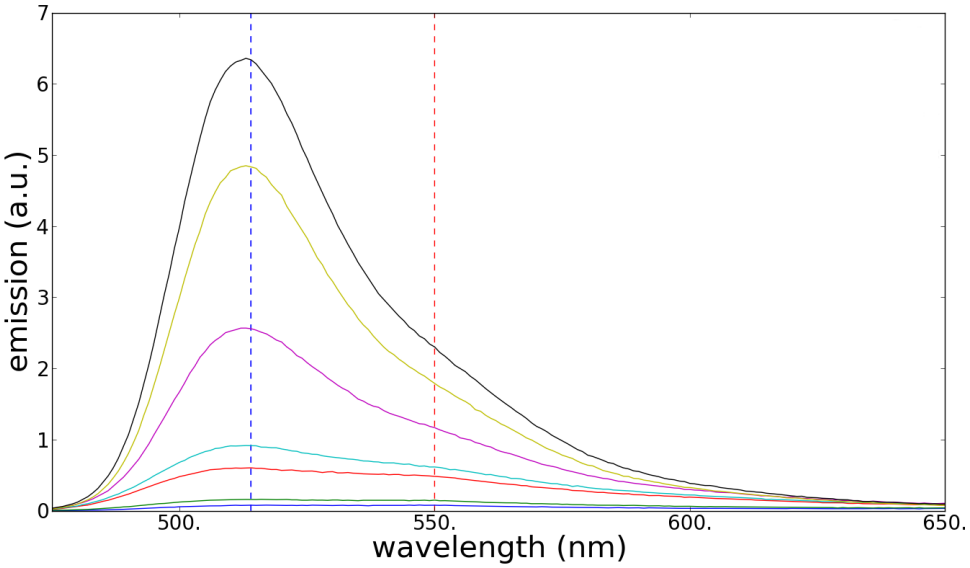

Figure B.3. Fluorescein emission in different pH buffers for calibration; 2.8 (blue), 3.7 (green), 4.6 (red), 5.4 (cyan), 6.3 (magenta), 7.2 (yellow), and 8.0 (black). Dashed lines indicates the wavelengths used in ratio-metric calculations; 514 nm (blue) and 550 nm (red). Actual data can be found in the place provided in Table D.8.

Eq. B.22 was then fitted to measured emission ratios $E_{514/550}$ found through the calibration experiments. Resultant sigmoid curve of fitting is shown in Fig. B.4.

Fitting parameters were found to be; $\alpha = 1.10$ (0.05), $\beta = 1.66$ (0.05), and $\gamma = 5.99$ (0.06). To further test our calibration, we used the reported pK_a value of Fluorescein between monoanion and dianion forms to recover ratio of emission efficiencies or equivalently quantum yields. Using the definition $\gamma = \log(\frac{\epsilon_{HA}(\lambda_2)}{\epsilon_{A^-}(\lambda_2)}) + pK_a$, and the exact values for quantum yields (0.93 for dianion, 0.37 for monoanion[155]), we found that the our calibration result for the ratio of quantum yields (0.39 (0.1)) was in agreement with the reported literature value (0.40).

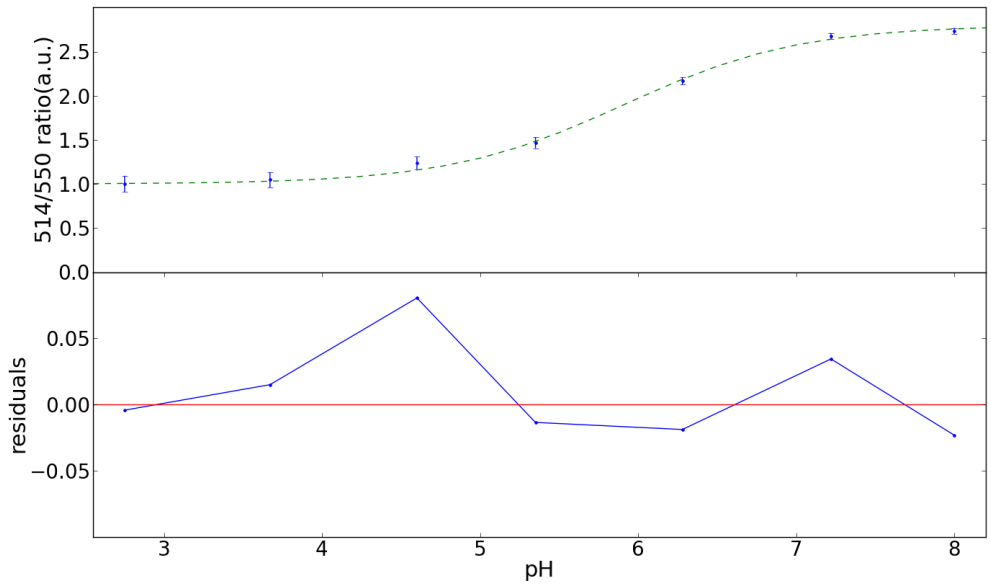

Figure B.4. $E_{514/550}$ ratios for Fluorescein at different pH values and corresponding fit of sigmoid.

APPENDIX C

SUPPLEMENTARY MATERIAL FOR SINGLE-MOLECULE STUDIES WITHIN SUB-MICRON SIZE DROPLETS

C.1 Confocal Imaging of Large Droplets

Confocal scanning images of large (micron) droplets show no evidence that the RNA used in this study is sequestered at the water-perfluorinate interface, as demonstrated in Fig. C.1. In this figure, the 16 base-pair duplex RNA labeled with Cy3 at a 5' termini (identical to that used for donor-only measurements in the text) was prepared at 16.7 μM in 20 mM Tris buffer with 200 nM NaCl . Droplets were created by adding 2 uL of RNA sample into 200 uL perfluorinated oil and surfactant solution as described earlier in the text. Sample is shaken 1 to 2 minutes , resulting in much larger droplets suitable for investigation by confocal scanning. Droplets were imaged at or very near to a glass boundary; the confocal image and corresponding line plot are centered at least one micron above a coverslip.

C.2 Burst Detection

A simple Bayesian model was constructed to separate photons originating from molecular fluorescence and those originating from background processes. Starting with the assumption of two Poisson processes, we assigned initial rates λ_{Burst} and λ_{BG} associated with each. Rather than considering directly the probability that the ith photon originates from either background or burst, we consider first that fast fluctuations between the two states are unphysical; a single "burst" photon between

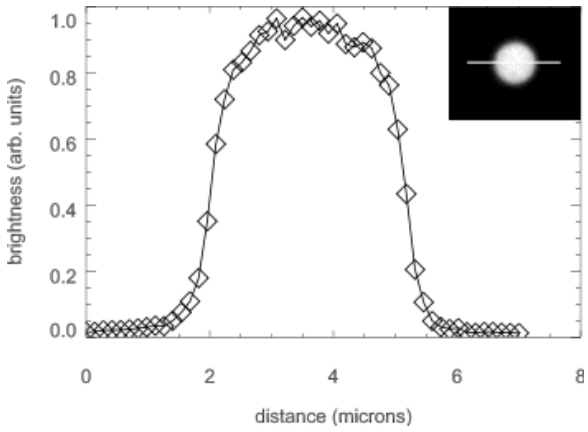

Figure C.1. Confocal scanning image of very large droplet in FC-40, as described above.

long stretches of background photons, and the opposite, should be avoided. We therefore consider the probability that $2N + 1$ sequential photons all originate from a burst:

$$P = \prod_{j=i-N}^{i+N} P(\tau_j|\lambda_{\text{Burst}}). \tag{C.1}$$

We compare this with the probability that the same photons originate from background:

$$Q = \prod_{j=i-N}^{i+N} P(\tau_j|\lambda_{\text{BG}}). \tag{C.2}$$

Defining

$$R = \frac{w_{\text{Burst}}P}{w_{\text{BG}}P + w_{\text{Burst}}Q}, \tag{C.3}$$

the ith photon is assigned to a burst if $S < R$, where S is a random number uniformly distributed on the interval 0 to 1. The weights w_{Burst} and w_{BG} are initially set equal to 1, and after the first iteration are calculated from the sample. This process converges by approximately 20 iterations for most data sets.

C.3 Photon Counting Histogram Analysis

In this work we have applied the photon counting histogram (PCH) method originally developed by Chen *et al.*[21] and later updated by Huang *et al.*[76]. This method provides a purely statistical analysis to find the molecular brightness and average number of molecules in the excitation volume for multiple species in a fluorescence fluctuation spectroscopy experiment. The data are binned in short time intervals and the resulting histogram of the number of photons in a bin is fit to a super-poissonian distribution. The model assumes an observation volume profile and for a one-photon excitation experiment a three dimensional gaussian is usually used. An extension of this introduces fitting parameters that describe the deviation of the detection volume from that of a three dimensional Gaussian and then taking a series expansion of the model to apply it. This is the approach developed by Huang *et al.* [76] and used to produce the fits in Fig. 2 of the text.

The data were modeled using a cylindrically symmetric three-dimensional Guassian with a waist of 260 nm and aspect ratio of 9:1, as suggested by a calibration of the instrument using FCS. We choose the arbitrary parameter[76] $Q = 6$ and used a bin time of 200 μs. The model has 2 fitting parameters, brightness, ϵ and molecular concentration, $\langle n \rangle$, for each species. All of the fits used a second order correction to the Gaussian detection volume, thereby adding two more fitting parameters F_1 and F_2. The parameter F_1, called the out-of-focus emission ratio, gives the ratio of the photons detected in the non-Gaussian part of the beam to the Gaussian part. When F_1 is large it becomes necessary to use a second parameter F_2 which increases the probability that a molecule in the non-gaussian part of the beam contributes two photons instead of just one.

