CONFORMATIONAL DYNAMICS OF PROTEINS BY FLUORESCENCE FLUCTUATION SPECTROSCOPY

By

Robel Birru Yirdaw

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

August, 2012

Nashville, Tennessee

Approved:

Professor Hassane S. Mchaourab

Professor M. Shane Hutson

Professor David W. Piston

Professor Thomas J. Weiler

Professor Kálmán Varga

Copyright © 2012 by Robel Birru Yirdaw
All Rights Reserved

To
My wife, the sun in my sky,
My mother, the embodiment of love,
My father, the beginning of my scientific journey

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Professor Hassane S. Mchaourab, for his support and guidance. I want to thank Professor Kálmán Varga, Professor Thomas J. Weiler, Professor M. Shane Hutson and Professor David W. Piston for serving on my committee. I also would like to express my thanks to current and former members of the Mchaourab lab for all their help.

I am blessed to have had the love, wisdom and unwavering support of my wife during this challenging process. Without her, this would not be possible. I would also like to express my deepest gratitude to my family for always supporting me. The journey that led to this point has taken years and extends past the time spent in graduate school. Numerous individuals have come into my life along the way and, in some form or another, have empowered me to take a step forward. I am grateful.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER IV

CHAPTER V

LIST OF FIGURES

LIST OF ABBREVIATIONS

APD ..... Avalanche photodiode

ATP ..... Adenosine triphosphate

CA ..... Cumulant analysis

CAMKII ..... Ca^{2+}/Calmodulin-dependent protein kinase II

cpsm ..... counts per second per molecule

ED ..... Essential dynamics

EPR ..... Electron paramagnetic resonance

FCS ..... Fluorescence correlation spectroscopy

FFS ..... Fluorescence fluctuation spectroscopy

FIDA ..... Fluorescence intensity distribution analysis

FRET ..... Fluorescence resonance energy transfer

MD ..... Molecular dynamics

MTS ..... Methanethiosulfonate

NMR ..... Nuclear magnetic resonance

PCA ..... Principal component analysis

PCH ..... Photon counting histogram

PMT ..... Photomultiplier tube

SMFS ..... Single molecule fluorescence spectroscopy

TAMRA ..... Tetramethylrhodamine

T4L ..... T4 lysozyme

WT ..... Wild-type

CHAPTER I

INTRODUCTION

I.1 Proteins: A Brief Historical Perspective

The biochemical classification of the molecules found in living organisms is deceptively simple. All of the molecules that make life possible can be grouped into four main classes: lipids, nucleic acids, proteins and carbohydrates. Of these, proteins are by far the most complex and as such have been the center of extensive research over the years, capturing the imagination of generations of scientists. An early account of the study of proteins can be traced back to the 18th century, to the work of Antoine Fourcroy, a prominent French chemist (1). At the time, these molecules were simply referred to as albumins and Fourcroy was engaged in obtaining the chemical understanding of albumins. Years after Fourcroy, the Dutch physician and chemist Gerrit Mulder probed the composition of albumins and fibrins and found them to have the same composition 'except for 1 atom of sulphur' (1). He communicated his findings to his colleague, the Swedish chemist Jacob Berzelius, who had also been actively examining similar molecules. In 1838, inspired by Mulder's findings, Berzelius coined the word "protein" to describe Mulder's molecules, a term derived from the Greek word *proteios*, meaning 'of the first rank' (1, 2).

A milestone event in the science of proteins occurred in 1902 with the simultaneous elucidation of the peptide bond, the primary bond in proteins, by Franz Hofmeister and Emil Fischer (1). Progress in our modern understanding of proteins then took a major leap with the work of James B. Sumner who, in 1926, showed that the enzyme urease was actually a protein. The following years brought on even more milestones. In 1949, Frederick Sanger, working with insulin, introduced protein sequencing. Less than a decade later, in 1958, John Kendrew determined the crystal structure of myoglobin, the first three dimensional atomic-scale structure

of a protein. The crystal structure of hemoglobin, a protein closely related to myoglobin, was determined by Max Perutz in 1959.

These were followed by other notable contributions to the field such as the work of Boyer and Walker in elucidating the function of a special enzyme, ATP Synthase, responsible for the production of ATP, the energy currency of life. The work of Christian Anfinsen and collaborators on ribonucleases, another special class of enzymes involved in breaking down nucleic acids, would lead to a fundamental insight into protein structure (2). Along the same line, Linus Pauling and his colleagues provided a description of the basic structural components of proteins, another important accomplishment. All of the above achievements, except those of Hofmeister, Fischer and Pauling, were recognized with the Nobel Prize in Chemistry.

I.2 Structural Basics of Proteins

The basic component of a protein is an amino acid molecule of which there are twenty variants. All of the amino acids, from which essentially all proteins are constructed, are composed of just five atoms: hydrogen, carbon, nitrogen, oxygen and sulfur. In a protein, individual amino acids are sequentially linked by a peptide bond to form a linear peptide chain and thus a protein is commonly referred to as a polypeptide (Fig. 1-1). In some cases, a protein may be composed of more than one peptide chain.

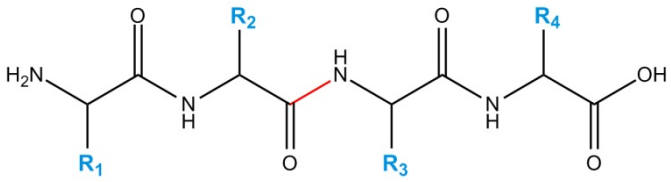


Figure 1-1. An example polypeptide with one peptide bond (red) and amino acid residues (R) indicated.

The sequence of amino acids for a given protein is genetically determined. This fact is captured by the central dogma of biology: "DNA makes RNA makes protein". The sequence of amino acids is the primary structure of a protein. A set of amino acids, in turn, can form three types of structures: an alpha helix, beta strand and loop structure (Fig. 1-2) (3). These are the secondary structures of proteins. The combinatorial arrangement of secondary structures, as observed in various proteins, gives specific types of structural motifs.

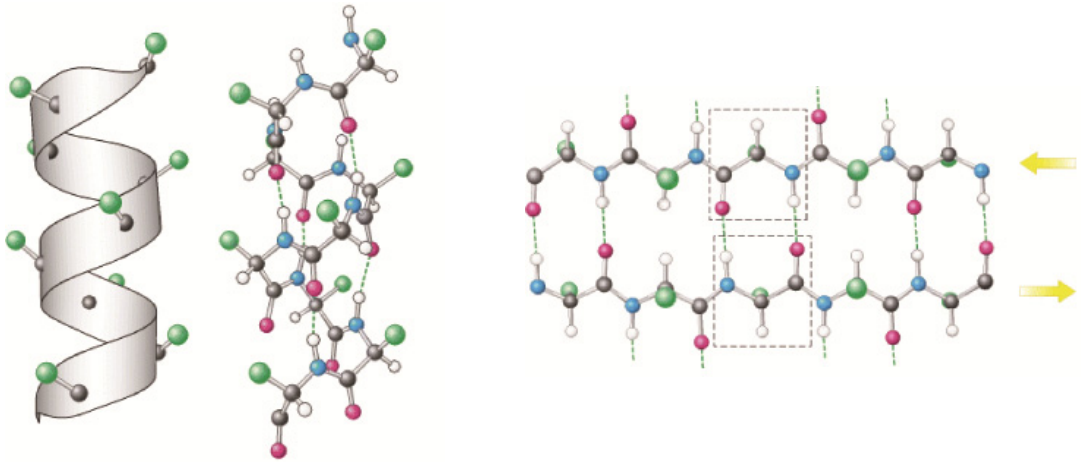


Figure 1-2. A ribbon of an alpha-helix and ball-and stick models of an alpha-helix and an anti-parallel beta strand arrangement. In the ribbon model of the helix, the alpha-carbon atoms (black) and side chains (green) are indicated. For both secondary structures, hydrogen bonding interactions are shown (dashed lines). Adopted from Ref. 2.

The three dimensional configuration of secondary structures defines a protein's tertiary structure or fold. In cases where a protein consists of multiple peptide chains, tertiary structures will combine to form a quaternary structure. The final and functional form a protein is therefore its tertiary or quaternary structure. The ability to determine the tertiary structure of a given protein from its primary structure remains the most important unanswered question of molecular biology, although significant advances have been made over the years. One notable finding comes from Christian Anfinsen's work on ribonucleases which revealed that all of the information required to obtain the three dimensional structure of a protein is contained within the amino acid sequence (2).

Various bonding interactions are at play in the tertiary and quaternary structures of proteins. The peptide bonds that hold the individual amino acids together are the only covalent bonds in proteins. In the alpha-helix and beta strands, the structures are held together by hydrogen bonding interactions (Fig. 1-2), as first revealed by Pauling (1, 2). More importantly, the tertiary structure is also held together by hydrogen bonding interactions as well as van der Waals and electrostatic interactions. In addition, proteins, in general, are either in aqueous or lipid environments and these solvent molecules also contribute to the stability of the tertiary structure. Thus, the functional structure of proteins is made possible by the intricate balance of various interactions from within the protein and from the surrounding molecules.

The primary method by which the three dimensional structure of a protein is determined at atomic-scale is X-ray crystallography. Since the monumental work of Perutz and Kendrew in the 1950s, this technique has allowed the determination of over 80,000 structures, as of this writing (PDB). Nuclear magnetic resonance (NMR) can also be used to determine protein structure. Unlike crystallography, only small proteins (less than ~ 30 – 50 kDa) (4, 5) can be studied by NMR. The primary advantage of NMR over crystallography is that a crystal of the given protein is not required for an NMR experiment. The process of protein crystallization is the least understood and possibly the most difficult aspect of protein crystallography and is still primarily trial-and-error (4). NMR also has the added advantage of providing information on the flexibility or dynamics of the protein, an important property of proteins and the theme of the dissertation work presented here.

I.3 "Kicking and Screaming"

In contrast to the static pictures obtained by X-ray crystallography, it has long been recognized that proteins exhibit internal motions at equilibrium, as succinctly described by Gregorio Weber almost 40 years ago:

> "Indeed the protein molecule model resulting from the X-ray crystallographic observations is a 'platonic' protein, well removed in its perfection from the kicking and screaming 'stochastic' molecule that we infer must exist in solution." (6)

The internal motions of proteins span multiple orders of magnitude in timescale. In general, these motions can be grouped into three types: motions involving individual atoms and/or residues, collective motions of atoms and/or residues and motions that are triggered externally (5, 7, 8).

The fastest motions in proteins are those of the first type involving individual residues and atoms. This includes bond vibrations and rotations of residue methyl groups and aromatic rings. Such motions occur on a timescale of a picosecond or less and are identified as small amplitude motions (5, 8, 9). The second type of motion involves the concerted movement of a group of atoms. These can be either local motions with timescales in the range of ps – ns or the motion of large regions of a protein such as hinge-bending motions, which are typically in the range of microseconds to milliseconds. The amplitude in collective motions can span 0.01 – 5 Å (5). Finally, the third type of motion involves the structural changes that occur as a result of an external source, such as the binding of a ligand to a protein. These can be the slowest of the three types of motions, taking up to minutes (5, 7, 8) and involve the largest displacement of atoms, up to 10 Å (5). Such large scale, induced, motions are of utmost importance as they are directly related to protein function. A crucial difference between these motions is that the first two are intrinsic protein motions - they occur without the requirement of an external molecule, unlike the third type. Nevertheless, all three types of motions are enabled primarily by the weak interactions that are responsible for the tertiary structure of proteins: hydrogen, electrostatic and van der Waals bonds.

The term protein dynamics, therefore, can refer to all three types of motions. At the heart of protein dynamics studies is the elucidation of the link between the three types of motions but most importantly the link between the intrinsic and induced motions. Although the fast timescale motions are thought to play a role in the functional motions of proteins (10), protein dynamics

studies focusing on the relationship of intrinsic collective motions to functional motions are finding considerable success. Examples of such studies are given in the following section.

I.4 The Role of Dynamics in Protein Function

Allostery refers to the process where the binding of a ligand to a protein at one location causes a structural change in the protein at a different location. This type of interaction is responsible for metabolic processes and gene regulation to a large degree (11). In an attempt to describe the underlying mechanism of allostery, two models were proposed almost 50 years ago. These models were developed within the context of oligomeric proteins as the well-known proteins at the time were the classic oligomers myoglobin and hemoglobin.

The first model, proposed by Koshland, Nemethy and Filmer in 1958 (12), is commonly referred to as the sequential model. The basic idea of this model is that ligand binding to a protein's subunit causes a change in that subunit and not the adjacent subunits. However, the interaction between the ligand-bound subunit and adjacent subunits will be affected due to the conformational change that resulted from the binding. This in turn results in an increase in binding affinity of the adjacent subunits. The subunit with bound ligand is said to be in the relaxed (R) state while those without ligand are in the tense (T) state (2). Thus, ligand binding results in a transition of one subunit from the T state to the R state while increasing the probability of such transitions in the adjacent subunits.

The second model, proposed by Monod, Wyman and Changeux (13), is the concerted model. In this case, the protein is initially assumed to be interconverting between states T and R. It is also assumed that the T state has a lower affinity for ligands than the R state. Ligand binding to one subunit then takes that subunit to the R state. However, as the protein is interconverting as a whole between the two states, the binding of ligand favors the R state for the other subunits, raising the binding affinity of the entire protein (2).

Although these two classical descriptions of allostery differ in their treatment of conformational changes, both prescribe to the presence of only two conformational states, T and R. In light of various experiments (14), this view of allostery has been amended. In the current view, the two-state assumption is replaced with an 'ensemble view' (11). The protein is no longer occupying only two states but exists in an ensemble of conformations. The effect of ligand binding is to simply shift the composition of conformational populations (11, 14-16). In other words, the conformational changes that occur upon binding are already being sampled by the intrinsic motions of the protein – it is the equilibrium motions that are involved in the allosteric conformational changes.

A second example illustrating the importance of intrinsic motions can be found in enzyme function. That protein dynamics and flexibility play a crucial role in enzyme function has been long recognized (17). Large scale motions such as domain motions allow access to the active site of a protein and enable the capture and isolation of substrates from the solvent. Such motions also prevent reaction intermediates from escaping (18). Moreover, fast and slow timescale motions are thought to allow proper coordination of the substrate and residues involved in the reaction (18, 19).

Although the significance of intrinsic motions has been acknowledged, a detailed and fundamental understanding of the relationship between protein motions and enzyme catalysis is slowly being revealed. There is now evidence showing that the conformational changes that occur during catalysis are those visited by the protein at equilibrium (20). Furthermore, Eisenmesser et al. (20) report that the rate of equilibrium motions is comparable to the catalytic turnover rate. Therefore, this suggests that the conformational changes required for catalysis of an enzyme are an inherent property of the enzyme.

Additional experiments have also shown that intrinsic motions can be the rate-limiting processes in enzyme catalysis (19, 21). The work of Wolf-Watz et al. (21) provides an illustrative example. Their study used NMR to probe the link between flexibility and catalysis in

the temperature sensitive enzyme, adenylate kinase (AdK). Specifically, their experiments involved comparison of two types of AdK: mesophille and hyperthermophile variants. Hyperthermophillic proteins have the capability to function in extreme temperatures (over 80 degrees Celsius), where other proteins like the mesophillic homolog will denature. It was previously observed that the catalytic activity of the hyperthermophillic enzyme is significantly reduced relative to the mesophillic counterpart at a given temperature. Wolf-Watz et al. found that the reason for the reduced catalytic activity is precisely due to the reduced dynamics of the hyperthermophile AdK at the lower temperature. Furthermore, they found that intrinsic motions were the rate-limiting processes in the activity of both types of enzymes.

Thus, experiments are slowly elucidating the role played by protein motions in enzyme function as well as protein allostery. Studies focusing on the contribution of the various types of protein motions in enzyme function make up one active area of investigation (10). Another involves the examination of the role of protein motions in the chemical reaction itself, in addition to the various processes that occur before the reaction takes place (22, 23).

I.5 Tools for the Study of Protein Dynamics

Over the years, various techniques have been developed and utilized for the study of motions in proteins. These tools can be categorized in terms of the timescales of the motion they are able to detect. Fast timescale motions in the region of fs – ps can be probed by vibrational spectroscopy methods, namely infrared and Raman spectroscopy. Moreover, Mossbauer spectroscopy, neutron scattering and fluorescence depolarization are able to extend this fast timescale to approximately 10 μs (7, 24).

Probably a more popular and powerful tool for the examination of protein motions in the fast timescale regime is molecular dynamics simulations (MD). Since the first simulation of a protein in 1977 (7), molecular dynamics has grown tremendously – today, highly accurate, quantum mechanical simulations are routinely performed, although these are limited to small

systems and short durations (~100 ps). The added advantage enjoyed by MD is the seemingly never-ending improvement of computational power.

For protein motions of biological interest, motions in the slow timescale regime (microseconds to seconds), the primary tools are X-ray crystallography, nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR) and single molecule fluorescence spectroscopy. Classical MD simulations currently fall short of the timescale of biological interest purely due to the lack of computational power. Also, the issue of force-field variability (9, 25) has yet to be fully addressed. The observation of protein motions using X-ray crystallography is possible if multiple crystal structures of the same protein showing differing conformations are obtained (9). The various conformations can be created experimentally, usually via specific point mutations. EPR enables the detection of protein motions with high sensitivity. In this case, the detection can be confined to local areas or whole regions of the protein.

Unlike X-ray crystallography and EPR spectroscopy, NMR spectroscopy has the double advantage of providing the timescale of motions along with the structure itself. This technique has been extensively used for the study of protein dynamics (9). Single molecule fluorescence spectroscopy (SMFS) is a relatively new technique that also allows proteins to be studied in their native environment. The three main variants of this technique are fluorescence correlation spectroscopy (FCS), cumulant analysis (CA) / photon counting histogram (PCH) and fluorescence resonance energy transfer (FRET). Of these, FCS is of utmost importance as it is the only SMFS technique that allows the determination of timescales. Single molecule fluorescence as whole has the advantage of being able to discern heterogeneity in the system by virtue of its single molecule nature.

Thus, a plethora of techniques are available that allow the observations of protein motions over all relevant timescales. Moreover, the combination of these tools can provide a more powerful approach to examination of the ‘kicking and screaming’ of proteins.

I.6 Description of Research

In the work described here, fluorescence correlation spectroscopy and cumulant analysis have been used to investigate the conformational dynamics and substrate-induced conformational changes of the protein T4 lysozyme. The Ca^{2+}/calmodulin-dependent protein kinase II (CAMKII) and membrane transporter MsbA have also been examined. In chapter II, the theoretical background of FCS and cumulant analysis is presented. FCS makes use of specific models for data analysis and those models that are used in this work will be derived in detail. Relative to FCS, cumulant analysis is a new technique and as such not widely used. In addition to describing the fundamental theory of cumulant analysis, the necessary corrections that need to be considered for proper data analysis will be discussed. Details on instrumentation and overall data acquisition and analysis approaches are presented in chapter III.

Chapter IV gives a conformational dynamics study of T4 lysozyme (T4L), a protein from the bacteriophage T4. T4L is involved in breaking down the *E.coli* cell wall during infection of the bacteria by the phage. Its structure consists of two domains joined by a long helix. Early crystal structures of this protein indicated the presence of domain motion at equilibrium. Various studies (EPR, NMR, MD) have probed the conformation of T4L and the presence of domain motion is now well established. However, the timescale of the motion is still unknown. Furthermore, the composition of the equilibrium conformation of T4L in solution is not understood. The work presented in this dissertation addresses these two questions. The substrate induced changes are also examined. A summary and a discussion on future directions including the application of FFS to CAMKII and MsbA is given in chapter V.

Supplementary FCS results for T4 lysozyme are available in the appendix A. As part of this research work, the following software tools have been developed using MATLAB: FCSLab, FCSSim and Brightness Analysis Tool (BAT). FCSLab and BAT are for the analysis of FCS and brightness analysis data respectively. FCSSim can be used to generate FCS curves. Details on all three can be found in appendix B.