When fitting the droplet and pH 4 solution data it was found that a fit with one or two species did not work well. Such fits resulted in a large χ^2, and/or non-random residuals and/or very large standard errors on some of the fit parameters. A model

sample	$\langle n_1 \rangle$	ϵ_1	$\langle n_2 \rangle$	ϵ_2	$\langle n_3 \rangle$	ϵ_3	F_1	F_2
FC40	0.0409(5)	2.45(37)	0.0189(12)	9.19(55)	0.0035(14)	15.2(1.1)	1.27(7)	0.012(3)
FC770	0.0532(21)	3.32(29)	0.0217(21)	8.41(37)	0.0013(3)	17.9(1.0)	1.45(4)	0.022(2)
pH 4	0.083(12)	1.11(69)	0.056(28)	3.24(62)	0.0064(13)	9.22(48)	1.38(fixed)	0.0258(5)
pH 7	0.18(11)	2.7(7)	0.0055(18)	6.5(5)	–	–	1.38(4)	0.048(4)

Table C.1. PCH fit parameters of the data in Fig. 3.2. Uncertainties are given in parentheses and represent the error on the last digits.

with three species gave values of χ^2 per degree of freedom near one in all cases. The pH 7 solution data were fit to two species. Fits to one species were unsatisfactory with large χ^2 or residuals, and attempts to add a species gave meaningless results for the third species. We chose to fix the parameter F_1 for pH 4 solution data to the same value found for the pH 7 data since this is a shape parameter that should be nominally the same for all solution data; small differences in alignment or index of refraction at low pH might account for the small change in F_2 that was required for a good fit. For droplet data we found it necessary to let the shape parameters vary to obtain good fits: It is reasonable to assume that the different indices of the oils and presence of the droplet may alter the shape of the detection volume. The final values of the shape parameters were only slightly different from those found in solution.

C.4 Proximity Histograms and Fit results

Proximity ratio histograms were fit to the probability density function of up to three beta distributions, representing the donor-only peak and up to two distinct FRET peaks:

$$P(x|\{A_i, \alpha_i, \beta_i\}) = \sum_{i=1}^{2\,or\,3} A_i \frac{x^{\alpha_i - 1}(1-x)^{\beta_i - 1}}{\mathrm{B}(\alpha_i, \beta_i)}. \tag{C.4}$$

The normalization constant is the beta function $\mathrm{B}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$. Each data bin was assigned to a specific peak i using Gibbs sampling. The donor-only peak was removed from the figures to simplify the comparison between the FRET peaks: the remaining amplitudes A_i were renormalized to exclude the donor-only peak.

For completeness, we include here the proximity ratio histograms of RNA confined to droplets in FC-77, evaluated at three different thresholds N_{th}, and taken under conditions otherwise identical to Fig. 3 and Fig. 5(c). The data in Fig. C.2 are distinct from those of Fig. 5(c) in the text; in particular the peak amplitudes are different, and the higher FRET peak also appears to be slightly shifted.

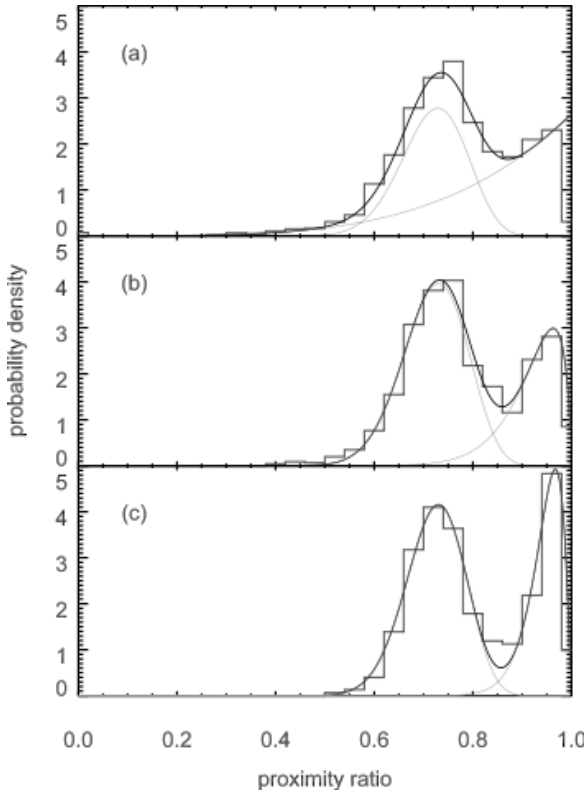

Figure C.2. Proximity histograms from single RNA molecules confined to freely-diffusing aqueous droplets in FC-77. Photon bin time is 2 ms. The three panels represent the same data but with different thresholds for inclusion in the histogram: (a) $N_{th} > 25$, (b) $N_{th} > 50$, and (c) $N_{th} > 75$. The data are fit with beta functions; fit parameters are given in Table C.2.

threshold	A	α	β	$\langle P \rangle$	σ_f	σ_s	bins	$\langle N \rangle$
25	0.471	32.60	11.91	0.732	0.066	0.0673	5373	53.99
	0.529	5.08	1.39	0.786	0.150	0.0633	5911	54.98
50	0.548	42.39	15.19	0.736	0.058	0.0518	2603	78.96
	0.452	5.13	1.18	0.813	0.144	0.0457	2099	87.58
75	0.615	45.69	16.43	0.736	0.056	0.0447	1332	101.73
	0.385	6.37	1.00	0.864	0.118	0.0365	865	117.62

Table C.2. Fit parameters for data of Fig. C.2.

threshold	A	α	β	$\langle P \rangle$	σ_f	σ_s	bins	$\langle N \rangle$
25	0.357	30.80	10.55	0.745	0.067	0.0659	1037	59.49
	0.643	4.17	1.47	0.739	0.170	0.0676	1805	56.19
50	0.496	31.25	10.52	0.748	0.066	0.0485	589	93.03
	0.504	5.93	1.62	0.786	0.140	0.0472	579	87.78
75	0.559	38.72	12.36	0.758	0.059	0.0412	335	117.21
	0.441	7.79	1.86	0.808	0.121	0.0379	287	116.85

Table C.3. Fit parameters for the data of Fig. 3.3 in the text.

pH	threshold	A	α	β	$\langle P \rangle$	σ_f	σ_s	bins	$\langle N \rangle$
7	25	0.460	10.92	4.65	0.701	0.112	0.0855	622	29.23
		0.540	16.45	10.73	0.605	0.092	0.0922	686	28.70
6	25	1.000	7.44	2.43	0.754	0.131	0.0745	2379	36.03
4	25	1.000	7.58	2.64	0.741	0.131	0.0759	1272	35.71

Table C.4. Fit parameters for the data of Fig. 3.4 in the text.

threshold	A	α	β	$\langle P \rangle$	σ_f	σ_s	bins	$\langle N \rangle$
25	0.780	7.68	2.79	0.734	0.130	0.0764	999	35.80
	0.220	7.76	2.32	0.770	0.126	0.0731	268	35.48
75	0.635	37.29	12.25	0.753	0.061	0.0417	381	116.33
	0.365	9.62	1.89	0.836	0.105	0.0371	246	115.13
75	0.767	35.86	12.76	0.738	0.062	0.0447	1657	105.97
	0.233	18.79	1.49	0.927	0.057	0.0260	527	112.48

Table C.5. Fit parameters for the data of Fig. 3.5 in the text.

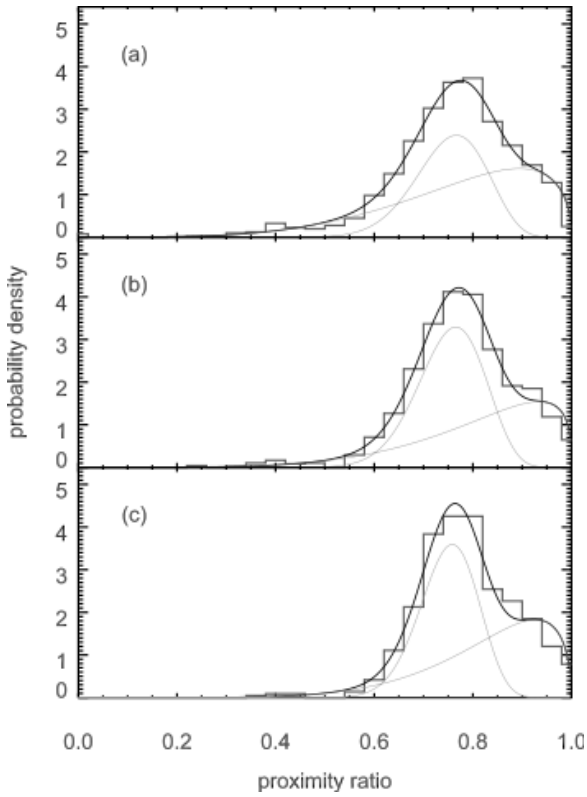

Figure C.3. Proximity histograms from single RNA molecules confined to freely-diffusing aqueous droplets in FC-40. Here 20 mM NaOH has been added just before droplet formation. Photon bin time is 2 ms. The three panels represent the same data but with different thresholds for inclusion in the histogram: (a) $N_{th} > 25$, (b) $N_{th} > 50$, and (c) $N_{th} > 75$. The data are fit with beta functions; fit parameters in Table C.6.

threshold	A	α	β	$\langle P \rangle$	σ_f	σ_s	bins	$\langle N \rangle$
25	0.432	26.48	8.75	0.752	0.072	0.0663	1377	52.66
	0.568	4.93	1.46	0.771	0.154	0.0653	1703	51.43
50	0.560	29.84	9.84	0.752	0.068	0.0506	692	79.35
	0.440	5.78	1.33	0.812	0.137	0.0458	507	78.88
75	0.469	7.19	1.49	0.829	0.121	0.0380	274	102.41
	0.531	40.06	13.45	0.749	0.059	0.0441	268	101.10

Table C.6. Fit parameters for the data shown in Fig. C.3.

C.5 Derivation of Analytical Auto-correlation Function Model for FCS

For a single fluorophore freely diffusing in solution with a known diffusivity (D), an analytical expression of its auto-correlation function (ACF) can be derived by assuming a Gaussian fluorescent detection profile [5]. Then the detected emission intensity profile is given by; [50]

$$I(\mathbf{r},t) = EQI_0 \exp(-\frac{2x^2+2y^2}{w_0^2}) \exp(-\frac{2z^2}{\omega^2 w_0^2}) c(\mathbf{r},t). \tag{C.5}$$

where E, Q, I_0 are the collection efficiency of the optical system, quantum yield of the fluorophore, and maximum excitation intensity in the focal plane correspondingly. w_0 is the radius of the $1/e^2$ contour of detection profile in lateral direction and ω is the aspect ratio or asymmetry parameter for the emission profile in axial direction. Finally $c(\mathbf{r},t)$ is the fluorophore concentration in the system.

Considering the fluctuations in the detection intensity so that $I(\mathbf{r},t) = \langle I(\mathbf{r},t)\rangle + \delta I(\mathbf{r},t)$, normalized correlation function is given by;

$$G(\tau) = \frac{\langle I(\mathbf{r}',t+\tau) I(\mathbf{r},t)\rangle}{\langle I(\mathbf{r},t)\rangle^2} = \frac{\langle \delta I(\mathbf{r},0)\delta I(\mathbf{r}',t)\rangle + \langle I\rangle^2}{\langle I\rangle^2} \tag{C.6}$$

In a typical FCS experiment excitation and emission optics are kept fixed and so the detection volume doesn't have time dependence. Thus the solution of the Eq. C.6 only requires the *a priori* knowledge of time and space dependence of fluorescent molecule concentration, $c(\mathbf{r},t)$. In other words, one needs to find the Green function for the system which eventually describes the relation between spatial coordinates and the time [41]. Then the δ terms in Eq. C.6 can be rewritten as;

$$\langle \delta I(\mathbf{r},0)\delta I(\mathbf{r}^{'},t)\rangle = \iiint_{-\infty}^{\infty} d\mathbf{r} \iiint_{-\infty}^{\infty} d\mathbf{r}^{'} (EQI_0)^2 \exp(-\frac{2(x-x^{'})^2+2(y-y^{'})^2}{w_0^2}) \times \exp(-\frac{2(z-z^{'})^2}{\omega^2 w_0^2}) \times g(\mathbf{r},\mathbf{r}^{'};t,t^{'}). \quad \text{(C.7)}$$

in which $g(\mathbf{r},\mathbf{r}^{'};t,t^{'})$ is the Green function which should be found by solving the diffusion equation which is given by;

$$\frac{\partial}{\partial t}\delta c(\mathbf{r},t) = D\nabla^2 \delta c(\mathbf{r},t) \quad \text{(C.8)}$$

For a collection freely diffusing fluorophores, Eq. C.8 can be considered as an initial value problem (IVP) and its solution is given by;

$$g(\mathbf{r},\mathbf{r}^{'};t,t^{'}) = \frac{\langle c\rangle}{[4D(t^{'}-t)]^{3/2}} \exp[-\frac{(\mathbf{r}^{'}-\mathbf{r})^2}{4D(t^{'}-t)}] \quad \text{(C.9)}$$

Gaussian form in Eq. C.9 can be simplified by changing time variables as $t^{'}-t \equiv t$. Solution of Eq. C.7 after changing time variables and plugging in the new form of Eq. C.9 [42], is given by Eq. 3.5 in the main text.

Although the procedure mentioned above requires the consideration of Eq. C.8 as an IVP, it can equivalently be considered as a boundary value problem (BVP). For the solution of BVP, one should treat the boundaries as they are at infinity. In BVP procedure, solution of diffusion equation turns into a separable form which allows the application of separation of variables technique on spatial part of Eq. C.8. Because the problem is actually a second order differential equation in spatial coordinates, BVP method provides Fourier series expansion of the solution which may consists of both real and imaginary exponential forms.

Aqueous droplets in our experiments can be considered spherical boundaries for the diffusion of fluorescent molecules within them. When there is a spherical boundary, Eq. C.8 can be re-written in spherical coordinates which provides periodic solutions in azimuthal and polar directions by the symmetry [16, 173]. After elimination

of angular components using the symmetry, the remaining radial part of the differential equation can be further reduced by separating the time component of the solution from space component $c(r,t) = c(r)\exp(-\lambda t)$ and it is given by;

$$-\lambda c(r) = D\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial c(r)}{\partial r}) \tag{C.10}$$

After some mathematical manipulations, one can find a Fourier series representation of the solution of Eq. C.10. The series solution without specifying the constants, is in following form;

$$c(r,t) \propto \sum_{n=1}^{\infty} a_n \frac{\sin(k_n r)}{r}\exp(-\lambda t) \tag{C.11}$$

Plugging Eq. C.11 into Eq. C.7 produces an integral equation which can not be solved by standard methods, as oppose to free diffusion case. This can easily be recognized by considering the solution in radial direction which does not allow one to calculate independent integrals in spatial coordinates.

Consequently, there is no available analytical form representing the ACF of diffusing fluorophores within aqueous droplets regardless of whether droplet are stationary in space (optically trapped) or not. To overcome this problem, we simulated the droplet system. We used the ACF we found from our simulations as numerical functional form. Then by changing the simulation parameters we fit this numeric model to experimental data. Result of this fitting procedure is shown in main text (Fig. 3.7). Location of all the data from simulation results as well as the codes used in the simulations are provided in Table D.10.

Here the simulation data did not only served us as a numeric functional form for ACF of droplet FCS data, it also provided us information on the source of multiple correlation times observed in experimental ACF. More specifically, by simulating the droplet FCS for case of single size droplets, we observed that the short correlation time in FCS can be attributed to diffusion of the molecule within a droplet whereas

the long one most likely corresponds to the diffusion of droplet itself. Also, in the case of a distribution of droplet sizes this long time scale is dictated by the most frequently observed droplet sizes or equivalently the diffusion time corresponding to the Mode of the size distribution (data not shown). For the completeness of the story, a comparison of ACFs from different but unique size droplets are shown in Fig. C.4.

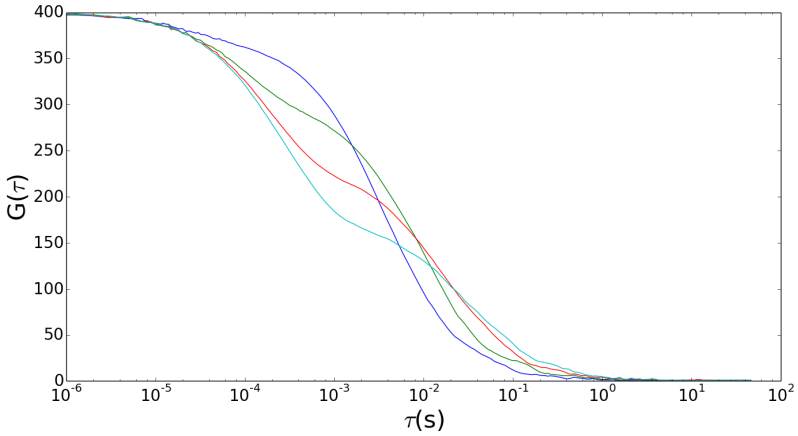

Figure C.4. Scaled synthetic ACFs for different droplet sizes; r=40 nm (blue), r=80 nm (green), r=120 nm (red), r=160 nm (cyan).

APPENDIX D

SUPPLEMENTARY MATERIAL FOR USED DATA IN FIGURES AND TABLES

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_3_3/	
Name in Text	**File Name**	**Date**
Sample 1	sample1.txt	June 12, 2013
Sample 2	sample3.txt	June 12, 2013
Sample 3	sample6.txt	June 12, 2013
Sample 4	FC40_pH75_200NaCl_1.txt	June 21, 2013

Table D.1. Storage place for raw data used in Fig. 2.3 and in Table 2.1 of main text. Here the data files consist of three columns. First column is showing the droplet sizes. The second is showing the percent intensities of each size. The last column is for the cumulative intensities for each size.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_3_4/	
Name in Text	**File Name (Sample Index)**	**Date**
Sample 1	Empty_200mM_Droplets_New.txt (2, 3)	June 1, 2013
Sample 2	FC40.txt (23, 24)	August 21, 2013

Table D.2. Storage place for raw data used in Fig. 2.4 and in Table 2.2 of main text. Data structure in the files are explained in Sec. D.1.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_3_6/	
Name in Text	**File Name (Sample Index)**	**Date**
t (min): 0	Empty_200mM_Droplets_New.txt (2, 3)	June 1, 2013
t (min): 10	Empty_200mM_Droplets_New.txt (8, 9)	June 1, 2013
t (min): 20	Empty_200mM_Droplets_New.txt (20, 21)	June 1, 2013
t (min): 30	Empty_200mM_Droplets_New.txt (26, 27)	June 1, 2013

Table D.3. Storage place for raw data used in Fig. 2.6 and in Table 2.3 of main text. Data structure in the files are explained in Sec. D.1.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_3_7/	
Name in Text	**File Name (Sample Index)**	**Date**
Sample 1	FC40.txt (1, 2, 3)	August 30, 2013
50 mM NaOH	FC40.txt (12, 13, 14)	August 30, 2013
100 mM NaOH	FC40.txt (9, 10, 11)	August 30, 2013

Table D.4. Storage place for raw data used in Fig. 2.7 (Sample 1) of main text. Storage place for other BTB data here is provided for consistency. Data structure in the files are explained in Sec. D.1.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_3_8/	
Name in Text	**File Name**	**Date**
0 mM NaOH	Droplets_10uM_0mMNaOH_Corr1.dat	November 6, 2013
1 mM NaOH	Droplets_10uM_1mMNaOH_Corr1.dat	November 6, 2013
5 mM NaOH	Droplets_10uM_5mMNaOH_Corr1.dat	November 6, 2013
20 mM NaOH	Droplets_10uM_20mMNaOH_FC40_Corr1.dat	November 6, 2013
50 mM NaOH	Droplets_10uM_50mMNaOH_Corr1.dat	November 6, 2013
100 mM NaOH	Droplets_10uM_1000mMNaOH_Corr1.dat	November 6, 2013

Table D.5. Storage place for raw data used in Fig. 2.8 and in Table 2.4 of main text.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Reference_Data/Q_ext/	
Name in Text	**File Name**	**Date**
Q_ext droplet reference in FC40	fc40_reference_data.txt	May 25, 2013
Q_ext droplet reference in FC77	fc77_reference_data.txt	May 25, 2013
Q_ext bead reference in water	polystyrene_reference_data.txt	May 25, 2013