CHAPTER II

FLUORESCENCE FLUCTUATION SPECTROSCOPY

II.1 Background

Fluorescence spectroscopy has been an invaluable tool for the study of biological systems. In the structural and dynamical studies of proteins, fluorescence spectroscopy is well suited because of its high experimental sensitivity, including sensitivity to a protein's environment, as well as its specificity for in vivo studies. The main requirement of the technique is a fluorescent molecule whose emission constitutes the primary observable. For proteins, this technique can take advantage of the intrinsic fluorescence of the three amino acid types capable of fluorescing: Trp, Tyr and Phe. On the other hand, a fluorescent molecule, or fluorophore, can be introduced onto the protein using site-directed mutagenesis.

Both approaches have their advantages but the second provides valuable flexibility. Using extrinsic fluorophores allows one to probe any region of the protein, while with intrinsic fluorescence one is restricted to the regions of the protein where the fluorescent amino acids are found. Moreover, the second approach provides flexibility in instrumentation as there are a variety of fluorophores available that can be matched with a given instrument. In addition, emission from intrinsic protein fluorescence is significantly weak compared to that of extrinsic fluorophores. Novel fluorophores are also routinely being designed and developed, making it easier to find a specific type of fluorophore.

Traditional fluorescence spectroscopy is an ensemble technique. Advances in detection instrumentation, excitation sources and more importantly techniques for increasing the signal-to-noise ratio have ushered in the era of single molecule fluorescence spectroscopy (SMFS) (26, 27). The ability to detect individual molecules is the maximum achievable sensitivity in the study of biological molecules such as proteins. SMFS currently has three main variants: fluorescence

correlation spectroscopy (FCS), brightness analysis thru cumulant analysis or photon counting histogram (PCH) and fluorescence resonance energy transfer (FRET). The first two techniques make up a sub-group commonly referred to as fluorescence fluctuation spectroscopy (FFS) (28). These two are the techniques used in this work and details of both types are given in the following sections.

In a typical FFS experiment (26, 27) a sample of very low concentration (~ nM) fluorescent molecules is used. An excitation source is directed into the sample through an inverted microscope arrangement and then focused with a high-NA objective. Either one-photon or two-photon excitation can be utilized. In the case of one-photon excitation, a proper pinhole needs to be used in order to limit background signal (confocal). For the case of two-photon excitation, the probabilistic nature of two-photon absorption will confine the emission to a very small region (~ fL) and therefore the use of a pinhole is not required. The emitted fluorescence is then detected by a photomultiplier tube (PMT) or avalanche photodiode (APD) and is transferred to a data acquisition card on a computer. For both FCS and brightness analysis the raw data is therefore the detected set of photon counts. The two techniques diverge in the way the raw data is analyzed.

II.2 Fluorescence Correlation Spectroscopy (FCS)

II.2.1 Introduction

Fluorescence correlation spectroscopy (FCS) was first introduced in 1972 (29). Conceptually, FCS utilizes fluctuations of fluorescence to extract physical parameters of one or more processes in a given system. FCS is able to obtain physical parameters through the correlation of the temporal fluctuations of fluorescence. In theory, then, any process that can cause fluorescence fluctuations can be examined by this technique. There are various types of

processes that can cause fluorescence fluctuations such as diffusion of molecules, chemical reactions, conformational dynamics and binding reactions (26, 27).

The analysis of these processes using FCS first requires the generation of the correlation curve from the fluorescence data (30, 31). This experimental correlation curve is then analyzed using a specific FCS model for the given process. Multiple FCS models have been derived that can describe a number of processes like those mentioned above. The analysis gives values that characterize the underlying process.

II.2.2 The Theory of FCS

II.2.2.1 Fundamental Formalism

Let the absorbance cross-section of a fluorophore and the incident excitation intensity be represented by σ_A and $I(r,t)$, respectively. The instantaneous absorbance or excitation of the fluorophore is then simply given by $\sigma_A I(r,t)$. If the spatial and temporal profiles of the incident excitation can be separated, the instantaneous excitation can be written as $\sigma_A I_0(t) S(r)$, where the temporal and spatial behaviors are represented by $I_0(t)$ and $S(r)$ respectively (32).

For a system composed of multiple fluorophores in a given volume (V), the number of photons absorbed is given as (32)

$$N_{abs}(t) = \int dV \sigma_A S(r) I_0(t) C(r,t) \tag{2.1}$$

where $C(r,t)$ gives the concentration at position r and time t. The corresponding detected emission or fluorescence, in general form, is therefore (32)

$$F(t) = \eta \phi N_{abs} \tag{2.2}$$

where the terms η and ϕ represent the fluorophore quantum yield and instrument detection efficiency, respectively. For the case of two-photon excitation, as utilized in this work, and rearranging the terms given in Eqs. 2.1 and 2.2, the detected fluorescence becomes (32)

$$F(t) = \frac{1}{2}\eta_2\phi(\overline{I_0^2(t)})\sigma_{2A}\int dVS(r)C(r,t) \tag{2.3}$$

with the quantum yield, average intensity and absorbance cross-section terms modified to indicate the two-photon process. The factor of ½ reflects the fact that a single photon is associated with the absorbance of two photons for excitation. In addition, the detected emission has to be taken as a time-average.

The normalized second-order temporal correlation of fluorescence fluctuations is defined as (27, 30)

$$G(t_i) = \frac{\langle \delta F(t)\delta F(t+t_i)\rangle}{\langle F\rangle^2} \tag{2.4}$$

where t_i is the lag time and the fluorescence fluctuation is defined as the deviation of fluorescence at time t from the average,

$$\delta F(t) = F(t) - \overline{F} \tag{2.5}$$

For stationary systems, the correlation term given in Eq. 2.5 will not be dependent on absolute time. Instead, the lag time t_i can be used,

$$G(t_i) = \frac{\langle \delta F(0)\delta F(t_i)\rangle}{\langle F\rangle^2} \tag{2.6}$$

With the description of fluorescence given in Eq. 2.3, the correlation becomes

$$G(t_i) = \frac{\int dV S^2(r) \int dV' S^2(r') \langle \delta C(r,0) \delta C(r', t_i) \rangle}{\left(\int dV S^2(r) \right)^2} \tag{2.7}$$

The above is the most basic form of the correlation. Therefore, the description of the system in terms of the concentration fluctuations is necessary for the determination of the correlation. The processes responsible for concentration fluctuations over time will determine the final form of the function and produce the various FCS models. In the following section, the case for the simplest process, a system of diffusing molecules, will be considered first followed by more complex processes.

II.2.2.2 System Specific Variations

Single Component Diffusion

For a system consisting of identical molecules in solution, where the molecules are allowed to freely diffuse through the observation volume, the physical source of fluorescence fluctuations is diffusion. The temporal dependence of the concentration is governed by the diffusion equation (the master equation) (27, 30), i.e.,

$$\frac{\partial}{\partial t}(\delta C(r,t)) = D \nabla^2 (\delta C(r,t)) \tag{2.8}$$

The diffusion coefficient of the molecules is given by D. To continue further, the solution to the above differential equation is required. This is done by taking the Fourier transform (27, 30) and substituting the resulting form into the concentration fluctuation term in Eq. 2.7 above gives

$$\left\langle \delta C(r,0)\delta C(r',t_i)\right\rangle = (2\pi)^{-3/2}\int dk Exp[-D|k|^2 t]Exp[-ik\cdot r']\left\langle \delta C(r,0)\delta \widetilde{C}(k,0)\right\rangle \tag{2.9}$$

Next, applying inverse Fourier transform gives

(2.10)

$$\left\langle \delta C(r,0)\delta C(r',t_i)\right\rangle = (2\pi)^{-3}\int dk Exp[-D|k|^2 t_i]Exp[-ik\cdot r']\int dr'' Exp[ik\cdot r'']\left\langle \delta C(r,0)\delta C(r'',0)\right\rangle$$

Since in a typical FCS experiment the sample is assumed to be dilute, the molecules do not possess spatial correlation. Furthermore, the number fluctuations can be taken to obey Poisson statistics (the variance is identical to the average), leading to (27, 30)

$$\left\langle \delta C(r,0)\delta C(r'',0)\right\rangle = \overline{C}\,\delta^3(r-r'') \tag{2.11}$$

Substituting this back into Eq. 2.10 and reducing the inner integral gives

$$\left\langle \delta C(r,0)\delta C(r',t)\right\rangle = (2\pi)^{-3}\overline{C}\int dk Exp[-D|k|^2 t]Exp[ik\cdot r]Exp[-ik\cdot r'] \tag{2.12}$$

With the above final form for the concentration fluctuation term, the correlation then becomes

$$G(t_i) = (2\pi)^{-3}\overline{C}\int dk Exp[-D|k|^2 t_i]\int dr S(r)Exp[ik\cdot r]\int dr' S(r')Exp[-ik\cdot r'] \tag{2.13}$$

The second and third integrals in the above are themselves Fourier transforms, i.e.,

$\int drS(r)Exp[ik \cdot r]$ is the Fourier transform of *S(r)* with a missing factor of $(2\pi)^{-3/2}$ and similarly for the third integral. Using this fact and the property of Hermitian symmetry of Fourier transforms on real functions, $\widetilde{S}(-k) = \left(\widetilde{S}(k)\right)^*$, where the asterisk represents the complex conjugate, the correlation takes the form

$$G(t_i) = \overline{C}\int dkExp[-D|k|^2 t_i]\left|\widetilde{S}(k)\right|^2 \tag{2.14}$$

Thus, the normalized second-order correlation has been reduced to a term consisting of the average concentration, while the integral contains two terms which we can call a PSF and non-PSF terms, i.e. a term with the diffusion coefficient and a term that is the square of the Fourier transform of the point spread function (PSF). The following three dimensional Gaussian for two-photon excitation, normalized to the intensity at the center of the observation volume (I_0^2), is commonly used for the PSF (27):

$$S(r) = I_0^2 Exp\left[-\frac{4}{w_0^2}(x^2+y^2)\right]Exp\left[-\frac{4z^2}{z_0^2}\right] \tag{2.15}$$

The dimensions of the observation volume are represented by w_0 and z_0. Taking the Fourier transform of the PSF above, substituting into Eq. 2.14 and evaluating the resulting integral will give

$$G(t_i) = \frac{1}{\overline{C}}\left[\frac{(8\pi)^{3/2}}{w_0^2 z_0}\left(1+\frac{8Dt_i}{w_0^2}\right)^{-1}\left(1+\frac{8Dt_i}{z_0^2}\right)^{-1/2}\right] \tag{2.16}$$

Finally, for a three dimensional Gaussian along with two-photon excitation, the volume is given by (33)

$$V = \frac{\pi^{3/2} w_0^2 z_0}{8} \tag{2.17}$$

Using the above and the relation $\overline{N} = V\overline{C}$ to simplify the correlation as shown in Eq. 2.16 gives the final form for a system of diffusing molecules (diffusion-only model) (34):

$$G(t_i) = \frac{1}{2\sqrt{2}\overline{N}} \left(1 + \frac{8Dt_i}{w_0^2}\right)^{-1} \left(1 + \frac{8Dt_i}{z_0^2}\right)^{-1/2} \tag{2.18}$$

Single Component Diffusion with Single Component Quenching

In some cases, it is possible that the fluorescence fluctuation is a result of not just diffusion but also an additional process that can modulate the fluorescence while the molecule is within the observation volume. For such cases, the additional process needs to be accommodated in the master equation (27, 30) which, in its general form, is given as

$$\frac{\partial}{\partial t}(\delta C_j(r,t)) = D_j \nabla^2 (\delta C(r,t)) + \sum_1^m k_{jl} \delta C_l(r,t) \tag{2.19}$$

The indexing is now required because, in terms of the fluorescence characteristics of the molecules, at a given time within the observation volume, there are *m* different types. The rates of the transitions are given by k_{jl}. The above system of linear differential equations is often represented in vectorized form (27, 30)

$$\frac{\partial}{\partial t}\begin{bmatrix} \delta C_1(r,t) \\ \vdots \\ \delta C_m(r,t) \end{bmatrix} = \mathrm{M}\begin{bmatrix} \delta C_1(r,t) \\ \vdots \\ \delta C_m(r,t) \end{bmatrix} \tag{2.20}$$

Using this form of the master equation and repeating the steps outlined in the previous section for the system of diffusing molecules will lead to the following form for the non-PSF term in the integral of Eq. 2.14

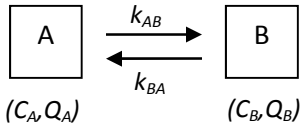
$$\sum_{j,l} Q_j Q_l \overline{C}_j \sum_s X_l^s Exp(\lambda^s t_i)(X^{-1})_l^s \tag{2.21}$$

The quantities X and λ are the eigenvectors and eigenvalues of the matrix M, respectively. More importantly, the inverse term is the left eigenvector of M (30, 35). With the evaluation of the PSF term already done in the previous section, what is left for this type of a system then reduces to evaluating the above for the desired number of species that make up the system. This task can be easily done using *Mathematica*. The following code listing will perform the evaluation of Eq. 2.21.

```
maxIndx=Length[q];
jSum=0;
For[j=1,j≤maxIndx,j++,
  lSum=0;
  For[l=1,l≤maxIndx,l++,
   sSum=0;
   For[s=1,s≤maxIndx,s++,
    sSum = sSum + x[[l,s]] Exp[λ[[s]]] xInv[[s,j]];
    ];
   lSum=lSum+q[[l]] sSum;
   ];
  jSum=jSum+q[[j]] c[[j]] lSum;
  ];
```

The terms c and q are the concentration and brightness vectors while the eigenvalues and eigenvectors are as described and must be constructed beforehand.

The simplest example for the current type of system in FCS is represented by the schematic below. In this case, the system consists of a single species that can occupy two different fluorescent states. A system with intersystem crossing is one example.



Here, the molecules are thought to be fluctuating between states *A* and *B*, with corresponding rates k_{AB} and k_{BA}. C_A and C_B give the concentrations of molecules in the respective states and the brightness of these molecules is represented by Q_A and Q_B.

Setting up the matrix *M* and following the steps as outlined in the previous section will lead to the following general form for the system:

$$G(t_i) = \frac{1}{2\sqrt{2}(Q_A\overline{C}_A + Q_B\overline{C}_B)^2 V_{3DG}}[\Omega]\left[\left(1+\frac{8Dt_i}{w_0^2}\right)^{-1}\left(1+\frac{8Dt_i}{z_0^2}\right)^{-1/2}\right] \tag{2.22}$$

where

$$\Omega = \frac{(Q_A\mathrm{K}+Q_B)(Q_A\overline{C}_A+Q_B\overline{C}_B)+(Q_A\overline{C}_A-Q_B\overline{C}_B\mathrm{K})(Q_A-Q_B)Exp\left[\frac{-t_i}{\tau_r}\right]}{\mathrm{K}+1}$$

$$\mathrm{K} = \frac{k_{BA}}{k_{AB}}, \tau_r = (k_{AB}+k_{BA})^{-1}$$

If we now rewrite the brightness as a ratio, $Q = \frac{Q_B}{Q_A}$

the correlation takes the form that is sometimes used in the literature for the type of system under consideration (36, 37)

$$G(t_i) = \frac{1}{2\sqrt{2}\overline{N}}\left[1 + \frac{\mathrm{K}(1-Q)^2}{(1+KQ)^2} Exp\left[\frac{-t_i}{\tau_r}\right]\right]\left[\left(1+\frac{8Dt_i}{w_0^2}\right)^{-1}\left(1+\frac{8Dt_i}{z_0^2}\right)^{-1/2}\right] \tag{2.23}$$

Finally, with the assumption that state *B* is non-fluorescent, i.e., $Q = 0$, we get the most familiar form for a system with two-state fluctuation and diffusion (diffusion/quenching model),

$$G(t_i) = \frac{1}{2\sqrt{2}\overline{N}}\left[1 + K \cdot Exp\left[\frac{-t_i}{\tau_r}\right]\right]\left[\left(1+\frac{8Dt_i}{w_0^2}\right)^{-1}\left(1+\frac{8Dt_i}{z_0^2}\right)^{-1/2}\right] \tag{2.24}$$

In this model, the respective populations of the two states (*A* and *B*) can be calculated from the amplitude term, *K*. Since this term is a ratio of rates (equilibrium constant), it follows that

$$\mathrm{K} = \frac{k_{BA}}{k_{AB}} = \frac{C_B}{C_A}$$

If we define the total population as $C_T = C_A + C_B$, the above relation can be written in terms of the fraction of molecules in state *B*, denoted by *F*:

$$\mathrm{K} = \frac{F}{1-F}$$

Single Component Diffusion with Two Component Quenching

In the FCS literature dealing with protein dynamics, an FCS model with two exponential components for the fluctuation term can be found (38-41). The form of this model is shown below:

$$G(t_i) = \left[1 + A_1 \cdot Exp\left(-\frac{t_i}{\tau_{R1}}\right) + A_2 \cdot Exp\left(-\frac{t_i}{\tau_{R2}}\right)\right] \cdot G_D(t_i) \tag{2.25}$$

where

$$G_D(t_i) = \frac{1}{2\sqrt{2}\overline{N}}\left[\left(1 + \frac{8Dt_i}{w_0^2}\right)^{-1}\left(1 + \frac{8Dt_i}{z_0^2}\right)^{-1/2}\right]$$

The primary difference between the above model and the simpler (diffusion/quenching) model derived in the previous section (Eq. 2.24) is that the use of the latter in the literature is not necessarily based on a specific description of a system. If one assumes a three-state, linear, system, the FCS derivation will lead to the same model as Eq. 2.25 (42). However, in this case the determination of population values from the two amplitudes, in a similar way as outlined in the previous section for the diffusion/quenching model, is not possible.

An alternative model to the above has been used for photophysical studies where the presence of two independent sources of fluctuation is assumed (43-45). This model has also been applied in protein dynamics studies (46). In this case, the form of the FCS model is given as

$$G(t_i) = \left[1 + A_1 \cdot Exp\left(-\frac{t_i}{\tau_{R1}}\right)\right] \cdot \left[1 + A_2 \cdot Exp\left(-\frac{t_i}{\tau_{R2}}\right)\right] \cdot G_D(t_i) \tag{2.26}$$

II.3 Cumulant Analysis

II.3.1 Introduction

In FCS, the sequence of detected photon counts is temporally correlated to obtain the evolution of fluorescence fluctuations. Thus, FCS relies on the temporal sequence of photon counts. However, instead of analyzing the emission temporally, it is also possible to perform

analysis in terms of the number of photons detected per unit time. In other words, generating a series of photon counts, per fixed time window, will give a histogram of photon counts which can also provide information about the underlying system. While FCS makes use of the temporal behavior of the fluorescence fluctuations, the amplitude behavior, as captured by the photon count histogram, is utilized by the second technique that makes up FFS, cumulant analysis. In this case, since temporal information is lacking, only two types of parameters are available: molecular brightness, which reflects the fluorescence property of a molecule, and the average number of molecules in the observation volume.

Over the years, the use of photon count statistics from FFS experiments has been presented in various forms. One early example employed the approach on an idealized system composed of two types of particles with different fluorescence properties (47). It was shown that the relative number of each type of particle could be determined by using their molecular brightness. Thus, a statistical approach without any temporal dimension has the potential to provide valuable information of a system.

A more rigorous approach to the use of photon counts and their respective moments has been presented by two types of techniques, photon counting histogram (PCH) (48) and fluorescence intensity distribution analysis (FIDA) (49). A characteristic difference between the two approaches is that PCH utilizes a specific mathematical form for the point spread function (PSF), such as a three dimensional Gaussian, while in FIDA an empirical description is taken. The advantage of the latter is that, theoretically any type of observation volume can be handled.

A fourth approach, introduced after PCH and FIDA, is cumulant analysis. This technique also utilizes photon count statistics but is mathematically significantly simpler than either PCH or FIDA. Unlike FCS, which has been developed and used for almost 40 years now, these photon count statistics approaches are relatively new. In addition, the techniques are not as standardized and established as FCS. In the following sections, a detailed description of cumulant analysis will be given.