Table D.6. Storage place for reference data lookup table used Mie scattering calculations.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Reference_Data/BTB_Calibration/	
Name in Text	**File Name (Sample Index)**	**Date**
pH 5.8	Bromothymol_pH_Series.txt (1)	January 29, 2013
pH 6.3	Bromothymol_pH_Series.txt (2)	January 29, 2013
pH 6.8	Bromothymol_pH_Series.txt (3)	January 29, 2013
pH 7.2	Bromothymol_pH_Series.txt (4)	January 29, 2013
pH 0.9	Indicators_Low_High_pH.txt (1)	August 31, 2013
pH 13.6	Indicators_Low_High_pH.txt (2)	August 31, 2013

Table D.7. Storage place for BTB calibration data. Data structure in the files are explained in Sec. D.1.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Reference_Data/Fluorescein_Calibration/	
Name in Text	**File Name**	**Date**
pH 2.8	pH2_10nM_Corr1.dat	November 2, 2013
pH 3.7	pH3_10nM_Corr1.dat	November 2, 2013
pH 4.6	pH4_10nM_Corr1.dat	November 2, 2013
pH 5.4	pH5_10nM_Corr1.dat	November 2, 2013
pH 6.3	pH6_10nM_Corr1.dat	November 2, 2013
pH 7.2	pH7_10nM_Corr1.dat	November 2, 2013
pH 8.0	pH78_10nM_Corr1.dat	November 2, 2013

Table D.8. Storage place for Fluorescein calibration data.

Computer :	goldnerlab.physics.umass.edu	
Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_on_Instruments/Peker/	
Experiment Type	**Folder Name**	**Original File Location**
DLS	DLS	Conte Building, Room B 564
DLS	DLS_Malvern	Goessmann Lab, Room 144
UV_Vis Absorption	UV_Vis	Hasbrouck Lab, Room 310
Fluorescent Emission	Fluorometer	Hasbrouck Lab, Room 314

Table D.9. Storage place for remaining data taken on droplets, dyes, and other chemicals used in text. Here all the data has two copies one on the machine next to instrument and the other on the specified location in Table. Only exception is the DLS data taken on Malvern instrument. There is only one copy of this data and it is placed to specified location in Table.

Computer :	goldnerlab.physics.umass.edu
Simulation Data Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_4_7/Simulation_Data/mu_4_88_sigma_0_2/
Experiment Data Location :	/home/lab/data/peker/Peker_Thesis_Data/Data_in_Figures_and_Tables/Figure_4_7/Experimental_Data/
Rest of Simulations Data Location :	/home/lab/data/peker/Peker_Thesis_Data/Simulation_Data/Constrained_Diffusion/

Table D.10. Storage place for simulation data shown in Fig. 3.7 and the rest of the simulations.

D.1 Data File Structures

Contents of a typical absorbance experiment data file is shown in Code D.1. For each experiment, the first line (here line 1) has the information about experiment date and company name. Every measurement is marked with a 3 columns of identification information (here lines 4 and 23). These are the type of the measurement (" Wavelength Scanning"), date of the measurement (as " 1/6/2013"), and the time of the measurement (as "19:01" and "19:26" for different measurements). Following the identification information, information for starting wavelength (here lines 6 and 25), end wavelength (here lines 7 and 26), and the measurement speed is provided (here lines 8 and 27). Then, measurement name (as **Sample number**) is given (here lines 10 and 29). Similarly, the data indices provided in Tables D.2, D.3, D.4, and D.7 all corresponds to these **Sample numbers**. Moreover, the data provided with multiple indices/Sample numbers in these tables should be understood as the average data out of given indices data. For example, in Table D.2, Sample 1 has two Sample indices as 2 and 3. This means, this data is actually the average of Sample 2 and Sample 3 in Empty_200mM_Droplets_New.txt. It is also important to note that before each experiment first the background measurements were done. But this information was directly used by the instrument and not kept in data files.

Code D.1. Example absorbance data

```
1  1-Jun-13        Biochrom Ltd.                                   20:05:22
2
3  0
4  Wavelength Scanning             1/6/2013        19:01
5
6  Start wavelength        400
7  End wavelength  850
8  Speed           Medium
9
10 Sample  1
11 Wavelength      Absorbance
12 400     -0.052
13 401     -0.052
14 402     -0.052
15 403     -0.052
16 404     -0.052
17 405     -0.051
18 406     -0.052
19 ...     ...
20 849     -0.038
21 850     -0.039
22 0
23 Wavelength Scanning             1/6/2013        19:26
24
25 Start wavelength        400
26 End wavelength  850
27 Speed           Medium
28
29 Sample  2
```

```
Wavelength      Absorbance
400     0.46
401     0.457
402     0.455
403     0.453
404     0.45
405     0.447
406     0.445
...     ...
849     0.079
850     0.077
0
```

D.2 A Mini "HowTo" for Calculation of Q_{ext} Values

A compiled and ready to use version of Mie scattering calculation code is provided on the PC *goldnerlab.physics.umass.edu*, in the location given in Table D.6. An executable file as well as the remaining necessary files for calculations is stored in the folder Mie-2-3-3.

To create Q_{ext} values for a collection of spherical particles, the given Python code fragment D.2 can be run as following; first the code needs to be saved as a Python script within the folder which has the executable file mentioned above. Then, the material properties for the system of interest should be provided to Python code. An example for this is shown between the lines 23-36 in example Python code (Code. D.2). Finally, after turning on the line 65, the script can be run in a shell environment using the command "nohup /usr/bin/python pythonscriptname.py &". Here the "nohup" and "&" are necessary for running the script in background without any interruption, because (despite it does multi-threading) calculation is intense and takes a long time to finish (up to a day).

Code D.2. Python code for generating Q_{ext} values

```
# -*- coding: utf-8 -*-
import numpy as np
import matplotlib.pyplot as plt
from scipy import *
from numpy import *
from scipy import stats
from scipy.optimize import leastsq, rosen, rosen_der, fmin_l_bfgs_b
from photon_tools import timetag_parse as ttp
from scipy import linalg as LA
from datetime import datetime
from scipy import special as spec
import sys, string, os
import subprocess
from scipy.interpolate import interp1d
import lmfit
import multiprocessing.pool
import tempfile
from scipy import interpolate
import commands
```

```
20
21 np.seterr(all='ignore')
22
23 index_continuous = 1.29
24 index_dipersed = 1.33
25 names = np.arange(400.,851.)
26 lambd = names / 1000.
27 num_wavelenghts = 450
28 tot_water = 1. #in uL
29 start_lambda = 400.
30 stop_lambda = 850.
31 executable = "./mie"
32 radius_0 = 50.
33 r_increment = 5
34 num_weights = 80
35 radius = np.linspace(10.,1000.,991)
36 radius_um = radius / 1000.
37
38
39 # Use below functions if you need to generate reference data !!!!!!!!
40 #-----------------------------------------------------------------------
41 pool = multiprocessing.pool.ThreadPool(16)
42 def compute_mie2(density, radius, lambd):
43     f = tempfile.NamedTemporaryFile()
44     v = locals()
45     v.update(globals())
46     full_command = '{exe} -d {dens} -m {index_c} \
47     -n {index_d} -r {radius} -l {lambd} -o {fname}'.format( \
48       exe=executable, dens=density, radius=radius, lambd=lambd, \
49       index_c=index_continuous, index_d=index_dipersed, fname=f.name)
50     os.system(full_command)
51
52     search = "grep 'Qext' %s | awk '{print $3}'" % f.name
53     Qext = float(subprocess.check_output(search, shell=True)) # Extinction efficiency
54     return (3./ (4.*radius)) * Qext # Absorption without density according to paper of
55                                     # M. D. Lechner
56
57 def compute_mie(density, radius, lambd):
58     Qext = np.zeros((len(radius),len(lambd)))
59     results = pool.map(lambda (i,j): ((i,j), compute_mie2(density, radius[i], lambd[j])),
60                        [(i,j) for i in range(len(radius)) for j in range(len(lambd))])
61     for ((i,j),v) in results:
62         Qext[i][j] = v
63     return Qext
64
65 #Qext = compute_mie(1.0, radius_um, lambd/1000.) # Turn this line on for reference data
66                                                  # calculation
```

BIBLIOGRAPHY

[1] Aitken, C. E.; Marshall, R. A.; Puglisi J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical Journal 94*, 5 (2008), 1826–35.

[2] Andrew, A. L. Correction of finite element eigenvalues for problems with natural or periodic boundary conditions. *BIT Numerical Mathematics 28*, 2 (1988), 254–269.

[3] Ångström, A. On the atmospheric transmission of sun radiation and on dust in the air. *Geografiska Annaler 11* (1929), 156–166.

[4] Anna, S. L.; Bontoux, N.; Stone H. A. Formation of dispersions using flow focusing in microchannels. *Applied Physics Letters 82*, 3 (2003), 364.

[5] Aragón, S. R.; Pecora R. Fluorescence correlation spectroscopy as a probe of molecular dynamics. *The Journal of Chemical Physics 64*, 4 (1976), 1791.

[6] Aramendia, P. F.; Negri, R. M.; Sanroman E. Temperature dependence of fluorescence and photoisomerization in symmetrical carbocyanines influence of medium viscosity and molecular structure. *Journal of Physical Chemistry 98*, 12 (1994), 3165–3173.

[7] Arayanarakool, R.; Shui, L.; Kengen S. W. M.; van den Berg A.; Eijkel J. C. T. Single-enzyme analysis in a droplet-based micro- and nanofluidic system. *Lab on a Chip 13*, 10 (2013), 1955–62.

[8] Badugu, R.; Lakowicz, J. R.; Geddes C. D. Wavelength-ratiometric near-physiological ph sensors based on 6-aminoquinolinium boronic acid probes. *Talanta 66*, 3 (2005), 569–74.

[9] Bayly, C. I.; Cieplak, P.; Cornell W.; Kollman P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *Journal of Physical Chemistry 97*, 40 (1993), 10269–10280.