Unfortunately, these techniques have not found a significant popularity. A literature search with the phrase "photon counting histogram" for the past ten years returns about 135 results while a similar search for "fluorescence correlation spectroscopy" returns over 2300 results. The reason for the lack of use of the photon count statistics approaches is at least two fold. First, the mathematical complexity of PCH and FIDA makes software implementation extremely difficult. Second, the concept of molecular brightness has an inherent arbitrariness in that a fluorescent molecule will have different fluorescence at different locations of the observation volume. As a result, the photon count distribution for just a single fluorophore will have a broad distribution, complicating the ability to resolve molecular species based on brightness alone. This second point is discussed further in the chapter V.

II.3.2 The Theory of Cumulant Analysis

II.3.2.1 Fundamental Formalism

In cumulant analysis, the starting point is the description of the integrated intensity reaching the detectors within a sampling time T (28, 48, 50, 51):

$$W(t) = \int_{-T/2}^{T/2} dt I(t) \tag{2.27}$$

For the case where the sampling time is short, the above reduces to

$$W(t) = I(t)T \tag{2.28}$$

Eq. 2.28 implies that the emission is assumed to be constant for the duration of time T. The specific parameters that determine the intensity (I) are the PSF, the intensity at the center of the

PSF, the fluorophore's quantum yield, absorbance cross-section and instrument factors (28, 48, 51). In addition, in photon count statistics, the moments of the integrated intensity are the key parameters. The r^{th} moment can then be written as

$$\begin{aligned} &< W^r(t) >=< (I(t)T)^r > \\ &< W^r(t) >= T^r < \lambda^r S^r(r) > \end{aligned} \tag{2.29}$$

In the above, the quantity *S(r)* gives the spatial distribution of the intensity as discussed in the previous section dealing with FCS. The quantity λ, which accounts for the properties of the fluorophore, instrument factors and the excitation intensity at the center of the PSF, is the molecular brightness parameter. This value gives the counts of photons detected per molecule per second (cpsm) and is the central parameter in cumulant analysis. The moments of the integrated intensity then become

$$< W^r(t) >= (T\lambda)^r \left(\frac{1}{V}\right) \int dV S^r(r) \tag{2.30}$$

Here, the probability of finding the molecule at a certain position within the volume is given by 1/V. Finally, the integral evaluates to the volume of the defined PSF, $V_{PSF,}$ and the volume contrast factor, γ_r, giving (28)

$$< W^r(t) >= (T\lambda)^r \left(\frac{1}{V}\right) \gamma_r V_{PSF} \tag{2.31}$$

In the discussion given thus far, the detected fluorescence has been treated in terms of intensity. In FFS experiments, however, photon counts are detected. A direct relationship exists

between the integrated intensity and the *factorial cumulants* of the photon counts $\kappa_{[r]}$ (28, 48, 50, 52, 53)

$$\kappa_{[r]} = < W^r(t) > \tag{2.32}$$

A valuable property of the factorial cumulants is that for a total of *N* particles, the total cumulant is given as the sum of the individual cumulants (28). Furthermore, a given system can be homogenous or mixed. Taking this into account, and using the relation between number of particles, volume and concentration *(N=VC)* to simplify the volume terms in the above, the factorial cumulants of photon counts will have the final, general, form as given below (28)

$$\kappa_{[r]} = T^r \gamma_r \sum_{i=1}^{m} \lambda_i^r N_i \tag{2.33}$$

From the experimentally determined factorial cumulants, the population and brightness parameters can be calculated. The above relation shows that two factorial cumulants are required per molecule.

II.3.2.2 Correction for Finite Sampling Time

In FFS experiments, the sampling time can be set to a desired value as permitted by the instrument. However, a sampling time that is considerably long relative to the diffusion of the molecules through the observation volume will create photon count undersampling. In addition, the assumption that the detected intensity is constant for the period of the sampling time, as was done in the previous section, is no longer valid. For experiments performed under this condition, the factorial cumulants of photon counts must be calculated with a correction term to account for

the finite sampling time effect. Unlike cumulant analysis, this scenario cannot be handled by the photon counting histogram approach (PCH) (48).

The correction term, also called the binning function, for a factorial cumulant of order *r* is given by the following (53)

$$B_r(T,\tau_D,\alpha) = \tau_D^r r! \int \cdots \int G(x_2, x_3, ..., x_r; 1, \alpha)(\frac{T}{\tau_D} - x_2 - ... - x_r) dx_2 ... dx_r \quad (2.34)$$

where τ_D is the characteristic diffusion time of the molecules, α is the square of the ratio of the axial to the lateral dimension of the observation volume, i.e., $\alpha=(z_0/w_0)^2$, and $G(x_2, ..., x_r)$ is the correlation function of order *r*. In the correlation function, $x_i = t_i/\tau_D$, where t_i is the lag time. Aside from the 2nd order correlation function, all higher order functions cannot be solved analytically.

Thus, the calculation of undersampling correction terms requires numerical integrations to be performed making this part of cumulant analysis the most time consuming and potentially computationally intensive. In order to avoid the numerical integration of the higher-order correlation functions, Wu and Muller (53) have suggested preparing a look-up table of the correction terms. During the analysis, the table values can be easily accessed. However, in this approach the correction values are calculated for a fixed value of the characteristic diffusion time. This can be overcome by calculating values for $\tau_D = 1$ and using this set of values to extrapolate to any arbitrary characteristic diffusion time (53).

A third approach developed in this dissertation work is to model the higher-order binning functions. The model is a function of *x* and has the following form

$$\widetilde{B}_r(x;a,b,c,d) = \frac{1}{\left(1 + \frac{x^a}{b} + \frac{x^c}{d}\right)} \quad (2.35)$$

The parameters *a, b, c* and *d* can be obtained through nonlinear fitting of the binning function using *Mathematica*. Fig. 2-1 shows that the third-order binning function values obtained from using the model in Eq. 2.35 is in good agreement with values from numerical integration of the third-order binning function performed using *Mathematica.*

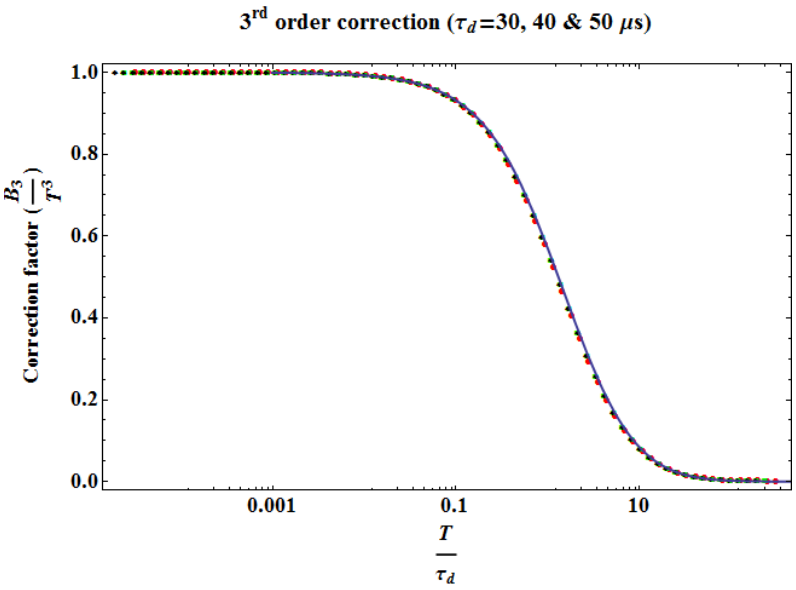


Figure 2-1. Third order correction values calculated by using the model (Eq. 2.35), for diffusion time values of 30, 40 and 50 µs, are shown in red, green and black points respectively. The solid line corresponds to values determined directly from the numerical integration of the third-order binning function.

The effect of the aspect ratio and characteristic diffusion time values on the corrections, calculated using the second order binning function, are shown in Figures 2-2 and 2-3 respectively. Varying the aspect ratio by a factor of two, for a given diffusion time value, changes the correction factor minimally. However, a factor of two difference in the diffusion time of the molecules produces a significant difference in the correction values. For both cases, the figures show that the correction is more severe with increasing sampling time, as expected.

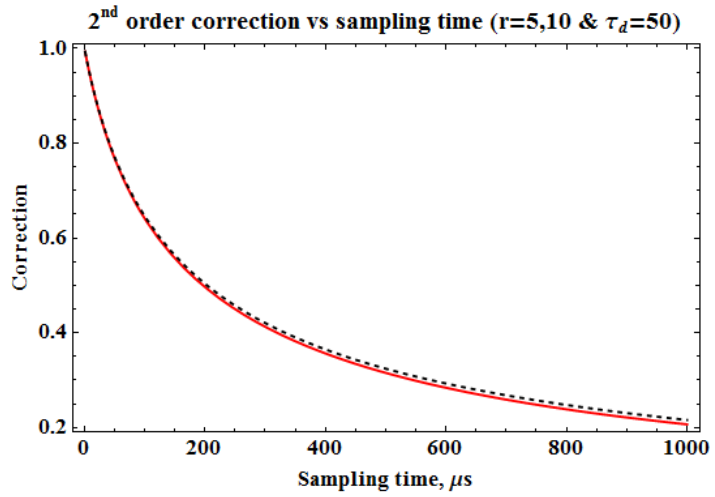


Figure 2-2. The dependence of the binning function values on the aspect ratio *(r)* is shown using the second order function. Doubling the aspect ratio has a minor effect on the correction values towards long lag times.

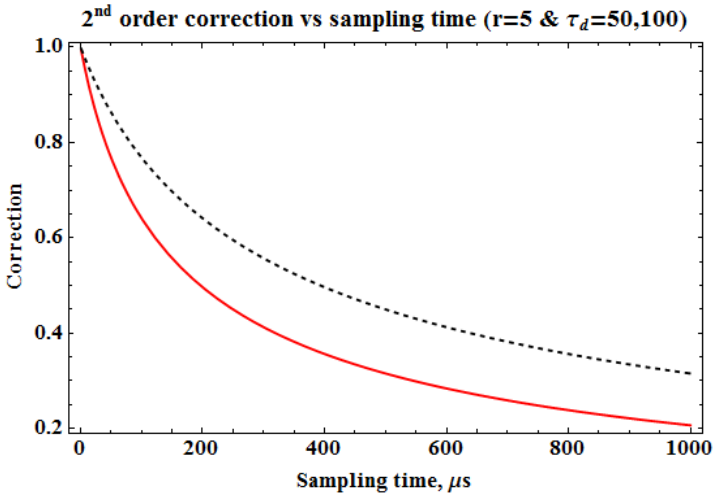


Figure 2-3. Unlike the dependence on aspect ratio, a factor of two difference in the diffusion time will cause a significant deviation in the calculated correction values.

II.3.2.3 Corrections for Detector Dead-time and Afterpulsing Effects

Photodetectors used in FFS experiments possess dead-time and afterpulsing effects. Dead-time refers to a period of time during which the detectors are not able to register a photon count. This occurs right after the detection of a photon and can be thought of as the recovery time of the detectors. In certain types of detectors, called paralyzable detectors, the arrival of a photon during the dead-time postpones the dead-time even longer (54, 55). Non-paralyzable detectors, on the other hand, possess a constant dead-time period.

Afterpulsing is the generation of two or more counts from a single photon detection event. The behavior of afterpulsing varies by detector types (55). In the case of avalanche photo diodes, afterpulses are generated when the released charge carriers get trapped. Their eventual release after a delay time will itself generate additional charge carriers resulting in multiple counts or afterpulses (55).

These two detector behaviors affect the determined photon counts distributions (50, 52, 54-59). In general, while afterpusling tends to increase the counts, dead-time reduces it. The calculations of the factorial cumulants have to take into account both of these effects in order to obtain the correct brightness and population values. The application of dead-time and afterpulsing corrections first requires the afterpulsing probability and dead-time values for the specific detectors to be determined. The afterpulsing and dead-time affected photon count moments must be used to calculate the factorial cumulants from which the proper molecular brightness and population values can be obtained (54, 60).

CHAPTER III

GENERAL METHODOLOGY

III.1 Fluorescence Fluctuation Spectroscopy

III.1.1 Instrumentation

Fluorescence fluctuation spectroscopy experiments were performed on an ISS Alba FCS unit (ISS, Champaign, IL). The instrument is illustrated by the schematic shown in Fig. 3-1. The excitation source consists of a mode-locked Ti:Sapphire laser (Mira 900-F) pumped by a Nd:YVO4 laser (Verdi V-5) (Coherent, Santa Clara, CA) for two-photon excitation at a wavelength of 800 nm. The beam was directed to the side port of an inverted Nikon TE2000-U microscope (Melville, New York) with a 60X (NA = 1.2) with a Plan-Apochromat water

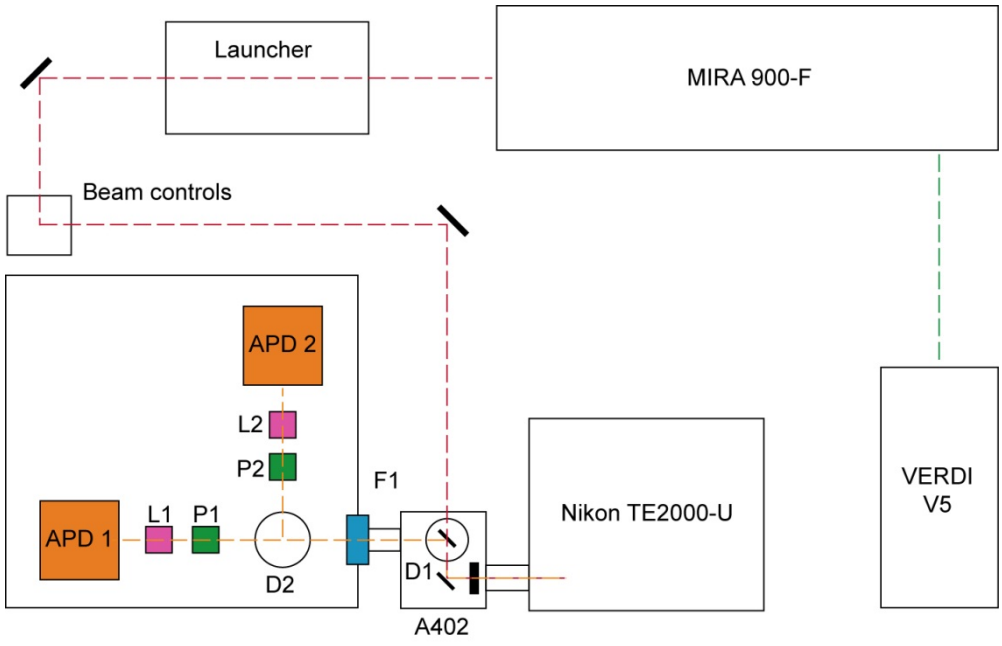


Figure 3-1. A schematic of the ISS Alba FCS instrument.

immersion objective. The launcher unit contains the shutter and the beam can be adjusted through the use of the beam controls. For all measurements unless specified, the power was 20 mW, as determined at a reference point outside the microscope.

The emission was separated from excitation with a dichroic mirror inside unit A402 (700dcxr) and a 25 mm shortpass filter (E700SP-2p, F1 in Fig. 3-1). For FCS experiments this is followed by a beam splitter (BS-21000) and for cumulant analysis, a dichroic/bandpass set (565dcsp/ET610/60-2p) (Chroma, Bellows Falls, VT) which are all located in a dichroic wheel (D2 in Fig. 3-1). For photon detection, the unit consisted of two avalanche photodiodes (APD 1 and APD 2 in Fig. 3-1), (SPCM-AQR-15, Perkin-Elmer, Waltham, MA). Each APD has a lens and a pinhole array (L1/2, P1/2 in Fig. 3-1). Since the experiments are performed with two-photon excitation, a pinhole is not necessary. Finally, output from the detectors was processed with a data acquisition board with time-mode and photon-mode acquisition capability and internal clock frequency settings of 24 and 50 MHz (61).

III.1.2 Data Acquisition

FCS data was acquired in time-mode at 500 kHz sampling rate and an internal clock of 24 MHz and in photon-mode with an internal clock of 50 MHz. The photon-mode data was then correlated at a resolution of 20 MHz. For all FCS experiments, photon counts were recorded with two detectors for two-channel autocorrelation in order to remove detector afterpulsing effects. Cumulant analysis data was collected at 800 kHz in time-mode and 50 MHz internal clock frequency. In this case, it is also important to use a reference molecular brightness value to make certain that the instrument is properly aligned and calibrated in order to maintain consistency between measurements (62). The data acquisition time was 45 – 60 seconds.

III.1.3 Data Analysis

For FCS experiments, calibration of the instrument was first performed using a dilute solution of TAMRA in the appropriate buffer. From these, the observation volume dimensions w_0 and z_0 were determined using a diffusion coefficient of 283.5 $\mu m^2/s$ for TAMRA, estimated from the known value of 300 $\mu m^2/s$ for rhodamine 6G (63) using the relation between diffusion coefficient and molecular weight ($D \propto (\text{molecular weight})^{-1/3}$) (64). During data analysis, these dimensions were held fixed while all other parameters were left free.

All fits were performed with FCS analysis software developed as part of this dissertation work. The software, called FCSLab, was written in object-oriented MATLAB (The MathWorks, Natick, MA). For the iterative minimization of χ^2 during non-linear least squares fitting, the trust-region-reflective algorithm was utilized. The standard deviations of the correlation points were calculated as previously described (65) and used as weights in the fitting. When FCS data have been averaged, the standard deviations have been propagated. Unless otherwise specified, the parameter values presented here are averages of multiple measurements. Correlation curves were normalized to the quantity $(2\sqrt{2} \langle N \rangle)^{-1}$.

Cumulant analysis was performed using software written in MATLAB, also developed as part of this dissertation work. The software (Brightness Analysis Tool) allows the inclusion of the necessary corrections for detector dead-time and afterpulsing as well as corrections for finite sampling time effects. Unless specified, the raw data of photon counts was rebinned to 10 kHz to minimize the effect of detector dead-time.

III.2 Steady-State Fluorometry

For all samples, bulk fluorescence emission was collected with a spectrofluorometer (Photon Technology International, Lawrenceville, NJ). Samples were excited at 535 nm and emission was recorded from 550 nm – 650 nm. Emission peaks at ~ 580 nm are reported. Measurements were taken using samples at a concentration of 2 μM.

CHAPTER IV

T4 LYSOZYME EQUILBIRUM CONFORMATIONAL DYNAMICS AND SUBSTRATE-INDUCED CHANGES

IV.1 Introduction

T4 lysozyme (T4L) is a 164 residue globular protein from Bacteriophage T4. In the process of infecting *Escherichia coli,* the virus utilizes T4L to break down the cell wall of the bacteria. The structure of T4L is composed of two domains that are joined by a long helix (Fig 4-1). The active site cleft, where hydrolysis of the glycosidic linkage takes place, is located at the interface between the two domains.

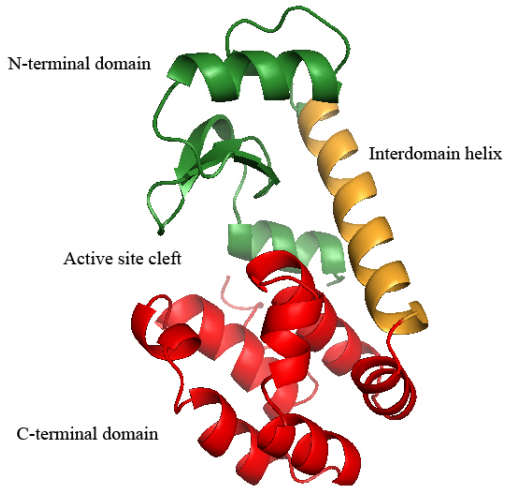


Figure 4-1. The structure of T4 lysozyme consists of an active site cleft in between two domains linked by a long helix (PDB ID 2LZM).

The hypothesis of equilibrium fluctuations in Bacteriophage T4 lysozyme (T4L) has its genesis in crystallographic studies of wild-type (WT) and point mutants. The observation of an occluded active site in early crystal structures of the wild-type (WT) T4L stimulated the conjecture of relative movement between the two domains that enables substrate access.

Subsequent structures with the M6I (66) and I3P (67) mutations gave the first indication of domain motion in T4L (68). Five different conformations were observed for the M6I mutation while the I3P mutation produced two conformations. All of the mutant conformations showed the active site open to various degrees relative to the WT, hinting at domain motion. This motion, referred to as a hinge-bending, consists of a rotation of one domain relative to the other, ranging from about 10 to 32 degrees, about an axis running through the interface of the two domains. This results in opening the active site cleft by as much as 8 Å (68).

EPR analysis of distances between the two domains provided direct evidence of a hinge-bending motion in solution, manifested as equilibrium between open and closed conformations (69). In contrast to the long held view that the closed conformation is energetically more favored, distances between spin label pairs with a member in each domain suggested a substantial population of open conformations in solution. This conclusion was reinforced by distance measurements between the same spin label pairs in a mutant background (T26E) which covalently traps the substrate in the active site. Relative to the WT structure, the crystal structure of this mutant shows a negative hinge-bending angle indicating a closing of the active site (70). The pattern of distance changes obtained by comparison of distances between substrate-bound and substrate-free mutants was in agreement with the closure of the active site cleft upon substrate binding.

In both the M6I and I3P structures, the mutations are located in the hinge-bending region (68). Therefore, it was unclear whether these mutations and/or crystal packing caused structural distortions. The EPR study, along with other mutant structures where the point of mutation is located away from the hinge-bending region, confirmed that the open structures are not merely distortions caused by the mutations and/or crystal packing but represent the consequence of intrinsic hinge-bending motion.

Although the presence of a hinge-bending motion in T4L is now widely accepted, the question of its timescale is still outstanding. For large scale domain motions, a theoretical

estimate of ~ 10 ps – 10 μs has been suggested (8). An upper bound estimate of ~ 160 μs for the domain motion in T4L was deduced from NMR studies of the average structure in solution (71). A number of molecular dynamics studies have also attempted to give a timescale for the domain motion in T4L. However, the computational cost of performing simulations for a reliable duration has made this task difficult. As a result, MD studies have been forced to resort to extrapolation techniques like principal component analysis (PCA) or essential dynamics analysis (EDA) in order to bypass the limited sampling constraint of simulations. In their study of T4L, de Groot et al., utilizing classical MD combined with EDA, observed a transition from the closed to an open conformation within 1 ns (71). Considering that the duration of the simulation was 1 ns and that the choice of force-field type seems to influence the simulations themselves (25), the reliability of the timescale put forth is questionable.

Moreover, fluorescence resonance energy transfer (FRET) has been applied to investigate the conformational dynamics of T4L in the presence of substrate (72, 73). It was observed that the formation of the enzyme-substrate complex occurs incrementally through six intermediate conformations. In addition, the formation of the complex, which involves hinge-bending motion from the open to the closed conformation, occurs in the order of milliseconds. The study pointed out that equilibrium hinge-bending motion (in the absence of substrate) is expected to be faster than the non-equilibrium timescale obtained in the presence of substrate.

A related question pertains to the nature of the equilibrium between T4 lysozyme conformations in solution. Current models, formulated on the basis of crystallographic analysis, postulate a two-state equilibrium between an open and a closed conformation. However, careful inspection of structures in different crystal environments reveals a range of over 50 degrees in hinge-bending angles relating the N- and C-terminal domains (68). Thus, the variability observed in the crystal suggests a continuum of conformations between the two extremes defined above as open and closed. Consistent with this model, MD simulations have identified a range of states

with intermediate opening of the active site (74). Nevertheless, for reasons related to ensemble averaging, NMR and EPR data were interpreted in terms of a two-state equilibrium.

Here, the single molecule aspect of fluorescence fluctuation spectroscopy was used to address the two key questions pertaining to T4L dynamics – the timescale of its hinge-bending motion and the nature of its conformational equilibrium. Both fluorescence correlation spectroscopy (FCS) and cumulant analysis (CA) have been applied in this study. FCS, through the analysis of fluorescence fluctuations, allows the temporal characterization of various processes spanning multiple orders of magnitude in timescale. Since its introduction in the early seventies (29), FCS has been used to examine a variety of systems including the photophysical behavior of fluorescent proteins (39, 45) as well as macromolecular diffusion and interactions (75, 76). In addition, FCS has been demonstrated to be a powerful tool for the investigation of equilibrium protein conformational and folding dynamics (38, 40, 41, 46, 77-80). Hence, this technique is inherently suited to study the type of domain motion observed in T4L with the ultimate goal of obtaining a fluctuation timescale. Cumulant analysis provides molecular brightness and relative population values, complementing information obtained from FCS (28).

To investigate the equilibrium dynamics of T4 lysozyme in solution, the fluorophore TAMRA (Tetramethylrhodamine) was introduced at residue pairs whose relative proximity is altered in the transition between open and closed conformations (68, 69). FCS measurements from multiple sites reveal T4 lysozyme dynamic conformational equilibrium involving opening and closing of the active site. Furthermore, the results establish the presence of multiple intermediate conformations in contrast to the prevailing two-state model.

IV.2 Materials and Methods

IV.2.1 Sample Preparation

T4 lysozyme was expressed and purified as described previously (81). T4L mutants with covalently bound substrate were constructed via the T26E mutation which enables the two sugar moieties of the peptidoglycan to be trapped in the active site (70). MTS-TAMRA, (2-((5(6)-Tetramethylrhodamine)carboxylamino)ethyl Methanethiosulfonate) (Toronto Research Chemicals, Ontario, Canada) (Fig. 4-2) labeling was carried out in 20 mM phosphate buffer pH 7.2 (100mM Sodium Chloride, 0.1mM Ethylenediaminetetra Acetate and 0.02% Sodium Azide) with 20% glycerol (v/v). TAMRA was added in 10-fold excess and the samples were incubated at room temperature for 4 hours and then on ice overnight. Excess TAMRA was removed by extensive dialysis with Spectra/Por MWCO 8000 Dialysis membranes (Sigma-Aldrich, St. Louis, Missouri), desalting using HiTrap column (GE Healthcare, Piscataway, New Jersey) and then size exclusion chromatography column (Superdex75, GE Healthcare, Piscataway, New Jersey).

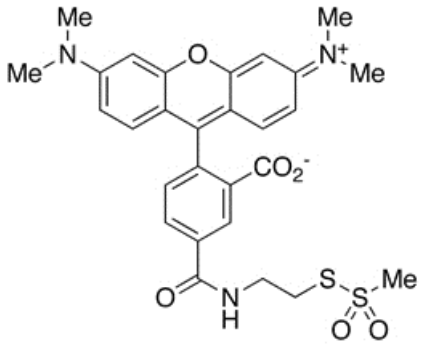


Figure 4-2. Structure of MTS-TAMRA.

Mutants with single TAMRA are designated as T4L-R, where R is the residue number where TAMRA has been introduced. Similarly, mutants with two TAMRA fluorophores are

designated as T4L-R1/R2 and those with covalently bound substrate as T4L-R1/R2S. All spectroscopic measurements were carried out in phosphate buffer.

IV.2.2 Experimental Methods

Fluorescence fluctuation spectroscopy and steady-state measurements were performed as detailed in chapter III. The data acquisition and analysis approaches are also given in chapter III.

IV.3 Results

IV.3.1 Experimental Design

In a typical FCS experiment, an excitation beam is focused by a high numerical aperture objective and directed into a solution of fluorescent molecules of very low concentration (~ nM). Furthermore, utilizing two-photon excitation produces an extremely small (~ fL) excitation volume. Because of the resulting small number of particles, diffusion through the observation volume causes the detected emission to fluctuate. The temporal correlation of this fluctuation, *G(t)*, produces a curve with a decay profile reflecting molecular diffusion. Analysis of *G(t)* with the appropriate FCS model (diffusion-only model, chapter II, Eq. 2.18) yields the corresponding diffusion coefficient. If other kinetic processes cause fluctuations in emission intensity while the molecules are within the observation volume, i.e., if these processes are faster than diffusion, the correlation curve will contain multiple decay profiles or components (38, 39) and FCS models other than the diffusion-only model will be required for analysis.

Hinge-bending motion in T4L results in changes in the relative distance and orientation of the N- and C-terminal domains. To detect the transition between the open and closed conformations in solution, Mchaourab et al. (69) measured short range distances between spin label pairs attached at sites selected to provide a distinctive pattern of distances for each conformation. In Fig. 4-3 *A*, the closed (top row) and open (middle row) structures have been

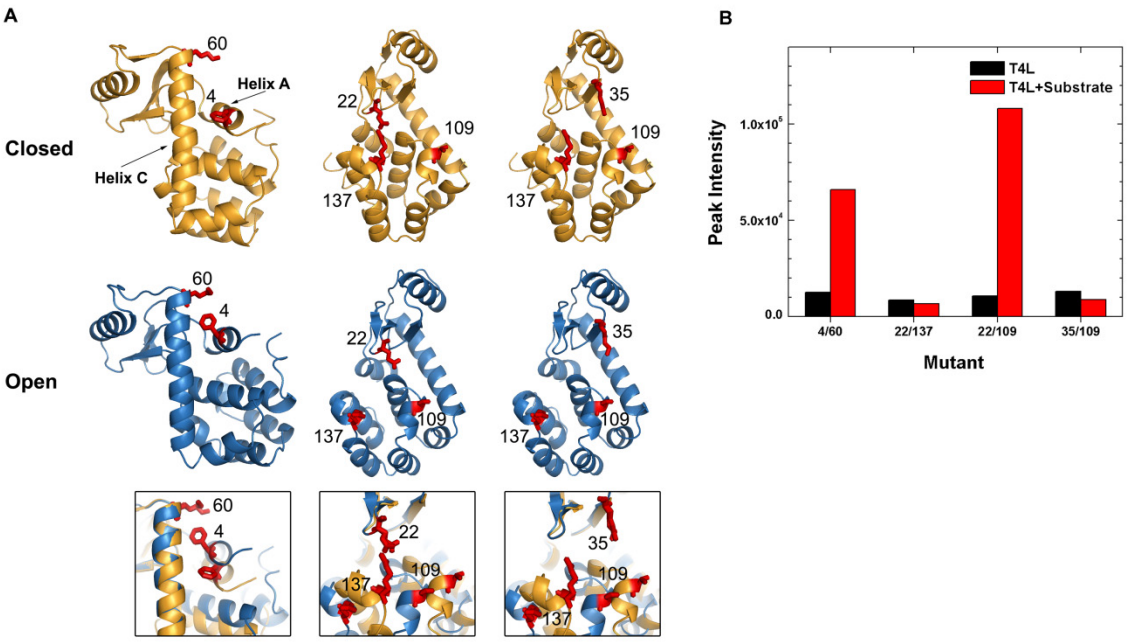


Figure 4-3. *(A)* Crystal structures of T4L in the closed (WT PDB ID 2LZM) and open (M6I mutant, PDB ID 150L, chain c) conformations are shown in the first and second rows respectively. The residues selected for labeling with TAMRA are shown in red. The bottom row is a close up view highlighting proximity changes in the selected residue pairs (Table 4-1). *(B)* Intensity changes from analysis of steady-state fluorescence (576 nm – 581 nm) at the selected residue sites, of the same pair in the WT and T26E backgrounds, demonstrating that TAMRA fluorophores sense the shift in equilibrium between open and closed conformations.

TABLE 4-1 C_α distances in the open and closed conformations of residue pairs selected for labeling with TAMRA, shown in Fig. 4-3 *A*

Mutant	Open (Å) *	Closed (Å) †
4/60	10.5	14.5
22/109	15.3	17.3
22/137	17.6	12.0
35/109	17.4	16.1
35/137	25.3	17.0

*Mutant (M6I) structure, PDB ID 150L (chain c)
†WT structure, PDB ID 2LZM

aligned (residues 15 – 60) to depict distance changes predicted to occur at these pairs (Table 4-1). For instance, residues 22 and 109 are separated by about 15 Å in the open conformation versus 17 Å in the closed conformation.

We reasoned that these pairs, which report hinge-bending by EPR spectroscopy, would show modulation of fluorescence by FCS if the appropriate fluorophore is used. Specifically, the fluorescence intensity should be quenched as a function of proximity, thus enabling the observation of the interconversion between open and closed conformations. For this purpose, we selected the fluorophore Tetramethylrhodamine (TAMRA) which has previously been shown to undergo self-quenching at short distances (77). Furthermore, this fluorophore is a derivative of rhodamine which belongs to the family of xanthenes. Multiple studies have examined the photophysical properties of rhodamines at considerable length demonstrating that they are able to form ground-state complexes, commonly referred to as H-dimers (82-85). H-dimers have two signature characteristics: a severe fluorescence quenching, of 90% or more, and a blue-shifted absorbance peak (86, 87). The distances reported in Table 4-1 suggest that the TAMRA pairs, which are expected to project towards each other, will be close enough for self-quenching and possibly for the formation of dark dimers.

To test the suitability of TAMRA for detection of T4L open to closed transition, we carried out steady-state fluorescence intensity analysis of the pairs reported in Table 4-1. The closed conformation was stabilized by the covalent trapping of the substrate in the T26E mutant background. We observed (Fig. 4-3 *B*) that 1) the distances between the sites in Table 4-1 allow for extensive quenching of fluorescence in one conformation and 2) that these pairs report the domain motion induced by covalent substrate binding leading to large intensity changes. These results predict intermittent quenching of TAMRA fluorescence as a result of T4L hinge-bending motion. If this fluctuation occurs with a relaxation time faster than the time of diffusion for T4L,

then it can be detected by FCS. Analysis of the correlation curve with the appropriate model (chapter II) yields characteristics of the underlying protein dynamics.

Excitation-driven processes, namely intersystem crossing and flickering, reported for some fluorophores and fluorescent proteins are two examples of kinetic processes (45, 88) that could cause intensity fluctuations. In order to exclude excitation-driven processes for TAMRA, a set of FCS measurements at various power levels was taken (Fig. 4-4, Table 4-2). The correlation curves showed only a single decay profile belonging to the diffusion of fluorophores through the observation volume. Thus, TAMRA does not show excitation-driven behavior.

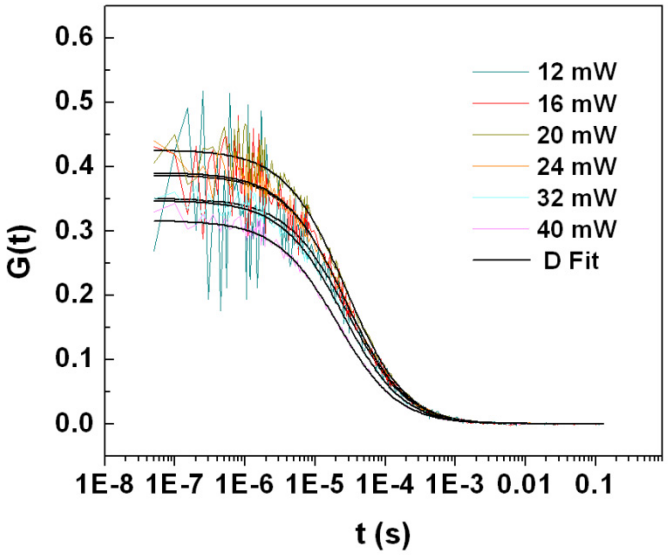


Figure 4-4. Correlation curves for TAMRA in phosphate buffer at power levels in the range of 12 – 40 mW (Table 4-2, below) are shown. Raw data, acquired in photon-mode with a 50 MHz internal clock, was correlated at 20 MHz. The curves have been fit to the diffusion-only model (black) and the corresponding parameters are given in the Table 4-2 (*D* was fixed to 283.5 $\mu m^2/s$). The single decay profile of these curves shows the lack of excitation-driven behavior for TAMRA.

TABLE 4-2 Parameters for diffusion-only model fits of correlation curves for TAMRA at various power levels, shown in Fig. 4-4

Power, mW	<N>	w_0 (μm)	z_0 (μm)	Reduced χ^2
12	0.999 ± 0.035	0.259 ± 0.009	1.50 ± 0.40	0.80 ± 0.08
16	0.916 ± 0.011	0.252 ± 0.003	1.54 ± 0.29	0.80 ± 0.08
20	0.834 ± 0.007	0.248 ± 0.002	1.46 ± 0.10	0.72 ± 0.08
24	0.897 ± 0.006	0.243 ± 0.003	1.52 ± 0.18	0.75 ± 0.04
32	1.012 ± 0.004	0.235 ± 0.002	1.33 ± 0.17	0.74 ± 0.05
40	1.132 ± 0.013	0.218 ± 0.001	1.68 ± 0.25	0.74 ± 0.05

IV.3.2 FCS Reveals T4L Equilibrium Hinge-Bending Motion

Detection of hinge-bending motion relies on the ability to distinguish between intermittently quenched and non-quenched fluorophores from the correlation curves. In the case of non-quenched pairs and the equivalent case of a single fluorophore, the correlation curve is expected to consist of a single component arising from molecular diffusion through the observation volume. To measure T4L diffusion by FCS and establish the single component nature of the corresponding correlation curve, we analyzed the double mutants T4L-61/135 and T4L-65/135 (Fig. 4-5) where the residue separation of about 35 Å is too large for fluorescence quenching. Consequently, the correlation curve for these mutants with a non-interacting TAMRA pair is identical to the curve of T4L-151 with a single TAMRA, showing a single decay component corresponding to diffusion of the molecules through the observation volume. The correlation curves are fit by the diffusion-only model (red curves) (chapter II). The determined diffusion coefficients were 85 $\mu m^2/s$ for the single and 82 $\mu m^2/s$ for the double mutants, in good agreement with the expected value of 88 $\mu m^2/s$ calculated through the relation between molecular weight and diffusion coefficient, i.e., $D \propto (\text{molecular weight})^{-1/3}$ (64).

The molecular brightness values determined for these mutants are given Fig 4-18. The cpsm value for T4L-151 is near that of a single TAMRA, as expected. For the doubly labeled mutants, the brightness determined is approximately 1.5 that of the value for a single TAMRA. In this case, the expected cpsm would be twice that of a single TAMRA. One explanation for this could be that not all of the T4L molecules are doubly labeled but instead, there are populations with a single fluorophore. As a result, the brightness value reflects an effective value.

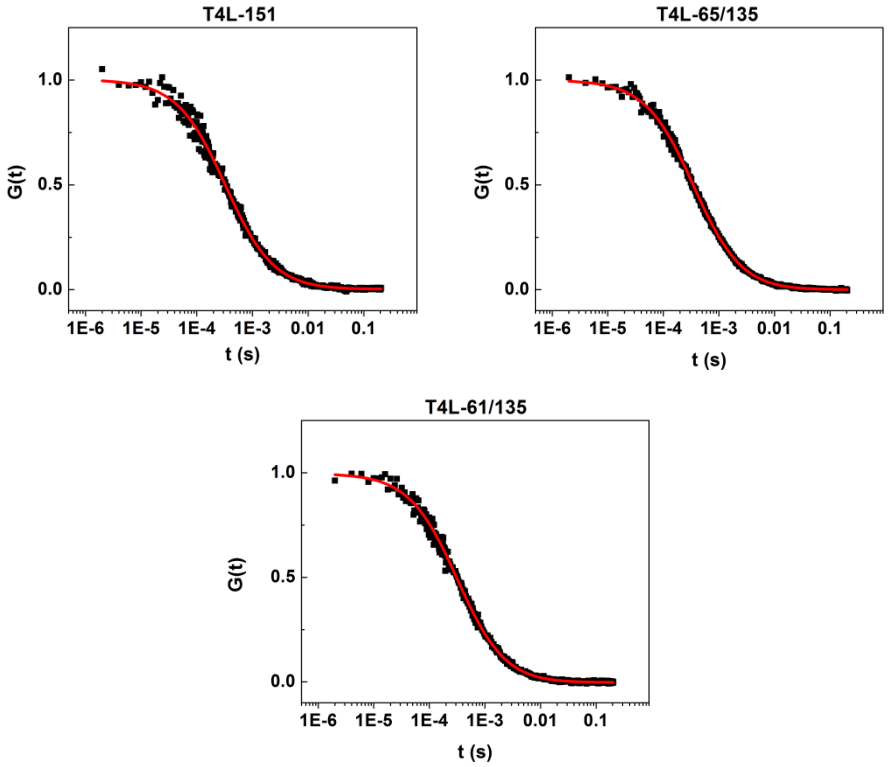


Figure 4-5. Representative correlation curves for singly (T4L-151) and doubly (T4L-65/135, T4L-61/135) labeled T4L mutants are shown. The curves consist of a single decay component arising from the diffusion of molecules through the observation volume. A fit to the diffusion-only model, Eq. 2.18, is superimposed (red curves).