[10] Beer, N. R.; Hindson, B. J.; Wheeler E. K.; Hall S. B.; Rose K. A.; Kennedy I. M.; Colston B. W. On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets. *Analytical Chemistry 79*, 22 (2007), 8471–8475.

[11] Best, R. B.; Merchant, K. A.; Gopich I. V.; Schuler B.; Bax A.; Eaton W. A. Effect of flexibility and cis residues in single-molecule fret studies of polyproline. *Proceedings of the National Academy of Sciences 104*, 48 (2007), 18964–18969.

[12] Bizzarri, R.; Arcangeli, C.; Arosio D.; Ricci F.; Faraci P.; Cardarelli F.; Beltram F. Development of a novel gfp-based ratiometric excitation and emission ph indicator for intracellular studies. *Biophysical Journal 90*, 9 (2006), 3300–14.

[13] Bloomfield, V. A.; Crothers, D. M.; Tinoco Jr. I. *Nucleic Acids: Structures, Properties, and Functions*. University Science Books, Sausalito, CA, 2000, ch. Bases, Nucleosides, and Nucleotides, pp. 13–43.

[14] Bohren, Craig F., and Huffman, Donald R., Eds. *Absorption and Scattering of Light by Small Particles*. Wiley-VCH Verlag GmbH, Weinheim, Germany, 1998.

[15] Boukobza, E.; Sonnenfeld, A.; Haran G. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *Journal of Physical Chemistry B 105*, 48 (2001), 12165–12170.

[16] Brindley, J.; Jivraj, N. A.; Merkin J. H.; Scott S. K. Stationary-state solutions for coupled reaction-diffusion and temperature-conduction equations i. infinite slab and cylinder with general boundary conditions. *Proceedings: Mathematical and Physical Sciences 430*, 1880 (1990), 459–477.

[17] Brunet, T.; Raffy, S.; Mascaro B.; Leng J.; Wunenburger R.; Mondain-Monval O.; Poncelet O.; Aristégui C. Sharp acoustic multipolar-resonances in highly monodisperse emulsions. *Applied Physics Letters 101*, 1 (2012), 011913.

[18] Buettner, A. V. Radiationless transitions in cyanine dyes. *Journal of Chemical Physics 46*, 4 (1967), 1398.

[19] Burada, P. S.; Hänggi, P.; Marchesoni F.; Schmid G.; Talkner P. Diffusion in confined geometries. *ChemPhysChem 10*, 1 (2009), 45–54.

[20] Case, D. A.; Darden T. A.; Cheatham, T. E. III; Simmerling-C. L.; Wang J.; Duke R. E.; Luo R.; Walker R. C.; Zhang W.; Merz K. M.; Roberts B. P.; Wang B.; Hayik S.; Roitberg A.; Seabra G.; Kolossvary I.; Wong K. F.; Paesani F.; Vanicek J.; Liu J.; Wu X.; Brozell S. R.; Steinbrecher T.; Gohlke H.; Cai Q.; Ye X.; Wang J.; Hsieh M. J.; Cui G.; Roe D. R.; Mathews D. H.; Seetin M. G.; Sagui C.; Babin V.; Luchko T.; Gusarov S.; Kovalenko A.; Kollman P. A. *AMBER 11*. University of California, San Francisco, 2010.

[21] Chen, Y.; Muller, J. D.; So P. T. C.; Gratton E. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophysical Journal 77*, 1 (1999), 553–567.

[22] Chen, L. J.; Hsu, M. C.; Lin S. T.; Yang S. Y. Salt effect on wetting/nonwetting behaviors. *Journal of Physical Chemistry 99*, 13 (1995), 4687–4697.

[23] Chung, H. S.; Louis, J. M.; Eaton W. A. Distinguishing between protein dynamics and dye photophysics in single-molecule fret experiments. *Biophysical Journal 98*, 4 (2010), 696–706.

[24] Cieplak, P.; Cornell, W. D.; Bayly C.; Kollman P. A. Application of the multimolecule and multiconformational resp methodology to biopolymers: Charge derivation for dna, rna, and proteins. *Journal of Computational Chemistry 16*, 11 (1995), 1357–1377.

[25] Clegg, R. M. Fluorescence resonance energy-transfer and nucleic-acids. *Methods in Enzymology 211* (1992), 353–388.

[26] Cooper, M.; Ebner, A.; Briggs M.; Burrows M.; Gardner N.; Richardson R.; West R. Cy3b: Improving the performance of cyanine dyes. *Journal of Fluorescence 14*, 2 (2004), 145–50.

[27] Cornell, W. D.; Cieplak, P.; Bayly C. I.; Gould I. R.; Merz K. M.; Ferguson D. M.; Spellmeyer D. C.; Fox T.; Caldwell J. W.; Kollman P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society 117*, 19 (1995), 5179–5197.

[28] Coupland, J. N.; McClements, D. J. Lipid oxidation in food emulsions. *Trends in Food Science & Technology 7*, 3 (1996), 83–91.

[29] Courtois, F.; Olguin, L. F.; Whyte G.; Bratton D.; Huck W. T. S.; Abell C.; Hollfelder F. An integrated device for monitoring time-dependent in vitro expression from single genes in picolitre droplets. *ChemBioChem 9*, 3 (2008), 439–46.

[30] Cox, A. J.; DeWeerd, A. J.; Linden J. An experiment to measure mie and rayleigh total scattering cross sections. *American Journal of Physics 70*, 6 (2002), 620.

[31] Crovetto, L.; Orte, A.; Talavera E. M.; Alvarez-Pez J. M.; Cotlet M.; Thielemans J.; De Schryver F. C.; Boens N. Global compartmental analysis of the excited-state reaction between fluorescein and (±)-n-acetyl aspartic acid. *Journal of Physical Chemistry B 108*, 19 (2004), 6082–6092.

[32] Culbertson, M. J.; Williams, J. T. B.; Cheng-W. W. L.; Stults D. A.; Wiebracht Emily R.; Kasianowicz J. J.; Burden D. L. Numerical fluorescence correlation spectroscopy for the analysis of molecular dynamics under nonstandard conditions. *Analytical Chemistry 79*, 11 (2007), 4031–9.

[33] Dias A. C. B.; Silva R. A. O.; Arruda, M. A. Z. A sequential injection system for indirect spectrophotometric determination of lactic acid in yogurt and fermented mash samples. *Microchemical Journal 96*, 1 (2010), 151–156.

[34] Dietrich, A.; Buschmann, V.; Muller C.; Sauer M. Fluorescence resonance energy transfer fret and competing processes in donor-acceptor substituted dna strands: a comparative study of ensemble and single-molecule data. *Reviews in Molecular Biotechnology 82*, 3 (2002), 211–231.

[35] Dittrich, P. S.; Jahnz, M.; Schwille P. A new embedded process for compartmentalized cell-free protein expression and on-line detection in microfluidic devices. *ChemBioChem 6*, 5 (2005), 811–4.

[36] Dix, J. A.; Hom, E. F. Y.; Verkman A. S. Fluorescence correlation spectroscopy simulations of photophysical phenomena and molecular interactions: a molecular dynamics/monte carlo approach. *The journal of physical chemistry. B 110*, 4 (2006), 1896–906.

[37] Djaković, L.; Dokić, P.; Radivojević P. Investigation on the particle size distribution of o/w emulsions. *Kolloid-Zeitschrift & Zeitschrift für Polymere 244*, 2 (1971), 324–332.

[38] Dupradeau, F. Y.; Pigache, A.; Zaffran T.; Savineau C.; Lelong R.; Grivel N.; Lelong D.; Rosanski W.; Cieplak P. The r.e.d. tools: advances in resp and esp charge derivation and force field library building. *Physical Chemistry Chemical Physics : PCCP 12*, 28 (2010), 7821–39.

[39] Dutta, P.; Bhattacharyya, K. Ultrafast chemistry in complex and confined systems. *Journal of Chemical Sciences 116*, 1 (2004), 5–16.

[40] Eck, T. F.; Holben, B. N.; Reid J. S.; Dubovik O.; Smirnov A.; O'Neill N. T.; Slutsker I.; Kinne S. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research 104*, D24 (1999), 31333.

[41] Elson, E. L.; Magde, D. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers 13*, 1 (1974), 1–27.

[42] Elson, E. L. Brief introduction to fluorescence correlation spectroscopy. In *Methods in enzymology*, 1 ed., vol. 518. Elsevier Inc., 2013, ch. Chapter 2, pp. 11–41.

[43] Erickson, D.; Li, D. Integrated microfluidic devices. *Analytica Chimica Acta 507*, 1 (2004), 11–26.

[44] Felderhof, B. U.; Sellier, A. Mobility matrix of a spherical particle translating and rotating in a viscous fluid confined in a spherical cell, and the rate of escape from the cell. *Journal of Chemical Physics 136*, 5 (2012), 054703.

[45] Fenn, E. E.; Wong, D. B.; Fayer M. D. Water dynamics at neutral and ionic interfaces. *Proceedings of the National Academy of Sciences 106*, 36 (2009), 15243–8.

[46] Fennema, O. R. *Food Chemistry, Third Edition*. Food Science and Technology - Marcel Dekker, Inc. Taylor & Francis, 1996.

[47] Ferry, M. S.; Razinkov, I. A.; Hasty J. *Microfluidics for synthetic biology: from design to execution.*, 1 ed., vol. 497. Elsevier Inc., 2011.

[48] Förster, T.; Sinanoglu, O. *Modern Quantum Chemistry*. Academic, New York, 1965, ch. Delocalized excitation and excitation transfer, pp. 93–160.

[49] Förster, T. Energiewanderung und fluoreszenz. *Naturwissenschaften 6* (1946), 166–175.

[50] Fradin, C.; Abu-Arish, A.; Granek R.; Elbaum M. Fluorescence correlation spectroscopy close to a fluctuating membrane. *Biophysical journal 84*, 3 (2003), 2005–20.

[51] Francis, M. J., and Pashley, R. M. A study of de-gassed oil in water dispersions as potential drug delivery systems. *Colloids and Surfaces A: Physicochemical and Engineering Aspects 260*, 1-3 (2005), 7–16.