In Fig. 4-6, the correlation curves for double mutants monitoring the active site cleft or the interface between helices A and C (residues 4/60) (Fig. 4-3 *A*) are shown. In stark contrast to the curves shown in Fig. 4-4, these curves display a distinctly multi-component profile. As expected, the diffusion-only model fails to fit the correlation curves (red curves); the lack-of-fit is particularly visible at short delay times.

In each case, a prominent second decay component, well separated from that of diffusion, is observed. This second decay component occurs in a faster time regime than diffusion. The diffusion/quenching model (chapter II, Eq. 2.24) provides a description for this additional process as highlighted by the fits of the correlation curves (green curves). The model approximates the kinetic process by equilibrium between two states. One of the states is assumed to be non-fluorescent arising from a T4L conformation or set of conformations wherein the fluorophores' separation enables substantial self-quenching (chapter II). In contrast, the second state is fluorescent and corresponds to conformation(s) of the protein where the fluorophores are separated by enough distance to reduce or prevent fluorescence quenching.

The extracted relaxation time values, fast-component amplitudes and non-fluorescent state populations that characterize the dynamics are given in Table 4-3. Fit parameters along with the diffusion coefficients and reduced chi-square values are given in appendix A. For these double mutants, a slight increase can be seen in the determined diffusion coefficient values, relative to those of the control mutants (Fig. 4-5). This deviation is likely due to the fluorescence quenching present in the double mutants but not the control mutants, as discussed in Nagy et al. (34).

The robustness and uniqueness of the fits are demonstrated by well-defined minima in chi-square surface search (89) (Fig. 4-7). Except for the 35/109 pair, the relaxation time values cluster in the range of 10 – 20 μs. The observed variability most likely arises from minor perturbations due to the incorporation of the fluorophore. This is particularly likely for 4/60 where Phe 4 engages in packing interactions in the open conformation. Although the substitutions do not perturb the overall fold or inhibit activity as evidenced by covalent trapping of substrate (see below), they may lead to minor shifts in relative stability of substates or alter the corresponding energy barrier for the transition. The non-fluorescent state population in Table 4-3 represents the fraction of T4L molecules where the TAMRA pairs are close enough to have a fully or substantially reduced fluorescence. Depending on the residue pair, this population may

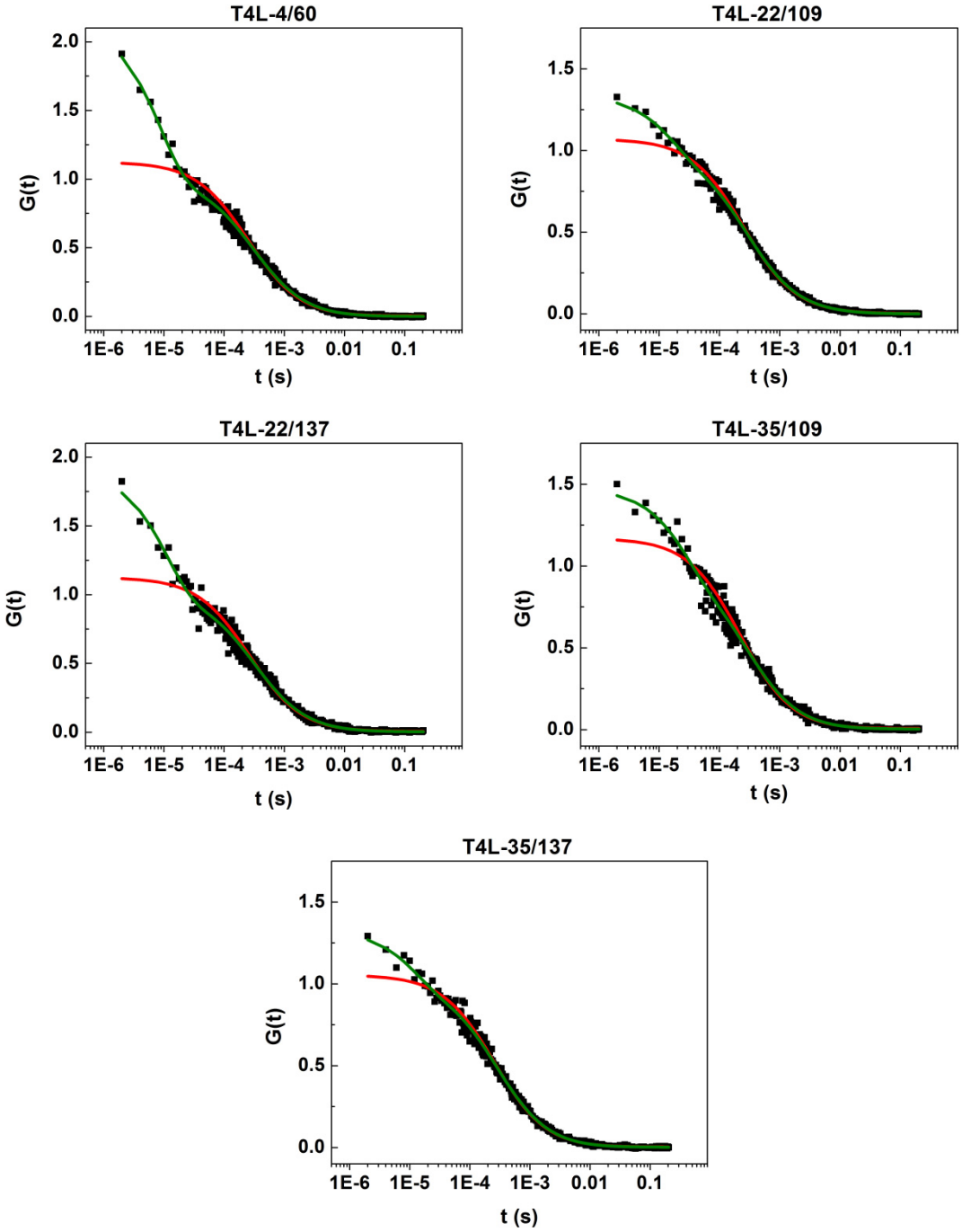


Figure 4-6. Correlation curves for T4L doubly labeled at residue sites shown in Fig. 4-2 *A*. These pairs monitor the active site region and the interface between helices A and C. All of the curves contain an additional decay component not observed in the control mutants (Fig. 4-5). The time regime for the new decay component is faster than that of diffusion with the best fit for these curves obtained from the diffusion/quenching model (green, Eq. 2.24). The corresponding parameters are given in Tables 4-3. As expected, the diffusion-only model (red, Eq. 2.18) yields a poor fit for short time delays.

TABLE 4-3 Parameters obtained from diffusion/quenching model fits of the correlation curves for the dynamic regions of T4L

Mutant	τ_R (μs)	Amplitude [Non-fluorescent state fraction (%)]
4/60	8.5 ± 1.1	1.20 ± 0.10 [54.4]
22/109	22.6 ± 11.1	0.33 ± 0.04 [25.0]
22/137	13.3 ± 2.7	0.78 ± 0.08 [43.7]
35/109	36.3 ± 11.1	0.54 ± 0.06 [34.9]
35/137	10.2 ± 2.1	0.37 ± 0.04 [26.7]

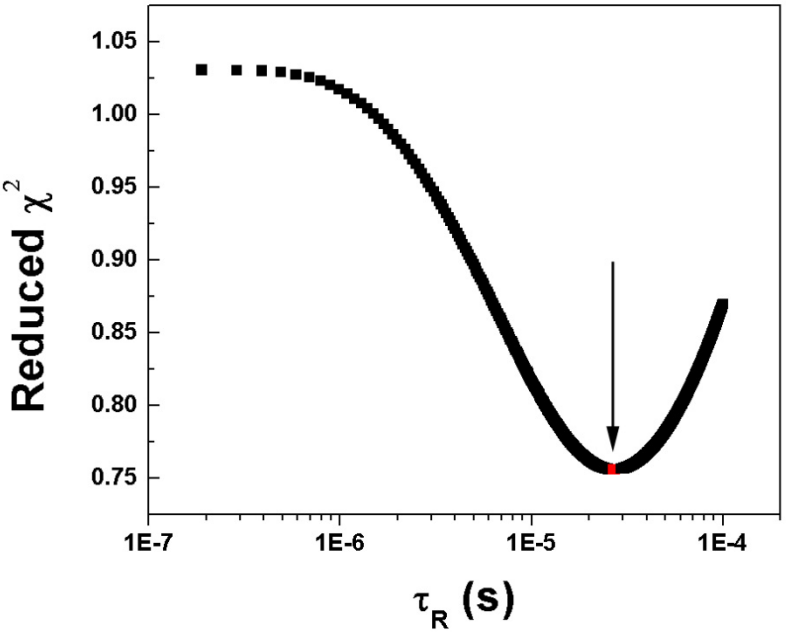


Figure 4-7. A representative chi-square dependence of the relaxation time for T4L-35/109, as determined by chi-square surface search, with the minimum shown in red.

be associated with the open or closed conformations. We note that the fraction of non-fluorescent molecules is determined relative to the total number of particles, <*N*>, (chapter II). Thus, <*N*> represents molecules that transition between the fluorescent and non-fluorescent states during diffusion through the observation volume.

Additional data was collected from two T4L mutants with TAMRA pairs in the vicinity of the helix A/C interface, T4L-5/60 and T4L-4/71 (Fig. 4-8). However, the correlation curves for these mutants are markedly different from that of T4L-4/60 (Fig. 4-6). The fast decay component is reduced for T4L-5/60 and completely lacking for T4L-4/71. These results taken together may indicate the role played by the orientation of the fluorophores; for the case of T4L-4/60, it is possible that the fluorophores are able to take relative orientations that allows for the type of interaction that leads to a prominent fast decay component.

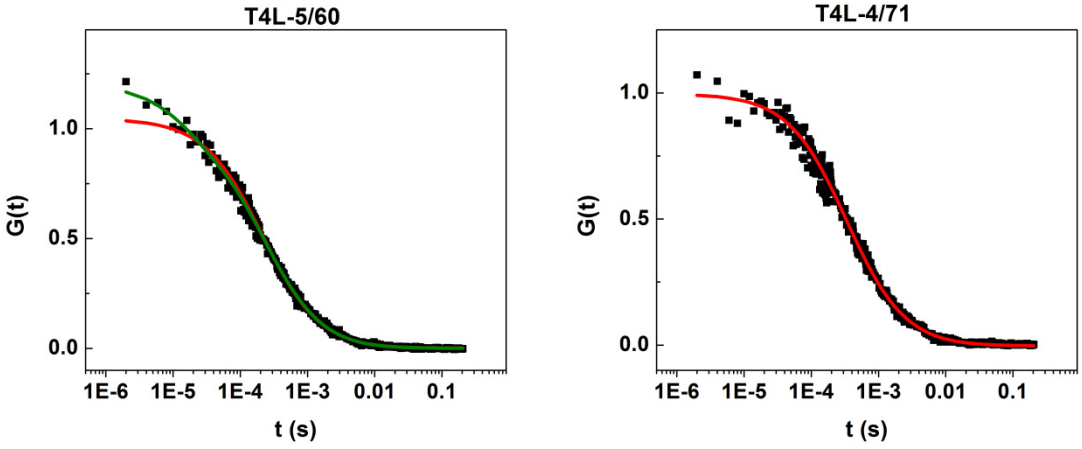


Figure 4-8. Representative correlation curves for T4L mutants monitoring the motion of helices A and C. Unlike the comparable pair 4/60 (Fig. 4-6), the fast component amplitude of 5/60 is visibly smaller and lacking in 4/71. These curves possibly point to the importance of the position of the fluorophore on the protein in sensing the underlying motion.

Cumulant analysis for these mutants shows brightness values that are lower than the value of a single TAMRA (Fig. 4-18). Thus, all indicate a considerable amount of quenching. In the context of the two-state FCS analysis, under the assumption that one of the states is fully quenched, the molecular brightness values would suggest that the fluorescent state itself has a considerable amount of quenching.

In order to further ascertain that the observed fluctuations are a result of domain motion associated with opening of the active site, FCS data on double mutants separated by a similar distance was also collected from a region of the protein unaffected by the crystallographic open to closed transformation. In the selected pairs, T4L-131/135, T4L-151/154, T4L-131/151 and T4L-131/154, the two fluorophores are located in close proximity such that their steady-state fluorescence is severely quenched (Fig. 4-19). Representative correlation curves for these mutants along with their fits are shown in Fig. 4-9. These mutants do not show a substantial second decay component that resembles the ones observed for the active site region or the interface between helices A and C. Although an incipient second component is observed at very short delay times for the three mutants other than T4L-151/154, it appears to reflect faster motion, possibly local backbone dynamics that modulates the proximity of the fluorophores. FCS analysis performed at a higher temporal resolution by increasing the sampling rate from 500 kilohertz (kHz) to 20 Megahertz (MHz) reveals that the second component primarily reflects sub-microsecond motion which will be discussed in section IV.3.4. Thus, the second component in pairs such as 22/137 cannot arise from the "flopping" of the fluorophores. The relaxation time for such fluctuations is expected to be on the order of nanoseconds or shorter, well faster than what is detectable in our measurements (38). The molecular brightness values show that there is severe quenching for all of these mutants indicating that the fluorophores do come into contact (Fig. 4-18).

Having excluded dynamics of secondary structures and/or the fluorophore pairs and excitation-driven photophysics of TAMRA, these findings indicate that the microsecond

relaxation time fluctuation observed for the active site region as well as the interface between helices A and C is a result of domain motion associated with transition between open and closed conformations.

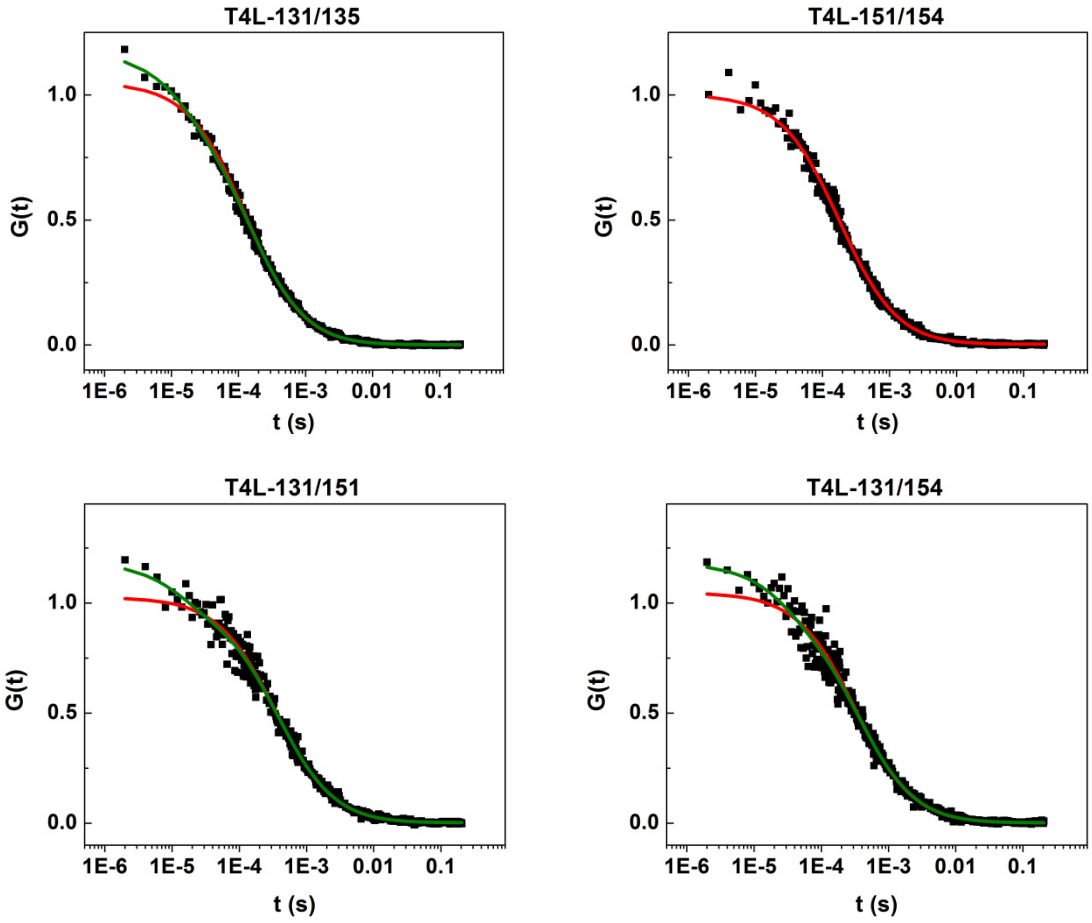


Figure 4-9. Correlation curves for T4L mutants with the fluorophore pair separated by less than 11 Å to allow for fluorophore contact. Curves for 131/135, 131/151 and 131/154 are best fit to the diffusion/quenching model (green) indicating the presence of an incipient fast component which is attributed to local structural dynamics. The curve for 151/154 can be fit to the diffusion only model (red).

IV.3.3 Substrate Induced Shifts in T4L Equilibrium Fluctuations

To confirm that the second component in the correlation curves reports hinge-bending dynamics, we carried out FCS measurements in the T26E mutant background. It has been shown that introducing the mutation T26E covalently traps the basic component of *E.coli* peptidoglycan – the two sugar moieties (N-acetyl/muramic acid or NAM and N-acetyl/glucosamine or NAG) and a peptide chain (70) in the active site during the purification of the protein from *E. coli*. In the crystal structure, the NAM component is covalently linked to the introduced glutamate. The structure determined by Kuroki et al. of this T4L-substrate complex showed a negative hinge-bending angle where the two domains further close the active site relative to the structure of the WT. Furthermore, an EPR study of T4 lysozyme (69) has shown that substrate binding selectively stabilizes the closed conformation in solution. Therefore, within the framework of the two-state FCS model, the conformation of this substrate-bound form of T4L makes up one of the fluctuation states. FCS experiments under this condition would provide us with a reference for the equilibrium measurements.

The mutants showing microsecond relaxation time fluctuations were prepared with a covalently bound substrate via the T26E mutation (section IV.2.1). FCS measurements were carried out under the same conditions as the substrate-free mutants. The correlation curves for these mutants are shown in Fig. 4-10 along with the corresponding fits. For all of the mutants (except 4/71), the curves show a decrease in the fast component amplitude relative to that of the substrate-free mutants (Fig. 4-6, Table 4-4). Conceptually, the lack of a fast component in a correlation curve corresponds to a lack of fluorescence fluctuations that are faster than those caused by diffusion. In the limit where all the T4L molecules trap their substrate, the contribution of hinge-bending motion should be frozen out. Hence, the protein is no longer undergoing structural fluctuations, the source for the fast decay component in the correlation curves. The decrease in the fast component amplitude is therefore in agreement with the shift in equilibrium expected upon substrate binding.

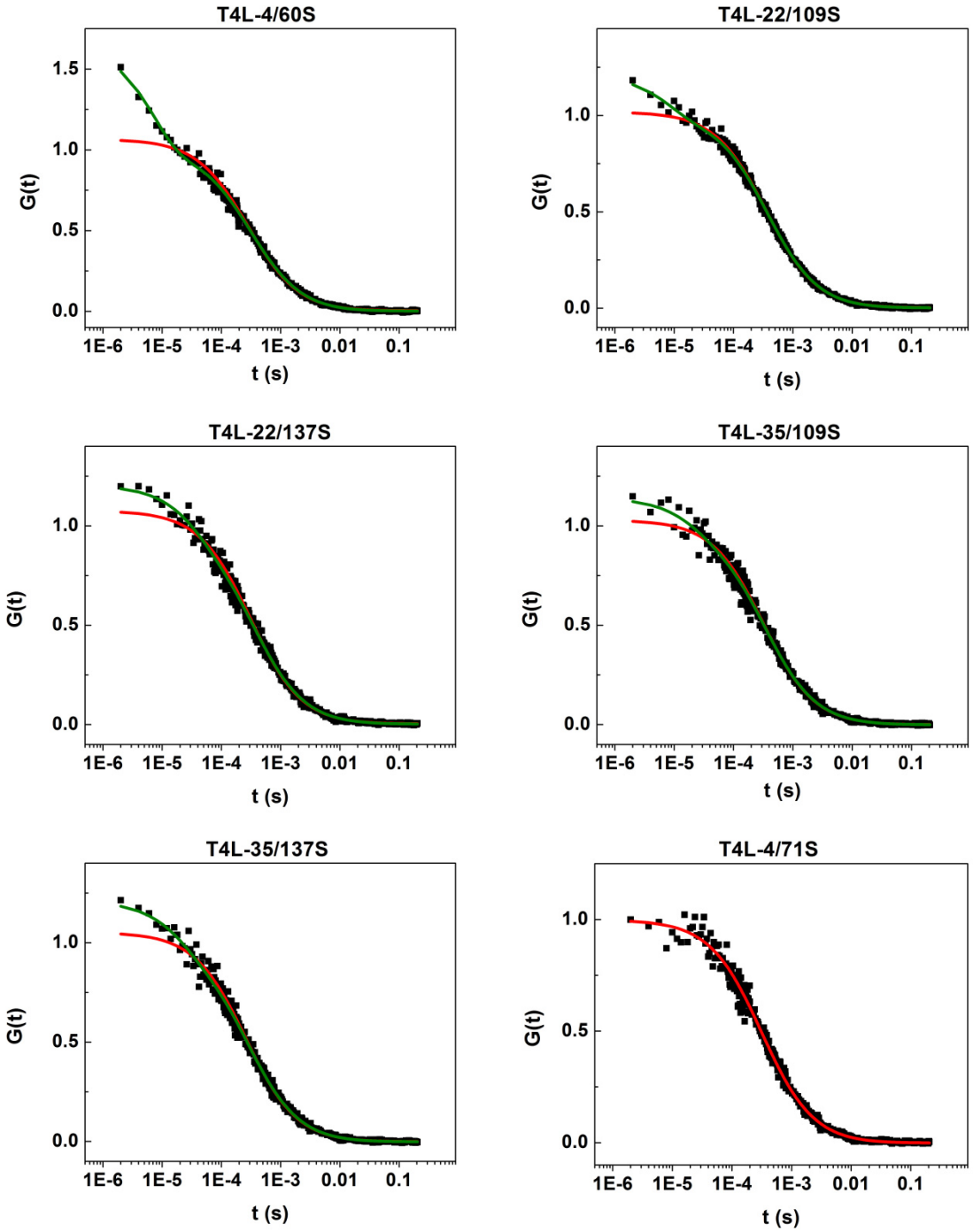


Figure 4-10. Correlation curves for T4L double mutants in the T26E background. The curves were fit to the diffusion-only and diffusion/quenching models, red and green lines respectively. With the exception of 4/71, relative to the substrate-free mutants (Fig. 4-6), these curves show a decrease in the fast component amplitude (Table 4-4), consistent with substrate-induced dampening of hinge-bending dynamics. The lack of complete suppression of the fast decay component in the correlation curves is likely due to the presence of T4L populations without substrate.

TABLE 4-4 Fast component amplitudes representing fluctuations between open and closed conformations from diffusion/quenching model fits of the correlation curves for the dynamic regions

Mutant	Amplitude (substrate-bound)	*Amplitude (substrate-free)*
4/60S	0.61 ± 0.04	*1.20 ± 0.10*
22/109S	0.22 ± 0.07	*0.33 ± 0.04*
22/137S	0.15 ± 0.05	*0.78 ± 0.08*
35/109S	0.11 ± 0.03	*0.54 ± 0.06*
35/137S	0.21 ± 0.03	*0.37 ± 0.04*

The changes observed for the molecular brightness values are also in agreement with this conclusion (Fig. 4-18). For mutants other than T4L-4/60 and T4L-22/109, there is a decrease in the molecular brightness values in agreement with the closing of the active site region. In the case of T4L-4/60, substrate binding increases the distance between sites 4 and 60, hence the increase in brightness. The separation distance of sites 22 and 109 also increases with substrate binding, leading to increased brightness. Thus, these results directly show that the structural changes in the active site region and the interface between helices A and C caused by substrate binding correspond to a dampening of protein dynamics.

It has also been shown (69, 70) that acid hydrolysis at 37°C reverses the T26E induced closure by removing the substrate product. FCS experiments were also performed under these conditions for T4L-4/60S and a representative curve is shown in Fig. 4-11. The data was best fit to the diffusion/quenching model and gave a dark population of approximately 48% compared to the initial value of ~ 38 %. This shows that the substrate has been successfully removed by the acid hydrolysis steps taken, increasing the dynamic T4L population.

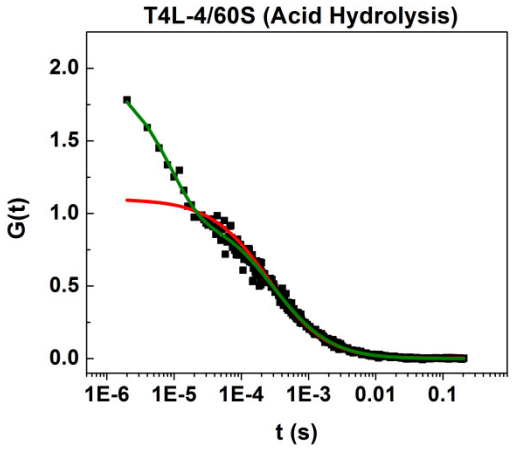


Figure 4-11. Through acid hydrolysis, mutant 4/60 in the T26E background shows a correlation curve with larger fast component amplitude indicating increased dynamics.

While the above experiments clearly show that the T26E background significantly attenuates the second component in the decay curves, the smaller, albeit discernible, fast decay component may reflect the presence of a small fraction of molecules without substrate. A similar incomplete trapping of substrate was also reported by Mchaourab et al. (69) and independently verified for TAMRA-labeled 4/60 by mass spectrometry and SDS-PAGE analysis .

IV.3.4 Dynamics of the Inter-domain Helix

The two domains of T4 lysozyme are connected by a helix 20 residues long (residues 61-80). Hinge-bending motion entails changes in the backbone torsion angles at the N and C-terminal residues of this helix (68). However, comparison of the open and closed crystal structures do not reveal alterations in the helix backbone.

With the aim of examining the dynamics of this helix in solution, a set of four double mutants spanning its lengths were prepared: T4L-65/69, T4L-65/72, T4L-65/76 and T4L-65/80. Fig. 4-12 *A* shows a correlation curve for T4L-65/69, which reports on the first two turns of the

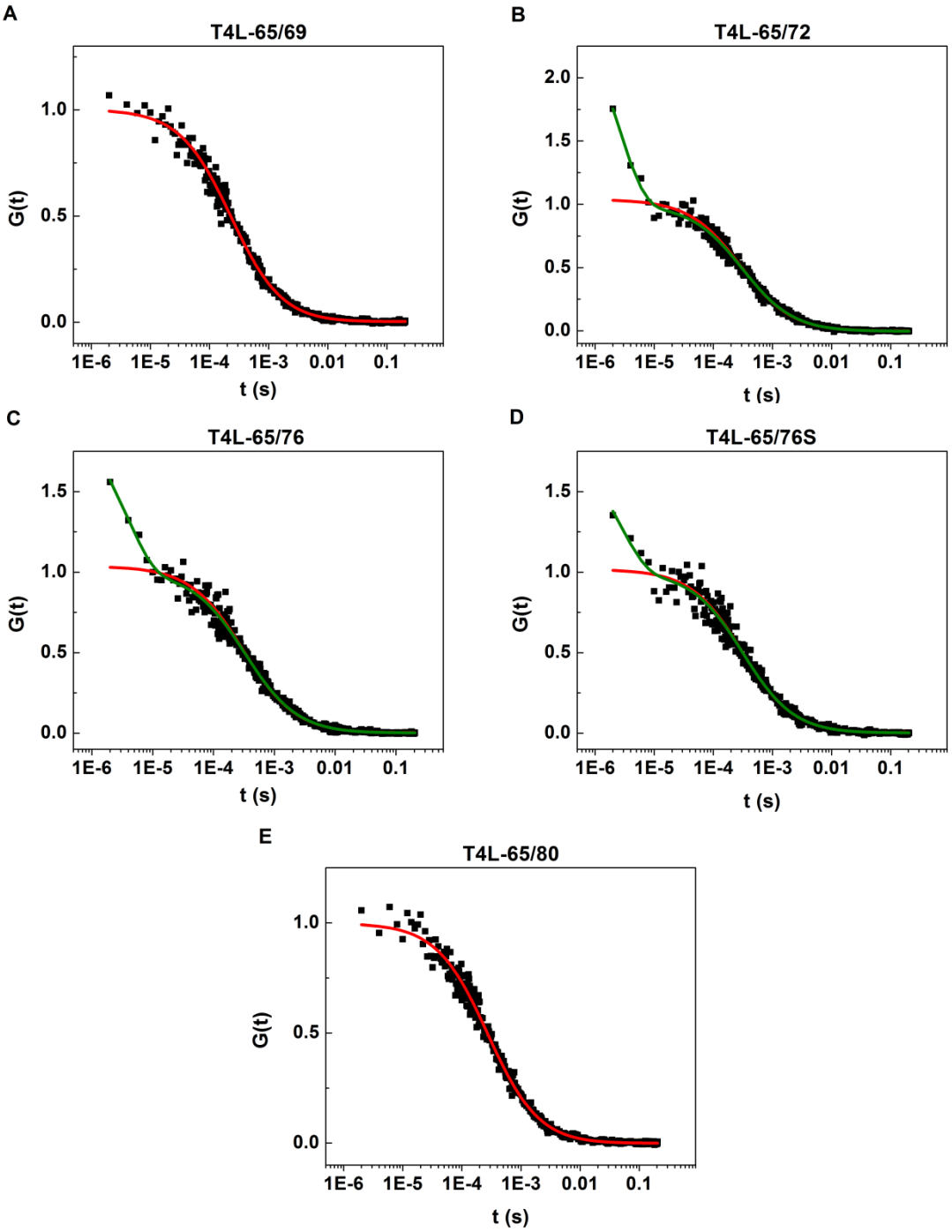


Figure 4-12. Dynamics of T4L inter-domain helix. *(A)* The single component correlation curve for T4L-65/69 is fit to the diffusion-only model (red). *(B)* and *(C)* Correlation curves for T4L-65/72 and T4L-65/76, fit to the diffusion/quenching model (green), indicate the presence of dynamic modes along the helix. *(D)* The correlation curve for the substrate-bound T4L-65/76, fit to the diffusion/quenching model (green), shows that the dynamics is not dampened by substrate binding. *(E)* T4L-65/80 shows a single component correlation curve similar to T4L-65/69.

helix. The shape of this curve is distinct from that observed for the dynamic regions of the protein in that it lacks a fast decay component and can be fit to the diffusion-only model. In contrast, a prominent fast component is clearly visible in the correlation curve for T4L-65/72 as shown in Fig. 4-12 *B*. The curve for T4L-65/76 (Fig. 4-12 *C*) is similar, with large fast-component amplitude. Considering that the two residues are separated by more than 15Å (three helix turns), this result implies significant bending of the helix. Interestingly, substrate binding did not decrease the fast component amplitude in T4L-65/76 (Fig. 4-12 *D*). Finally, the correlation curve for T4L-65/80 (Fig. 4-12 *E*) does not show a fast decay component as observed for T4L-65/72 and T4L-65/76 and can be fit to the diffusion only model. Cumulant analysis shows that, like the other three mutants, the molecular brightness for T4L-65/80 is lower than that of a TAMRA. Thus, similar to T4L-4/71, although the correlation curves do not show a fast decay component, the fluorophore pairs are still able to come into contact to be quenched.

TABLE 4-5 Diffusion/quenching model fit parameters of correlation curves for T4L mutants along the inter-domain helix

Mutant	τ_R (μs)	Amplitude [Non-fluorescent state fraction (%)]
65/72	3.7 ± 0.4	1.26 ± 0.19 [55.6]
65/76	4.3 ± 1.0	0.89 ± 0.16 [46.8]
65/76S	2.3 ± 0.5	1.07 ± 0.21 [51.2]

Together, the correlation curves for T4L-65/72 and T4L-65/76 are the first reported evidence that the long helix is dynamic. The observed fluctuations are faster than those associated with the hinge-bending motion as evidenced by their relaxation time values (Tables 4-3 and 4-5). The fact that substrate binding did not reduce the fast component amplitude indicates that the motion persists even in the closed conformation. A flexible inter-domain helix may be required to facilitate the opening and closing of the active site

IV.3.5 Detection of T4 Lysozyme Sub-microsecond Dynamics

The FCS results presented in the presence and absence of substrate establish a microsecond relaxation time for T4 lysozyme domain dynamics. Furthermore, FCS measurements clearly capture the substrate-induced structural changes across the active site and the interface between helices A and C. However, other dynamic modes are expected and their presence is evident in incipient second components (example T4L-131/135 in Fig. 4-9 *A*) observed at short time delays, suggesting a faster relaxation time. Therefore, additional FCS experiments were performed at a higher temporal resolution by increasing the resolution of the correlation from 500 kilohertz (kHz) to 20 Megahertz (MHz).

Fig. 4-13 *A* shows that the correlation curve for T4L-61/135 at the higher sampling rate has a similar shape to the one obtained at the lower sampling rate (Fig. 4-4 *B*), both of which can be adequately fit to the diffusion-only model. Because the two fluorophores are far apart, there is no mechanism for fluorescence quenching. In contrast, extending the correlation curve of T4L-22/137 to shorter time delays uncovers additional fluctuations (Fig. 4-12 *B*). The correlation curve has been fit to the diffusion/quenching model as well as a double-exponential diffusion/quenching model (chapter II, Eq. 2.25). With increased sampling rate, the correlation curves can no longer be adequately fit to the two-state diffusion/quenching model. Correlation data was also collected for the control mutants T4L-131/135 and T4L-151/154 (Fig. 4-12, *C-D* respectively). As in the case of T4L-22/137, T4L-131/135 is best fit to the double-exponential model. Unlike the single-exponential diffusion/quenching model which is derived for a two-state system, the double-exponential model is phenomenological. Thus, the two processes are not assumed to be coupled nor do they represent equilibrium between dark and fluorescent populations. The fit to the double-exponential model was verified by a chi-square surface search for the second relaxation time (89). A representative figure from this analysis (Fig. 4-13) shows a clear minimum for the second relaxation time parameter (red), in the sub-microsecond regime.

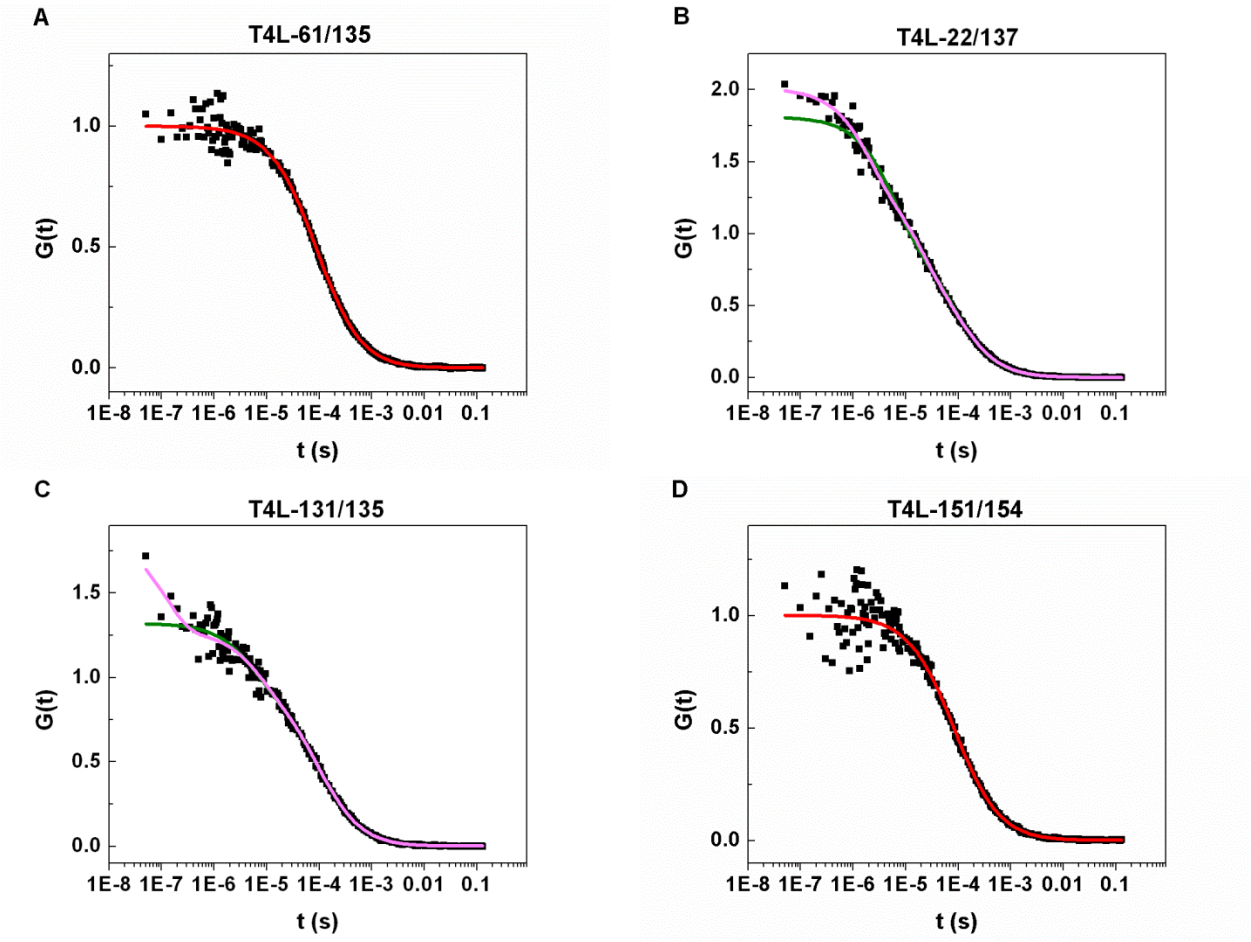


Figure 4-13. Representative correlation curves at 20 MHz resolution, for *(A)* the doubly labeled control mutant shown in Fig. 4-5, *(B)* a location in the active-site region of T4L, and *(C)*-*(D)* mutants with TAMRA pairs in close proximity. The curves in *(A)* and *(D)* are fit to the diffusion-only model (red). In *(B)* and *(C),* the increased resolution has revealed additional fluctuations and the curves are best fit to the double-exponential model (magenta, Eq. 2.25) instead of the single-exponential model (green, Eq. 2.24).

TABLE 4-6 Parameters from the double-exponential model fit of the 20 MHz correlation curves

Mutant	τ_{R1} (μs)	A_1	τ_{R2} (μs)	A_2
22/137	15.1 ± 4.3	0.36 ± 0.04	2.2 ± 0.4	0.59 ± 0.07
131/135	11.5 ± 5.2	0.20 ± 0.04	0.4 ± 0.2	0.49 ± 0.19

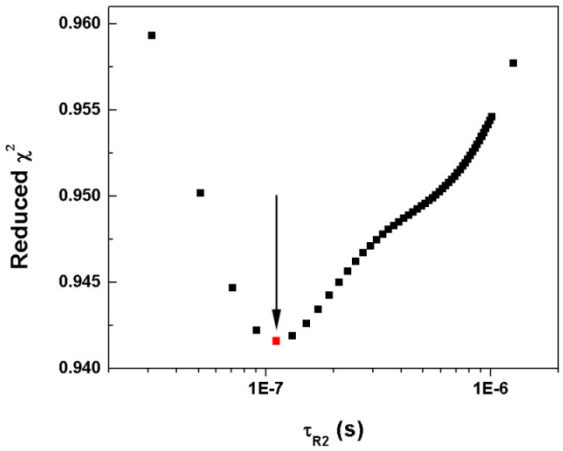


Figure 4-14. Chi-square dependence of the second relaxation time, $\tau_{R2,}$ for T4L-131/135 shows a well-defined minimum in the sub-microsecond time regime.