[52] Francis, M. J.; Pashley, R. M. The effect of de-gassing on the dispersion of fine oil droplets in water. *Colloids and Surfaces A: Physicochemical and Engineering Aspects 287*, 1-3 (2006), 36–43.

[53] Gamari, B. D.; Zhang, D.; Buckman R. E.; Milas P.; Denker J. S.; Chen H.; Hongmin L.; Goldner L. S. Inexpensive electronics and software for photon statistics and correlation spectroscopy. *American Journal of Physics* (2013), in press.

[54] Goldner, L. S.; Jofre, A. M.; Tang J. Droplet confinement and fluorescence measurement of single molecules. *Methods in Enzymology 472*, 10 (2010), 61–88.

[55] Gopich, I. V.; Szabo, A. Theory of photon statistics in single-molecule förster resonance energy transfer. *Journal of Chemical Physics 122*, 1 (2005), 014707–014725.

[56] Gopich, I. V.; Szabo, A. Theory of single-molecule fret efficiency histograms. In *Single-molecule biophysics: Experiment and theory*, M.; Takahashi S.; Yang H.; Silbey R. J. Komatsuzaki, T.; Kawakami, Ed., vol. 146 of *Advances in Chemical Physics*. 2012, pp. 245–297.

[57] Gopich, I. V. private communication, 2012, 2012.

[58] Gopich, I. V., and Szabo, A. Single-molecule fret with diffusion and conformational dynamics. *Journal of Physical Chemistry B 111* (2007), 12925–12932.

[59] Graaff, R.; Aarnoudse, J. G.; Zijp J. R.; Sloot P. M.; de Mul F. F.; Greve J.; Koelink M. H. Reduced light-scattering properties for mixtures of spherical particles: a simple approximation derived from mie calculations. *Applied Optics 31*, 10 (1992), 1370–6.

[60] Granick, S. Motions and relaxations of confined liquids. *Science 253*, 5026 (1991), 1374–9.

[61] Granovsky, A. A. Firefly version 7.1.g.

[62] Grebenkov, D. Residence times and other functionals of reflected brownian motion. *Physical Review E 76*, 4 (2007), 041139.

[63] Grunwell, J. R.; Glass, J. L.; Lacoste T. D.; Deniz A. A.; Chemla D. S.; Schultz P. G. Monitoring the conformational fluctuations of dna hairpins using single-pair fluorescence resonance energy transfer. *Journal of the American Chemical Society 123*, 18 (2001), 4295–4303.

[64] Ha, T. J.; Ting, A. Y.; Liang J.; Caldwell W. B.; Deniz A. A.; Chemla D. S.; Schultz P. G.; Weiss S. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences 96*, 3 (1999), 893–898.

[65] Ha, T. J.; Zhuang, X. W.; Kim H. D.; Orr J. W.; Williamson J. R.; Chu S. Ligand-induced conformational changes observed in single rna molecules. *Proceedings of the National Academy of Sciences 96*, 16 (1999), 9077–9082.

[66] Han, J.; Burgess, K. Fluorescent indicators for intracellular ph. *Chemical Reviews 110*, 5 (2010), 2709–28.

[67] Hanson, G. T.; McAnaney, T. B.; Park E. S.; Rendell M. E. P.; Yarbrough D. K.; Chu S.; Xi L.; Boxer S. G.; Montrose M. H.; Remington S. J. Green fluorescent protein variants as ratiometric dual emission ph sensors. 1. structural characterization and preliminary application. *Biochemistry 41*, 52 (2002), 15477–88.

[68] Haratifar, S.; Bazinet, L.; Manoury N.; Britten M.; Angers P. Impact of redox potential electrochemical modification and storage conditions on the oxidation reaction prevention in dairy emulsion. *Dairy Science & Technology 91*, 5 (2011), 541–554.

[69] Harvey, B. J., and Levitus, M. Nucleobase-specific enhancement of cy3 fluorescence. *Journal of Fluorescence 19*, 3 (2009), 443–448.

[70] Harvey, B. J.; Perez, C.; Levitus M. Dna sequence-dependent enhancement of cy3 fluorescence. *Photochemical & Photobiological Sciences 8*, 8 (2009), 1105–10.

[71] Hoefling, M.; Lima, N.; Haenni D.; Seidel C. A. M.; Schuler B.; Grubmueller H. Structural heterogeneity and quantitative fret efficiency distributions of polyprolines through a hybrid atomistic simulation and monte carlo approach. *PLOS One 6*, 5 (2011), e19791.

[72] Holtze, C.; Rowat, A. C.; Agresti J. J.; Hutchison J. B.; Angilè F. E.; Schmitz C. H. J.; Köster S; Duan H.; Humphry K. J.; Scanga R. A.; Johnson J. S.; Pisignano D.; Weitz D. A. Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab on a Chip 8*, 10 (2008), 1632–9.

[73] Hornak, V.; Abel, R.; Okur A.; Strockbine B.; Roitberg A.; Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics 65* (2006), 712–725.

[74] Hoxter, G. Suggested isosbestic wavelength calibration in clinical analyses. *Clinical Chemistry 25*, 1 (1979), 143–6.

[75] Huang, D.; Ou, B.; Prior R. L. The chemistry behind antioxidant capacity assays. *Journal of Agricultural and Food Chemistry 53*, 6 (2005), 1841–56.

[76] Huang, B.; Perroud, T. D.; Zare R. N. Photon counting histogram: One-photon excitation. *ChemPhysChem 5*, 10 (2004), 1523–1531.

[77] Huebner, A.; Bratton, D.; Whyte G.; Yang M.; Demello A. J.; Abell C.; Hollfelder F. Static microdroplet arrays: a microfluidic device for droplet trapping, incubation and release for enzymatic and cell-based assays. *Lab on a Chip 9*, 5 (2009), 692–8.

[78] Iqbal, A.; Wang, L.; Thompson K. C.; Lilley D. M. J.; Norman D. G. The structure of cyanine 5 terminally attached to double-stranded dna: Implications for fret studies. *Biochemistry 47*, 30 (2008), 7857–7862.

[79] Iqbal, A.; Arslan, S.; Okumus B.; Wilson T. J.; Giraud G.; Norman D. G.; Ha T.; Lilley D. M. J. Orientation dependence in fluorescent energy transfer between cy3 and cy5 terminally attached to double-stranded nuclelic acids. *Proceedings of the National Academy of Sciences 105*, 32 (2008), 11176–11181.

[80] Ivanov, I. B.; Danov, K. D.; Kralchevsky P. A. Flocculation and coalescence of micron-size emulsion droplets. *Colloids and Surfaces A: Physicochemical and Engineering Aspects 152*, 1-2 (1999), 161–182.

[81] Jofre, A.; Tang, J. Y.; Greene M. E.; Lowman G. M.; Hodas N.; Kishore R. B.; Helmerson K.; Goldner L. S. Hydrosomes: Femtoliter containers for fluorescence spectroscopy studies. *Proceedings of SPIE - The International Society for Optical Engineering 6644* (2007), 66440E.

[82] Joo, C.; Balci, H.; Ishitsuka Y.; Buranachai C.; Ha T. J. Advances in single-molecule fluorescence methods for molecular biology. *Annual Review of Biochemistry 77* (2008), 51–76.

[83] Ju, L. K.; Lee, J. F.; Armiger W. B. Enhancing oxygen transfer in bioreactors by perfluorocarbon emulsions. *Biotechnology Progress 7*, 4 (1991), 323–329.

[84] Junge, C. E. *Air Chemistry and Radioactivity.* Academic Press Inc., 1963.

[85] Jungwirth, P.; Tobias, D. J. Specific ion effects at the air/water interface. *Chemical Reviews 106*, 4 (2006), 1259–81.

[86] Kalinin, S.; Sisamakis, E.; Magennis S. W.; Felekyan S.; Seidel C. A. M. On the origin of broadening of single-molecule fret efficiency distributions beyond shot noise limits. *Journal of Physical Chemistry B 114*, 18 (2010), 6197–6206.

[87] King, M. D.; Byrne, D. M. A method for inferring total ozone content from the spectral variation of total optical depth obtained with a solar radiometer. *Journal of the Atmospheric Sciences 33*, 11 (1976), 2242–2251.

[88] Klotz, E.; Doyle, R.; Gross E.; Mattson B. The equilibrium constant for bromothymol blue: A general chemistry laboratory experiment using spectroscopy. *Journal of Chemical Education 88*, 5 (2011), 637–639.

[89] Kottalam, J., and Case, D. A. Dynamics of ligand escape from the heme pocket of myoglobin. *Journal of the American Chemical Society 110*, 23 (1988), 7690–7697.

[90] Krafft, M. Emulsions and microemulsions with a fluorocarbon phase. *Current Opinion in Colloid & Interface Science 8*, 3 (2003), 251–258.

[91] Kunz, W.; Lo Nostro, P.; Ninham B. W. The present state of affairs with hofmeister effects. *Current Opinion in Colloid & Interface Science 9*, 1-2 (2004), 1–18.

[92] Lechner, M. D. Influence of mie scattering on nanoparticles with different particle sizes and shapes: Photometry and analytical ultracentrifugation with absorption optics. *Journal of the Serbian Chemical Society 70*, 3 (2005), 361–369.

[93] Lee, J. Y.; Okumus, B.; Kim D. S.; Ha T. J. Extreme conformational diversity in human telomeric dna. *Proceedings of the National Academy of Sciences 102*, 52 (2005), 18938–43.

[94] Lemke, E. A.; Gambin, Y.; Vandelinder V.; Brustad E. M.; Liu H. W.; Schultz P. G.; Groisman A.; Deniz A. A. Microfluidic device for single-molecule experiments with enhanced photostability. *Journal of the American Chemical Society 131*, 38 (2009), 13610–2.

[95] Lewandowski, G.; Meissner, E.; Milchert E. Special applications of fluorinated organic compounds. *Journal of Hazardous Materials 136*, 3 (2006), 385–91.