The faster processes that give rise to the new component in the correlation curve of T4L-131/135 are likely to reflect local dynamics. For T4L-151/154, the fluorophores are in closest proximity with both located on the same helix. In both cases, one source of a second decay component would be the orientational freedom of the fluorophores themselves.

Thus, the result shows that fluorophore dynamics alone does not produce the type of second decay component observed for the active site region and the helix A/C interface of the protein. On the other hand, the curve for T4L-131/135 shows clear evidence for fast fluctuation processes. However, in this case, one of the fluorophores is located on a helix and the other on a loop structure. Thus, the presence of a second fast decay component is likely a result of local dynamics that modulate the proximity of the TAMRA pair. Similar results for T4L-131/151 and T4L-5/60 can be found in appendix A.

Overall, the results presented here for the control mutants are distinct from those of T4L-22/137 and reinforce the conclusion from data obtained at the lower sampling rate (Figs. 4-5 and 4-8). The fluctuation for the active site region is characterized by relaxation times in the

microsecond regime while T4L-131/135 shows a sub-microsecond behavior (Table 4-6). More importantly, this sub-microsecond component is the dominant component as can be seen from the corresponding amplitude values. In addition, these findings suggest that the observations for the inter-domain helix, where the fluorophores are on the same helix, are similar to observations for the active site region, giving further support to the contention that the long helix facilitates hinge-bending motion.

IV.3.6 Equilibrium Conformational States and Populations

As noted above, the average number of molecules in the observation volume can be experimentally determined from fitting the correlation curves ($<N>$ in the FCS models, chapter II) or from cumulant analysis (chapter II). This number reflects TAMRA molecules that sample a fluorescent state during diffusion. Using the experimentally determined observation volume, an expected number of particles can be calculated based on UV-Vis absorbance of T4L at 280 nm. We found that concentrations calculated from absorbance measurements agree with those determined experimentally by CA (Fig. 4-15 *A*) and FCS (Fig. 4-15 *B*) for mutants where the fluorophores are too far apart for quenching (T4L-65/135). In contrast, a significant discrepancy was observed for doubly labeled mutants where the fluorophores are expected to come into contact, either due to the dynamics of the protein (e.g., 22/109) or their close proximity (e.g., 151/154). This discrepancy implies that only a fraction of the T4L molecules samples the fluorescent state as they pass through the observation volume.

In order to determine the source of the difference in particle numbers, a separate accounting of the number of T4L and TAMRA molecules was carried out for each mutant. All samples were initially prepared at the same concentration (2 μM) as determined by UV-Vis absorbance and then diluted by ~20 – 200 fold for CA and FCS measurements. Dithiothreitol (DTT) was then added to release TAMRA fluorophores which are conjugated to the protein via disulfide bonds. The number of particles was determined by CA and FCS prior and subsequent to

DTT addition (Fig. 4-15 *A* and *B*). The measured values in the absence of DTT (black bars) reflect the number of T4L molecules while, in the presence of DTT, the number (purple bars) is expected to be twice the number of T4L, assuming all T4L double mutants are stoichiometrically labeled. Labeling efficiency was verified from absorbance measurements and for selected mutants by mass spectrometry.

Fig. 4-15 *A* and *B* show that for T4L-65/135 the particle numbers calculated by CA and FCS and UV-vis absorbance are approximately equal. In contrast, for a majority of the other mutants, the TAMRA populations are significantly larger than twice the corresponding T4L populations. Thus, there exists a fraction of T4L molecules that are undetected, presumably in a non-fluorescent state for the duration of diffusion through the observation volume.

Consistent with this conclusion, UV-Vis absorbance spectra for all T4L double mutants, excluding T4L-61/135 and T4L-65/135, show the presence of a blue-shifted peak. Given the short distances between residue sites in the T4L mutants, the fluorophores are close enough to form dark or H-dimers. The fraction of these can be calculated using the following relation:

$$1 - \frac{< N >}{0.5 < N >_{DTT}} \tag{4.1}$$

In the above relation, the average number of molecules in the observation volume with and without DTT is given by $<N>_{DTT}$ and $<N>$, respectively. Fig. 4-16 shows the fraction of dark dimers deduced from the equation above. Double mutants that do not experience hinge-bending dynamics, but where the fluorophores are close enough for formation of H-dimers, have a substantial fraction of particles that are undetected.

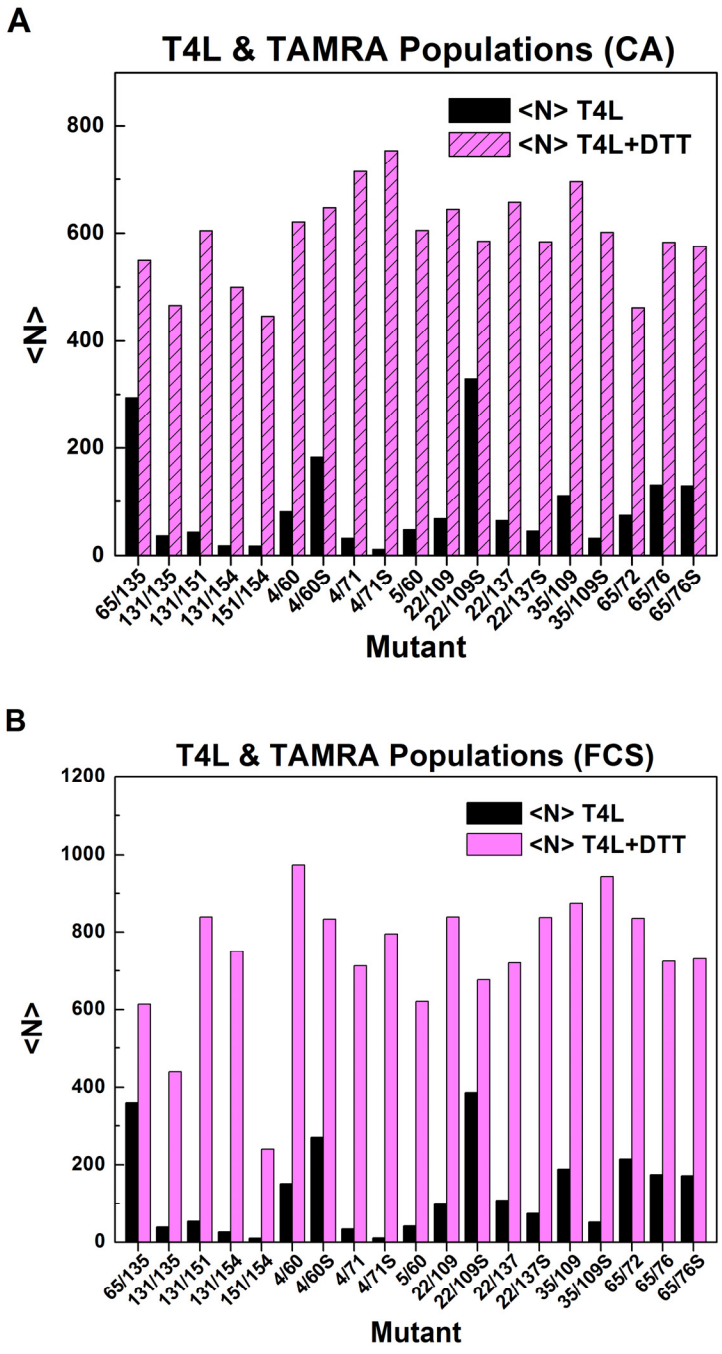


Figure 4-15. The average number of T4L molecules (black) from *(A)* CA and *(B)* FCS for T4L samples of the same concentration (2 μM) and the corresponding TAMRA molecules (magenta) determined after the addition of DTT are shown. For T4L-65/135, the released number of TAMRA is approximately twice the number of T4L-65/135, as expected for a doubly labeled mutant. For all other mutants, the number of released TAMRA is greater than twice the number of T4L molecules, revealing the presence of dark T4L populations.

For these mutants, the distance between the fluorophores makes fluorescent states highly improbable. For instance, the molecular brightness for T4L-151/154 as determined by CA (Fig. 4-17), is well below that of TAMRA, illustrating the extent of quenching. The fact that emission is detected for this mutant is likely due to the fluorophores disengaging and assuming a configuration that reduces quenching. That the corresponding correlation curve did not show a fast decay component reflects the fact that fluorophore dynamics is faster than the resolution of our instrument, likely in a time regime of nanoseconds or faster (38).

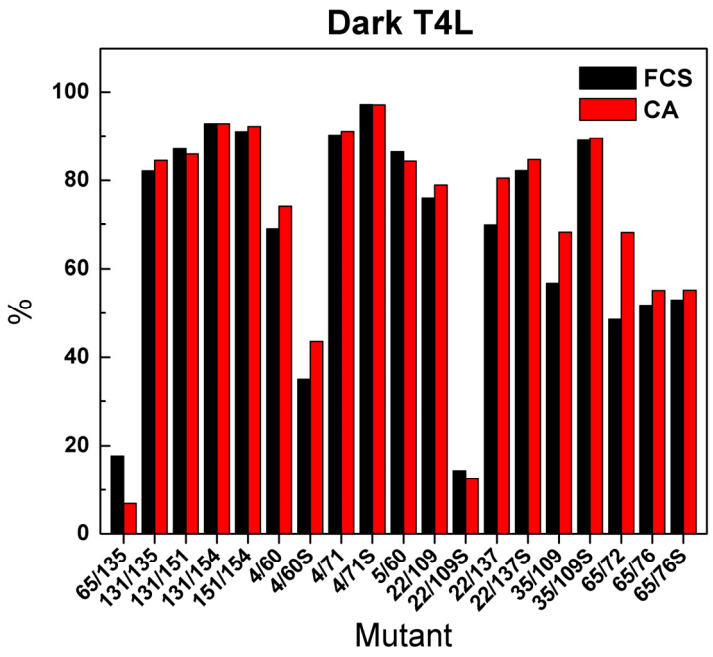


Figure 4-16. Fraction of dark T4L populations corresponding to the population values shown in Fig. 4-15 and calculated by the relation given in Eq. 4.1.

For double mutants where the distance between fluorophores is dependent on the hinge-bending angle such as T4L-4/60 and T4L-22/109, the percentage of dark dimers is modulated by the binding of substrate. Thus, similar to the two-component correlation curves, the dark dimers in these mutants represent conformational states associated with domain motion. For the 4/60 double mutant, substrate binding, which stabilizes the closed conformation, causes a movement

of helix A away from site 60 (Fig. 4-2); hence the decrease in the population of dark dimers. At 22/109, EPR analysis shows that residue 22 moves away from 109 upon closing of the active site. Here the residual percentage of dark dimers is indistinguishable from that of 65/135 where the distance between the fluorophores precludes the formation of such dark populations. A similar change in dark dimer population is observed for 35/109 pointing to a decrease in distance between 35 and 109 with substrate binding.

Taken together, these results demonstrate that dark dimers reflect T4L conformations related to each other by hinge-bending. A substantial fraction of these dimers is not detectable in the FCS analysis because they do not transition to a conformation in which they are fluorescent during diffusion through the observation volume. The discrepancy between the non-fluorescent population calculated after DTT-induced TAMRA release and those from FCS (Table 4-3) arises because the latter is normalized to the detectable number of particles which undercount the total number of particles in the observation volume. Thus, T4L samples a series of equilibrium conformations wherein the TAMRA pairs are close enough for substantial fluorescence quenching. While the number of these states and their structures cannot be deduced from this analysis, their existence revealed by particle accounting excludes the model wherein T4L is in a two-state equilibrium.

IV.4 Discussion

The work presented here gives a fluorescence fluctuation spectroscopy study of T4 lysozyme dynamics in solution. FCS captures conformational dynamics in the active site as well as at the interface between helices A and C – regions of the protein that are displaced in the transformation between the crystallographic open and closed conformations (68). Direct evidence of hinge-bending motion is manifested in the correlation curves as a second component distinct from that of diffusion. This component arises from fluctuations in the fluorescence as the two TAMRA fluorophores are stochastically brought into close proximity, leading to quenching

and the formation of dark dimers. At pairs such as 4/60 and 22/109, both fluorescence and spin labels (69) converge to report conformational heterogeneity as the protein samples conformations with different distance and orientation between the two domains. Furthermore, substrate binding reduced the amplitude of the second component in the correlation curve, in agreement with a shift in equilibrium towards the closed conformations. We find evidence of dynamics along the inter-domain helix which has not been reported previously. Helix fluctuations persist upon substrate binding and transition to the closed state. The exclusion of fluorophore photophysics and secondary structure/fluorophore dynamics as sources of the observed second components supports our contention that the fluctuations represent dynamic modes of the protein. Although T4L hinge-bending motion has long been postulated (68), FCS enables a determination of its timescale in the range of approximately 10 to 20 μs. Relaxation times in the range of microseconds to hundreds of microseconds have been reported for fluctuations arising from domain motions as well as transitions to partially unfolded states (38, 78, 80). An FCS study of calmodulin reported a ~100 μs relaxation time for the dynamics of the two domains in the protein while intradomain dynamics has a 30 to 40 μs relaxation time (79). Chattopdhayay et al. (78) have performed a pH-induced unfolding study of the intestinal fatty acid binding protein (IFABP) using FCS, revealing a fluctuation relaxation time of ~35 μs under native conditions (78). FCS analysis of apo-Myoglobin showed the presence of two relaxation times, 8 and 100 μs, reflecting interconversion between the folded and molten globule conformations under native conditions (38).

The model of a two-state equilibrium between open and closed conformations has its origins in the distribution of hinge-bending angles as observed in the crystal structures of 42 T4L mutants (68). Although the hinge-bending angle values were quite dispersed, on a coarse level they seem to segregate into two groups: those in the range of -10 and 10 degrees and those in the range of 25 and 40 degrees. However, the sparse population in the intermediate range, which led to the reductionist two-state model, could be a consequence of crystal packing similar to the

conformational selectivity that almost exclusively favors the closed structure. EPR analysis of T4L structure in solution did not have the resolution to distinguish intermediate hinge-bending angles while in the NMR study interpretation in terms of an average structure was given.

In the FCS analysis, a two-state equilibrium was assumed so that a relaxation rate could be extracted. Here the two-state model describes those T4L molecules which transition from quenched or dark to fluorescent states or vice versa during transit time in the observation volume (section II.2.2.2). A single exponential term accounted for the fast decay component of the correlation curves. That the single exponential diffusion/quenching model is the appropriate model is not only supported by comparing reduced chi-square values but can be gleaned visually from superposition of the data and the respective fits. For example, the average of the reduced chi-square values for T4L-22/137 decreased from 1.978 to 0.758 in going from the diffusion-only to the single-exponential diffusion/quenching model. Furthermore, we can test whether the addition of a second exponential term is necessary in a similar way. The reduced chi-square values for the single and double-exponential diffusion/quenching model (Eq. 2.25) fits for T4L-22/137 were found to be 0.758 and 0.726, respectively. Thus, addition of the second exponential term does not improve the quality of the fit, making the single-exponential model the appropriate choice.

Number of particles determined from FCS and CA combined with the molecular brightness values determined from CA revealed a large fraction of T4L molecules where the TAMRA fluorophores remain dark during diffusion through the observation volume. These must reflect T4L intermediates wherein the fluorophore pairs are in close proximity enabling H-dimer formation. One possible arrangement that can account for the presence of such large numbers of dark populations within the framework of a dynamic equilibrium is given in Fig. 4-17. In this case, instead of a transition between an open and a closed conformation, transitions occur between multiple conformations with intermediate hinge-bending angles. The requirement imposed by the FFS results is that the majority of these states are non-fluorescent (D). The

predominance of dark states is not unexpected given the relative size of TAMRA with respect to the distances between the sites. It is also likely that heterogeneity exists in the transition from the non-fluorescent states (D) to the fluorescent state (F) and may be indicated by the necessity of two exponential terms to account for the fast decay component in the 20 MHz correlation data.

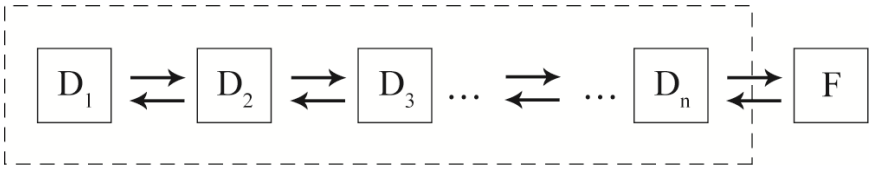


Figure 4-17. Model of T4L conformational equilibrium consistent with FCS and CA results. Multiple non-fluorescent conformations (D) are required to account for the experimental data. These conformations have intermediate values of hinge-bending angles between the open and closed conformations.

The conclusion that T4L equilibrium involves multiple conformations with distinct hinge-bending angles is also supported by a recent study probing T4L non-equilibrium (in the presence of substrate) conformational dynamics through the use of an electronic circuit – a carbon nanotube field-effect transistor (90). Analysis of hinge-bending motion showed that closure motion occurs in a single step while the opening motion occurs in a minimum of two steps. Further experiments are required to explore the multi-state model of T4 lysozyme dynamics.

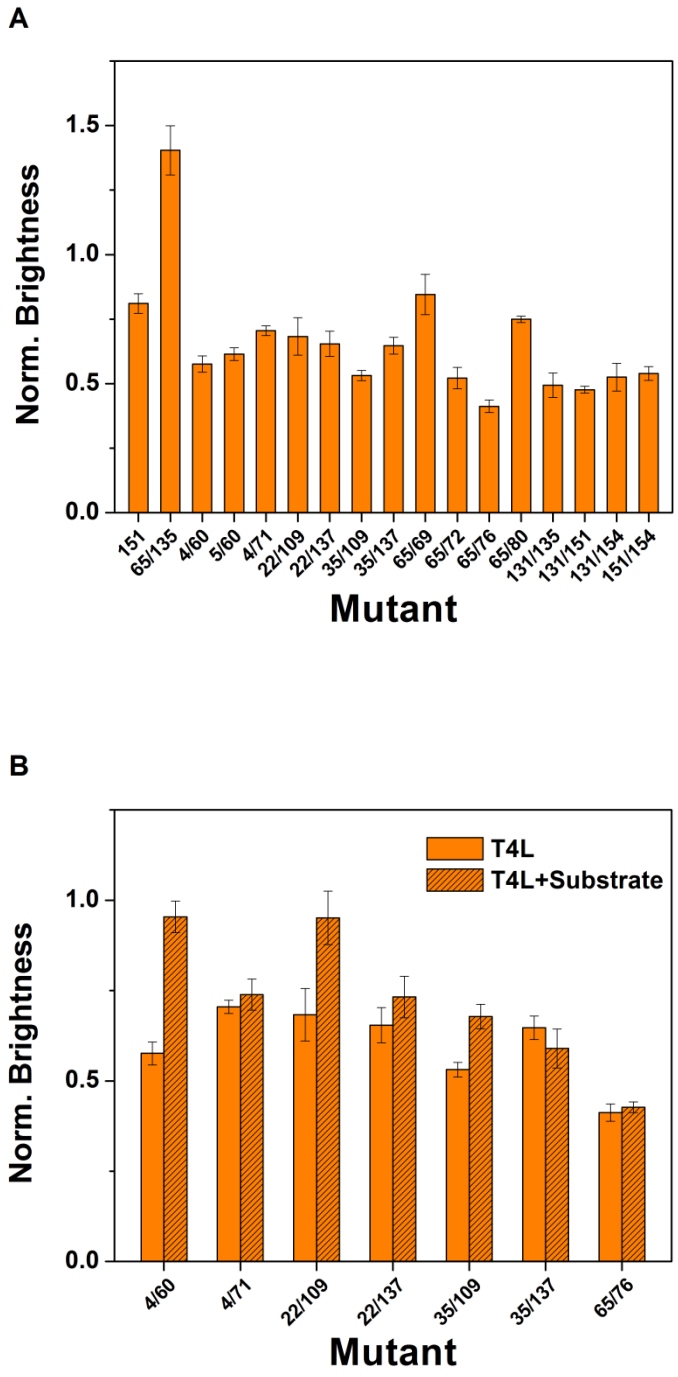


Figure 4-18. *(A)* Normalized molecular brightness values for singly and doubly labeled T4L mutants. The values have been normalized to the brightness of a single TAMRA fluorophore. *(B)* Relative changes in molecular brightness (normalized) upon substrate binding.