[96] Li, P. T. X.; Goldner, L. S. *RNA Nanotechnology*. Pan Stanford Publishing, in press, ch. Application of Single-Molecule Fluorescence in RNA Biology.

[97] Lilley, D. M J. *The structure and folding of branched RNA analyzed by fluorescence resonance energy transfer.*, 1 ed., vol. 469. Elsevier Inc., 2009.

[98] Lilley, D. M. J.; Wilson, T. J. Fluorescence resonance energy transfer as a structural tool for nucleic acids. *Current Opinion in Chemical Biology 4*, 5 (2000), 507–517.

[99] Lin, B.; Yu, J.; Rice S. Direct measurements of constrained brownian motion of an isolated sphere between two walls. *Physical Review E 62*, 3 Pt B (2000), 3909–19.

[100] Lowe, K. C. Perfluorochemical respiratory gas carriers: applications in medicine and biotechnology. *Science progress 80*, 2 (1997), 169–93.

[101] Lowe, K. C. Perfluorinated blood substitutes and artificial oxygen carriers. *Blood Reviews 13* (1999), 171–184.

[102] Lukic, B. *Brownian Motion Studied with an optical trap.* PhD thesis, École Polytechnique Fédérale De Lausanne, 2007.

[103] Luong, Trung Quan. *Terahertz and Infrared Spectroscopy of Confined Water.* Phd., Ruhr University, Bochum, 2012.

[104] Malicka, J.; Gryczynski, I.; Fang J.; Kusba J.; Lakowicz J. R. Photostability of cy3 and cy5-labeled dna in the presence of metallic silver particles. *Journal of Fluorescence 12*, 314 (2002), 439 –447.

[105] Marinova, K. G.; Alargova, R. G.; Denkov N. D.; Velev O. D.; Petsev D. N.; Ivanov I. B.; Borwankar R. P. Charging of oilwater interfaces due to spontaneous adsorption of hydroxyl ions. *Langmuir 12*, 8 (1996), 2045–2051.

[106] Mason, T. G.; Wilking, J. N.; Meleson K.; Chang C. B.; Graves S. M. Nanoemulsions: formation, structure, and physical properties. *Journal of Physics: Condensed Matter 18*, 41 (Oct. 2006), R635–R666.

[107] McClements, D. J. *Food Emulsions: Principles, Practice, and Techniques.* CRC Press, 2005.

[108] McDonald, J. C.; Duffy, D. C.; Anderson J. R.; Chiu D. T.; Wu H.; Schueller O. J.; Whitesides G. M. Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis 21*, 1 (2000), 27–40.

[109] Milas, Peker. personal communication.

[110] Milas, Peker, Gamari, Ben D, Parrot, Louis, Krueger, Brent P, Rahmanseresht, Sheema, Moore, James, and Goldner, Lori S. Indocyanine dyes approach free rotation at the 3' terminus of A-RNA: a comparison with the 5' terminus and consequences for fluorescence resonance energy transfer. *Journal of Physical Chemistry B 117*, 29 (2013), 8649–58.

[111] Milas, Peker, Rahmanseresht, Sheema, Ramos, Kieran P, Gamari, Ben D, and Goldner, Lori S. Single-Molecule-Sensitive FRET in Attoliter Droplets. *arXiv*, 1312.0854 (2013).

[112] Moilanen, D. E.; Fenn, E. E.; Wong D.; Fayer M. D. Geometry and nanolength scales versus interface interactions: water dynamics in aot lamellar structures and reverse micelles. *Journal of the American Chemical Society 131*, 23 (2009), 8318–28.

[113] Moilanen, D. E.; Fenn, E. E.; Wong D.; Fayer M. D. Water dynamics at the interface in aot reverse micelles. *Journal of Physical Chemistry B 113*, 25 (2009), 8560–8.

[114] Mourant, J. R.; Fuselier, T.; Boyer J.; Johnson T. M.; Bigio I. J. Predictions and measurements of scattering and absorption over broad wavelength ranges in tissue phantoms. *Applied Optics 36*, 4 (1997), 949–57.

[115] Muñoz Losa, A.; Curutchet, C.; Krueger B. P.; Hartsell L. R.; Mennucci B. Fretting about fret: Failure of the ideal dipole approximation. *Biophysical Journal 96*, 12 (2009), 4779–4788.

[116] Mudeme, S. *Drop formation and rupture in shearing during processing of highly concentrated emulsions.* Ms., Cape Peninsula University of Technology, Cape Town, 2009.

[117] Murphy, M. C.; Rasnik, I.; Cheng W.; Lohman T. M.; Ha T. J. Probing single-stranded dna conformational flexibility using fluorescence spectroscopy. *Biophysical Journal 86*, 4 (2004), 2530–2537.

[118] Myers, D. *Surfactant Science and Technology.* John Wiley & Sons, Inc., Hoboken, NJ, USA, 2005.

[119] Nettels, D.; Schuler, B. *Single-Molecule Biophysics*, vol. 146 of *Advances in Chemical Physics.* 2012, ch. Single-Molecule FRET of Protein-Folding Dynamics, pp. 230–48.

[120] Norman, D. G.; Grainger, R. J.; Uhrin D.; Lilley D. M. J. Location of cyanine-3 on double-stranded dna: Importance for fluorescence resonance energy transfer studies. *Biochemistry 39*, 21 (2000), 6317–6324.

[121] Novotny, L.; Hecht, B. *Principles of Nano-Optics.* Cambridge University Press, 2006.

[122] Ntwampe, S. K. O.; Williams, C. C.; Sheldon M. S. Water-immiscible dissolved oxygen carriers in combination with pluronic f 68 in bioreactors. *African Journal of Biotechnology 9*, 8 (2010), 1106–1114.

[123] Olson, W. K.; Bansal, M.; Burley S. K.; Dickerson R. E.; Gerstein M.; Harvey S. C.; Heinemann U.; Lu X. J.; Neidle S.; Shakked Z.; Sklenar H.; Suzuki M.; Tung C. S.; Westhof E.; Wolberger C.; Berman H. M. A standard reference frame for the description of nucleic acid base-pair geometry. *Journal of Molecular Biology 313*, 1 (2001), 229–237.

[124] Ouellet, J.; Schorr, S.; Iqbal A.; Wilson T. J.; Lilley D. M. J. Orientation of cyanine fluorophores terminally attached to dna via long, flexible tethers. *Biophysical Journal 101*, 5 (2011), 1148–1154.

[125] Pashley, R. M.; Francis, M. J.; Rzechowicz M. The hydrophobicity of non-aqueous liquids and their dispersion in water under degassed conditions. *Current Opinion in Colloid & Interface Science 13*, 4 (2008), 236–244.

[126] Pedregosa, F.; Varoquaux, G.; Gramfort A.; Michel V.; Thirion B.; Grisel O.; Blondel M.; Prettenhofer P.; Weiss R.; Dubourg V.; Vanderplas J.; Passos A.; Cournapeau D.; Brucher M.; Perrot M.; Duchesnay E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res. 12* (2011), 2825–2830.

[127] Pettersen, E. F.; Goddard, T. D.; Huang C. C.; Couch G. S.; Greenblatt D. M.; Meng E. C.; Ferrin T. E. Ucsf chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry 25*, 13 (2004), 1605–12.

[128] Piletic, I. R.; Moilanen, D. E.; Levinger N. E.; Fayer M. D. What nonlinear-ir experiments can tell you about water that the ir spectrum cannot. *Journal of the American Chemical Society 128*, 32 (2006), 10366–7.

[129] Piletic, I. R.; Moilanen, D. E.; Spry D. B.; Levinger N. E.; Fayer M. D. Testing the core/shell model of nanoconfined water in reverse micelles using linear and nonlinear ir spectroscopy. *Journal of Physical Chemistry A 110*, 15 (2006), 4985–99.

[130] Porras, M.; Solans, C.; González C.; Gutiérrez J.M. Properties of water-in-oil (w/o) nano-emulsions prepared by a low-energy emulsification method. *Colloids and Surfaces A: Physicochemical and Engineering Aspects 324*, 1-3 (2008), 181–188.

[131] Psathas, P. A.; Sander, E. A.; Ryoo W.; Mitchell D.; Lagow R. J.; Lim K. T.; Johnston K. P. Interfacial studies of the formation of microemulsions of water in carbon dioxide with fluorinated surfactants. *Journal of Dispersion Science and Technology 23*, 1-3 (2002), 81–92.

[132] Qian, H.; Elson, E. L. Analysis of confocal laser-microscope optics for 3-d fluorescence correlation spectroscopy. *Applied Optics 30*, 10 (1991), 1185–1195.

[133] Randolph, J. B.; Waggoner, A. S. Stability, specificity and fluorescence brightness of multiply-labeled fluorescent dna probes. *Nucleic Acids Research 25*, 14 (1997), 2923–2929.

[134] Randsoe, T.; Hyldegaard, O. Effect of oxygen breathing and perfluorocarbon emulsion treatment on air bubbles in adipose tissue during decompression sickness. *Journal of Applied Physiology 107*, 6 (2009), 1857–63.

[135] Rasnik, I.; Myong, S.; Cheng W.; Lohman T. M.; Ha T. J. Dna-binding orientation and domain conformation of the e-coli rep helicase monomer bound to a partial duplex junction: Single-molecule studies of fluorescently labeled enzymes. *Journal of Molecular Biology 336*, 2 (2004), 395–408.

[136] Reid, J. S.; Eck, T. F.; Christopher S. A.; Hobbs P. V.; Holben B. Use of the ångstrom exponent to estimate the variability of optical and physical properties of aging smoke particles in brazil. *Journal of Geophysical Research 104*, D22 (1999), 27473.

[137] Reiner, J. E.; Crawford, A. M.; Kishore R. B.; Goldner L. S.; Helmerson K.; Gilson M. K. Optically trapped aqueous droplets for single molecule studies. *Applied Physics Letters 89*, 1 (2006), 013904.

[138] Reynolds, R. A.; Stramski, D.; Wright V. M.; Woźniak S. B. Measurements and characterization of particle size distributions in coastal waters. *Journal of Geophysical Research 115*, C8 (2010), C08024.

[139] Rhoades, E.; Gussakovsky, E.; Haran G. Watching proteins fold one molecule at a time. *Proceedings of the National Academy of Sciences 100*, 6 (2003), 3197–202.

[140] Riess, J. G. Understanding the fundamentals of perfluorocarbons and perfluorocarbon emulsions relevant to in vivo oxygen delivery. *Artificial Cells, Blood Substitutes and Biotechnology 33*, 1 (2005), 47–63.

[141] Rigler, R.; Mets, U.; Widengren J.; Kask P. Fluorescence correlation spectroscopy with high count rate and low-background - analysis of translational diffusion. *European Biophysics Journal 22*, 3 (1993), 169–175.

[142] Rogers, S. A.; Lisicki, M.; Cichocki B.; Dhont J. K. G.; Lang P. R. Rotational diffusion of spherical colloids close to a wall. *Physical Review Letters 109*, 9 (2012), 098305.

[143] Roig, A. R.; Alessandrini, J. L. Particle size distributions from static light scattering with regularized non-negative least squares constraints. *Particle & Particle Systems Characterization 23*, 6 (2006), 431–437.

[144] Rosen, M. J. *Surfactants and Interfacial Phenomena*, 3rd ed. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2004.

[145] Roux, B. The calculation of the potential of mean force using computer simulations. *Computer Physics Communications 91*, 1-3 (1995), 275–282.

[146] Sadtler, V. M.; Krafft, M. P.; Riess J. G. Reverse water-in-fluorocarbon emulsions as a drug delivery system: an in vitro study. *Colloids and Surfaces A: Physicochemical and Engineering Aspects 147*, 3 (1999), 309–315.

[147] Salinas, S. V.; Ning, C. B.; Liew S. C. Characterization of aerosol physical and optical properties from a combination of ground-based and hand-held sunphotometer data of singapore. *2009 IEEE International Geoscience and Remote Sensing Symposium* (2009), II–531–II–534.

[148] Sanborn, M. E.; Connolly, B. K.; Gurunathan K.; Levitus M. Fluorescence properties and photophysics of the sulfoindocyanine cy3 linked covalently to dna. *Journal of Physical Chemistry B 111*, 37 (2007), 11064–11074.

[149] Schmitt, J. M.; Kumar, G. Optical scattering properties of soft tissue: a discrete particle model. *Applied Optics 37*, 13 (1998), 2788–97.

[150] Schubert, H.; Engel, R. Product and formulation engineering of emulsions. *Chemical Engineering Research and Design 82*, 9 (2004), 1137–1143.

[151] Schuler, B.; Lipman, E. A.; Steinbach P. J.; Kumke M.; Eaton W. A. Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proceedings of the National Academy of Sciences 102*, 8 (2005), 2754–2759.

[152] Schwierz, N.; Horinek, D.; Netz R. R. Anionic and cationic hofmeister effects on hydrophobic and hydrophilic surfaces. *Langmuir 29*, 8 (2013), 2602–14.

[153] Schwille, P.; Kummer, S.; Heikal A. A.; Moerner W. E.; Webb W. W. Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins. *Proceedings of the National Academy of Sciences 97*, 1 (2000), 151–6.

[154] Sindbert, S.; Kalinin, S.; Nguyen H.; Kienzler A.; Clima L.; Bannwarth W.; Appel B.; Mueller S.; Seidel C. A. M. Accurate distance determination of nucleic acids via forster resonance energy transfer: Implications of dye linker length and rigidity. *Journal of the American Chemical Society 133*, 8 (2011), 2463–2480.

[155] Sjöback, R.; Nygren, J.; Kubista M. Absorption and fluorescence properties of fluorescein. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 51*, 6 (1995), L7–L21.

[156] Skinner, J. L.; Pieniazek, P. A.; Gruenbaum S. M. Vibrational spectroscopy of water at interfaces. *Accounts of Chemical Research 45*, 1 (2012), 93–100.

[157] Speelman, A. L.; Muñoz-Losa, A.; Hinkle K. L.; VanBeek D. B.; Mennucci B.; Krueger B. P. Using molecular dynamics and quantum mechanics calculations to model fluorescence observables. *Journal of Physical Chemistry A 115*, 16 (2011), 3997–4008.

[158] Spiess, B. D. Perfluorocarbon emulsions as a promising technology: a review of tissue and vascular gas dynamics. *Journal of Applied Physiology 106*, 4 (2009), 1444–52.

[159] Spiriti, J.; Binder, J. K.; Levitus M.; van der Vaart A. Cy3-dna stacking interactions strongly depend on the identity of the terminal basepair. *Biophysical Journal 100*, 4 (2011), 1049–1057.

[160] Spry, D. B.; Goun, A.; Glusac K.; Moilanen D. E.; Fayer M. D. Proton transport and the water environment in nafion fuel cell membranes and aot reverse micelles. *Journal of the American Chemical Society 129*, 26 (2007), 8122–30.

[161] Tadros, T. F. *Applied Surfactants: Principles and Applications*. 2005.

[162] Taha, G.; Box, G. P. New method for inferring total ozone and aerosol optical thickness from multispectral extinction measurements using eigenvalue analysis. *Geophysical Research Letters 26*, 20 (1999), 3085–3088.

[163] Tang, J.; Jofre, A. M.; Lowman G. M.; Kishore R. B.; Reiner J. E.; Helmerson K.; Goldner L. S.; Greene M. E. Green fluorescent protein in inertially injected aqueous nanodroplets. *Langmuir : the ACS journal of surfaces and colloids 24*, 9 (2008), 4975–8.

[164] Taylor, P. Ostwald ripening in emulsions. *Advances in Colloid and Interface Science 75*, 2 (1998), 107–163.

[165] Thompson, N. L.; Lakowicz, J. R. *Topics in Fluorescence Spectroscopy*. Plenum Press, 1991, ch. Fluorescence Correlation Spectroscopy, pp. 337–374.

[166] Tian, C. S.; Shen, Y. R. Structure and charging of hydrophobic material/water interfaces studied by phase-sensitive sum-frequency vibrational spectroscopy. *Proceedings of the National Academy of Sciences 106*, 36 (2009), 15148–53.

[167] Tonova, K.; Lazarova, Z. Reversed micelle solvents as tools of enzyme purification and enzyme-catalyzed conversion. *Biotechnology Advances 26*, 6 (2008), 516–32.

[168] Tremper, K. K.; Anderson, S. T. Perfluorochemical emulsion oxygen transport fluids: a clinical review. *Annual Review of Medicine 36* (1985), 309–13.

[169] Urnavicius, L.; McPhee, S. A.; Lilley D. M. J.; Norman D. G. The structure of sulfoindocarbocyanine 3 terminally attached to dsdna via a long, flexible tether. *Biophysical Journal 102*, 3 (2012), 561–568.

[170] van Gunsteren, W. F.; Berendsen, H. J. C. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angewandte Chemie International Edition 29*, 9 (1990), 992–1023.

[171] VanBeek, D. B.; Zwier, M. C.; Shorb J. M.; Krueger B. P. Fretting about fret: Correlation between kappa and r. *Biophysical Journal 92*, 12 (2007), 4168–4178.

[172] Vazdar, M.; Pluhařová, E.; Mason P. E.; Vácha R.; Jungwirth P. Ions at hydrophobic aqueous interfaces: Molecular dynamics with effective polarization. *Physical Chemistry Letters 3*, 15 (2012), 2087–2091.

[173] Verga, A D. Brownian motion constrained to enclose a given area. *Journal of Physics A: Mathematical and General 24*, 10 (1991), L561–L564.

[174] Wang, X.; Pogue, B. W.; Jiang S.; Song X.; Paulsen K. D.; Kogel C.; Poplack S. P.; Wells W. A. Approximation of mie scattering parameters in near-infrared tomography of normal breast tissue in vivo. *Journal of Biomedical Optics 10*, 5 (2013), 051704.

[175] Wang, Y. J.; Shen, J.; Liu W.; Sun X. M.; Dou Z. H. Non-negative constraint research of tikhonov regularization inversion for dynamic light scattering. *Laser Physics 23*, 8 (2013), 085701.

[176] Wennmalm, S.; Edman, L.; Rigler R. Conformational fluctuations in single dna molecules. *Proceedings of the National Academy of Sciences 94*, 20 (1997), 10641–10646.

[177] Widengren, J., and Schwille, P. Characterization of photoinduced isomerization and back-isomerization of the cyanine dye cy5 by fluorescence correlation spectroscopy. *Journal of Physical Chemistry A 104*, 27 (2000), 6416–6428.

[178] Widengren, J.; Rigler, R.; Mets U. Triplet-state monitoring by fluorescence correlation spectroscopy. *Journal of Fluorescence 4* (1994), 255–258.

[179] Widengren, J.; Mets, U.; Rigler R. Fluorescence correlation spectroscopy of triplet states in solution: A theoretical and experimental study. *Journal of Physical Chemistry 99*, 36 (1995), 13368–13379.

[180] Woodson, S. A. Metal ions and rna folding: a highly charged topic with a dynamic future. *Current Opinion in Chemical Biology 9*, 2 (2005), 104–9.

[181] Yasuda, R.; Masaike, T.; Adachi K.; Noji H.; Itoh H.; Kinosita K. The atp-waiting conformation of rotating f-1-atpase revealed by single-pair fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences 100*, 16 (2003), 9314–9318.

[182] Zhang, J.; Bright, F. V. Nanosecond reorganization of water within the interior of reversed micelles revealed by frequency-domain fluorescence spectroscopy. *Journal of Physical Chemistry 95*, 20 (1991), 7900–7907.

[183] Zhu, X.; Shen, J.; Liu W.; Sun X.; Wang Y. Nonnegative least-squares truncated singular value decomposition to particle size distribution inversion from dynamic light scattering data. *Applied Optics 49*, 34 (2010), 6591–6.

[184] Zhuang, X. W.; Kim, H.; Pereira M. J. B.; Babcock H. P.; Walter N. G.; Chu S. Correlating structural dynamics and function in single ribozyme molecules. *Science 296*, 5572 (2002), 1473–1476.