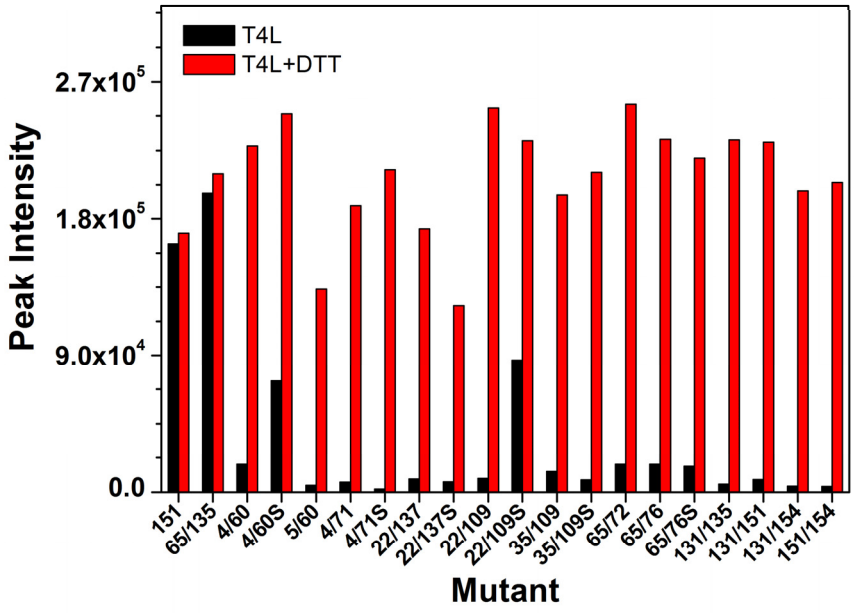


Figure 4-19. Peak emission intensity (574 nm – 581 nm) determined from steady-state fluorometry for T4L with and without DTT.

CHAPTER V

SUMMARY AND FUTURE DIRECTIONS

V.1 T4 Lysozyme

In recent years, a number of studies have produced results that have expanded our knowledge of protein dynamics while bringing the topic to the forefront (9, 11, 15, 16, 20). The importance of protein dynamics studies has been clearly highlighted by the journal Science, which dedicated a special issue to the topic in April of 2009. Thus it is becoming increasingly clear that the knowledge of intrinsic protein motions is one significant piece of the puzzle we face in obtaining the full picture of a protein system.

The observation of dynamics in T4 lysozyme has its roots in early crystallographic studies probing the fundamental nature of protein structure, stability and folding. From the onset, the WT crystal structure of T4L appeared incapable of accommodating the peptidoglycan substrate. Thus, it was suggested that the active site region needed to open up to allow access to the substrate (91). Support for this supposition was to come from the multitude of crystal structures produced by the various crystallographic studies. The abundance of available structures was such that a distribution of hinge-bending angles could be observed leading to the proposition that the domain motion involves a transition between average open and closed conformations (68).

Although this was a plausible proposal, clear evidence was lacking. In the work presented here, the equilibrium dynamics of T4L has been investigated by taking advantage of the sensitivity of single molecule fluorescence spectroscopy. The outcome of this study is two-fold: first, a relaxation time for the domain motion in T4L has been obtained, and second, the results shed light into the nature of the conformational equilibrium of T4L in solution.

At the heart of this study was the use of the self-quenching of TAMRA pairs in close proximity. In addition, the strategy was to investigate T4 lysozyme at multiple locations which were selected to allow for the optimal detection of the domain motion, as illustrated by the EPR study (69). The FCS results from these carefully selected locations on the protein showed correlation curves with decay profiles that are ultimately attributed to the hinge-bending motion of T4L. In general, this type of motion serves to capture substrate and then release final products (18). Particularly for T4L, the domain motion opens up the active site region to allow access for the substrate, as initially suggested.

Prior to the FFS study reported here, the domain motion in T4L has been investigated by NMR and MD (71, 74). The dipolar-coupling study by NMR was able to place the domain motion at or below 160 μs while the results from the MD study indicated a much faster motion, in the order of nanoseconds. Therefore, precise temporal information has been lacking. FCS is well suited for determining relaxation time values as demonstrated by various studies (38, 39, 78-80), and the results give a value in the range of 10 – 20 μs for the domain motion in T4L.

Furthermore, investigation of the inter-domain helix showed the presence of dynamics along the helix. The type of hinge-bending motion observed in T4L has been classified as a "double-hinged alpha-helix" construct – the ends of the helix act as the connecting regions (18, 92). This type of structural arrangement has been observed in other proteins as well and it was suggested that the helix can act as a spring, in effect facilitating the hinge-bending motion (92).

The two-state description of the hinge-bending motion in T4L was proposed on the basis of the distribution of hinge-bending angles observed (68). This conjecture raises a crucial question that Goto et al. (71) explored in their study – "What is the average conformation of bacteriophage T4 lysozyme in solution?" The conclusion from their dipolar coupling study was that on average, the active site region for T4L in solution is more open than what was observed in the WT structure. Furthermore, one important finding of the EPR study was that the open

conformation is favored. Thus, conformational heterogeneity is indicated by both studies and this is in agreement with a model consisting of two-states.

The results presented here, on the contrary, show that the two-state model is insufficient to describe the nature of the conformational equilibrium. The presence of dark T4L populations requires that various intermediate hinge-bending angles be sampled. The non-fluorescent nature of these states further implies that the intermediate conformations are such that the TAMRA pairs are in close proximity to enable fluorescence quenching. Although the results point to intermediate conformations, the extent of opening of the active site region is unclear. That these intermediates are non-fluorescent could be mistaken to imply that the closed conformation is favored. However, we have to note that the linker of the fluorophore may also be enabling the fluorophore pairs to come into contact.

A second evidence showing the inadequacy of the two-state model comes for the FCS results. The FCS analysis, specifically the description of a two-state system encapsulated by the diffusion/quenching model given in Eq. 2.24, should suffice for such a system. However, the mismatch between the non-fluorescent populations determined from this analysis (Fig. 4-6, Table 4-3), and those determined directly through the use of DTT to determine T4L populations (Fig. 4-16) shows that the two-state model is insufficient.

What permitted the revelation of intermediate conformations was the H-dimer formation capability of the fluorophore TAMRA. An immediate follow-up to this would be the examination of the make-up of the intermediate states. Since H-dimers are non-fluorescent, the use of xanthene fluorophore pairs makes it difficult to probe into the set of intermediate conformations. Instead, the use of fluorescence resonance energy transfer (FRET) may be able to garner information not obtained by FFS. Since FRET utilizes non-radiative long range dipole-dipole interaction of a fluorophore pair, a direct fluorophore contact is not required in order to observe conformational changes in the protein. Thus, with the appropriate fluorophore pair, FRET may be able to reveal characteristics of the intermediate conformations of T4L.

In addition, the behavior of the fluorophore pair can be explored further. In the FCS results presented here, in particular those showing microsecond timescale fluctuations, the control experiments have shown that the fluctuation cannot arise from fluorophore dynamics. This is in agreement with the accepted timescale for the dynamics of a conjugated fluorophore (38). Because of the fast timescale, the behavior of fluorophores conjugated to a protein can be examined by molecular dynamics simulations. Such studies could produce important insights that can aid in the selection of a fluorophore for FFS experiments.

V.2 Application of FFS to Complex Protein Systems

Two properties of T4 lysozyme that make it a good system for FFS experiments is its relatively small size (19 kDa) and good stability. Unfortunately, other systems of interest may not possess these characteristics. Two such systems are the Ca^{2+}/calmodulin-dependent protein kinase II (CAMKII) and the membrane transporter MsbA. The application of FFS on these types of systems presents challenges not encountered with T4L.

V.2.1 Ca^{2+}/calmodulin-dependent protein kinase II (CAMKII)

Ca^{2+}/calmodulin-dependent protein kinase II (CAMKII) is a Ser/Thr kinase that is involved in various signaling pathways. Its role in the regulation of proper heart function as well as synaptic transmission in the brain have been documented (93). CAMKII is a dodecameric holoenzyme with each monomer consisting of a catalytic, regulatory and association domains (Fig. 5-1 *A*) (94). A monomer construct (Fig. 5-1 *B*) with only the catalytic (residues 1 – 237) and the regulatory domains (residues 274 – 314) was utilized by a recent EPR study investigating the regulatory domain (93). This study has found that the regulatory domain (shown in orange in Fig. 5-2 *B*) exhibits a complex structural behavior. In particular, it was found that the Ca^{2+}/CaM-binding segment of this domain, the upper portion of the helix with residues 297 – 314, exists in a dynamic equilibrium in the apo state, between docked and undocked conformations.

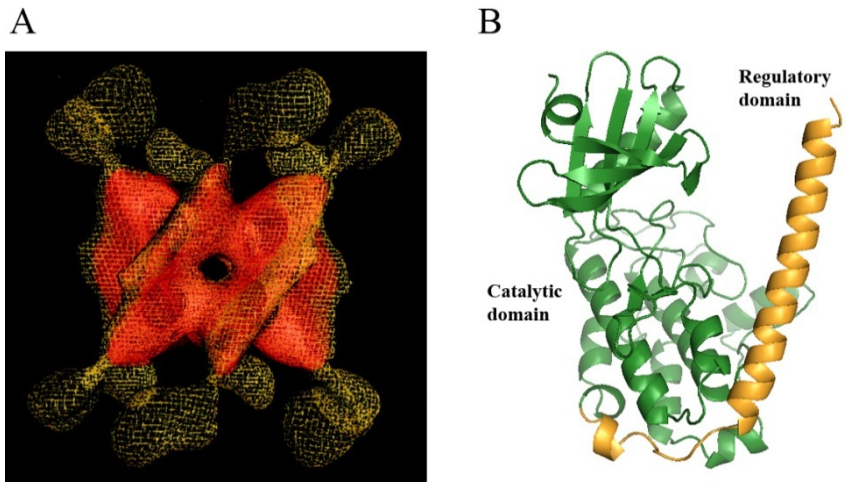


Figure 5-1. *(A)* The three-dimensional structure of CAMKII showing the twelve subunits forming two hexameric rings. *(B)* The monomer construct used in the EPR and FFS experiments.

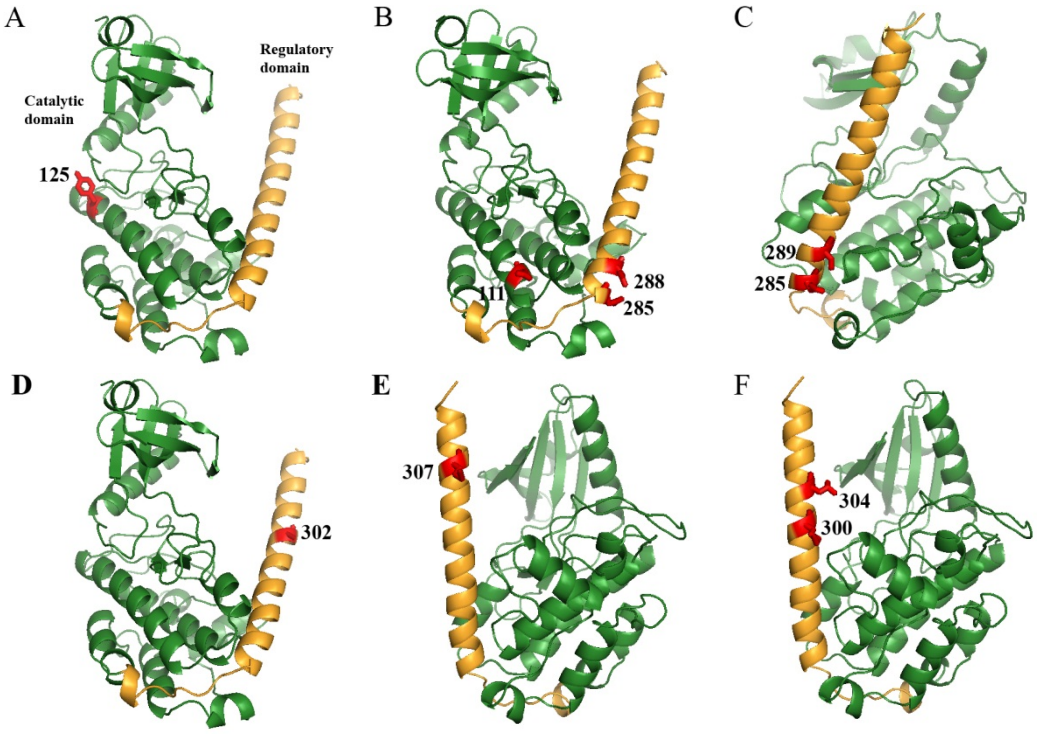


Figure 5-2. *(A)* A control mutant labeled with a single TAMRA at site 125. Also shown are residues covering the lower *(B and C)* and upper *(D - F)* segments of the regulatory domain.

Furthermore, this dynamic equilibrium is modulated by ATP binding. On the other hand, the lower segment of the regulatory domain, residues 274 – 291, was found to be docked to the catalytic domain. The binding of Ca^{2+}/CaM rearranges the entire domain with the lower segment becoming undocked and less structured while the Ca^{2+}/CaM binding segment becomes structured.

Both segments of the regulatory domain were investigated by FFS. The specific residue sites selected for labeling with TAMRA are shown in Fig. 5-2. CAMKII with TAMRA at site 125 was used as a control (Fig. 5-2 *A*). In the lower segment of the regulatory domain, four doubly labeled mutants were prepared (Figs. 5-2 *B* and *C*): 111/285, 111/288, 285/289 and 285/289E. In the first two, site 111 is located on the catalytic domain while 285 and 289 are on the regulatory domain. Thus site 111 essentially serves as a static site with respect to sites 285 and 289 which are expected to show dynamics. 285/289E contains the mutation T286E which is a phosphorylation mimic. The phosphorylation of T286 has been observed to destabilize the lower segment (93). For the upper segment, single mutants 302 and 307 and the double mutant 300/304 were prepared (Figs. 5-2 *D* – *F*). These are all located on the regulatory domain and allow us to examine the dynamics of that specific region.

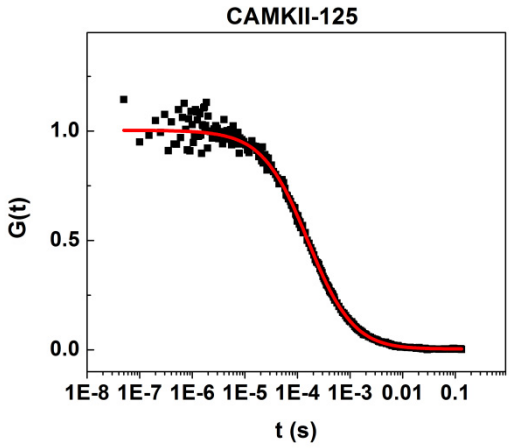


Figure 5-3. A representative correlation curve for the control mutant shows a single decay profile. The fit to the diffusion-only model is shown in red.

Fig. 5-3 shows the correlation curve for the control mutant, CAMKII-125. The curve displays a single decay profile and is fit to the diffusion-only model, giving a diffusion coefficient of 65.1 $\mu m^2/s$, in good agreement to the value expected based on molecular weight scaling, 67 $\mu m^2/s$.

Correlation curves for the upper segment of the regulatory domain, CAMKII-302, CAMKII-307 and CAMKII-300/304 are shown in Fig. 5-4. In all cases a fast decay component is observed indicating fluctuations in this region. For the singly labeled mutants, the fluorophore is possibly being quenched as a result of collisions with nearby residues. All three curves have been fit to the double-exponential diffusion/quenching model (chapter II, Eq. 2.25) shown in magenta

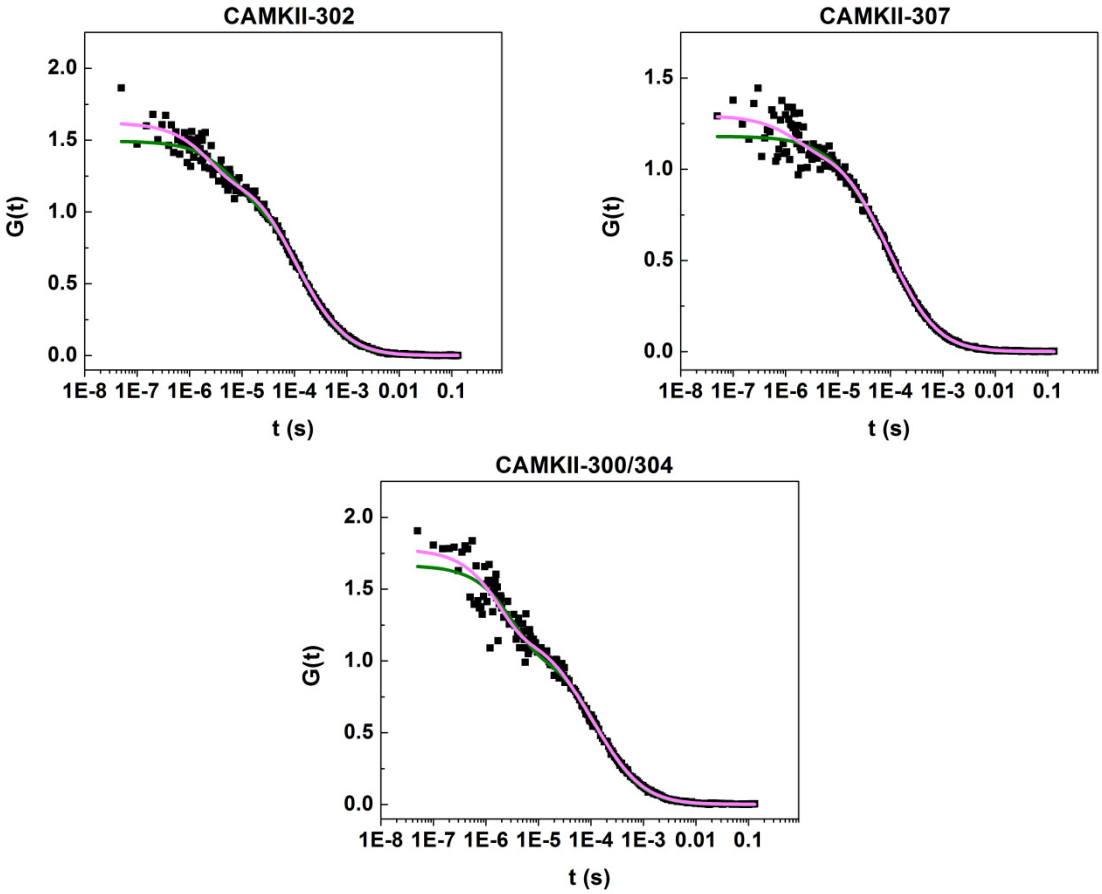


Figure 5-4. Correlation curves for the upper segment of the regulatory domain show fast decay components. The curves have been fit to the single and double exponential diffusion/quenching models shown in green and magenta, respectively. The presence of fluctuation in this region in the apo state is in agreement with the EPR findings.

and the values are given in Table 5-1. The parameters characterizing the fluctuation are reasonably similar aside from the lower amplitude values of CAMKII-307. The extent of quenching as seen from brightness analysis (Fig. 5-9) is comparable for the three mutants.

TABLE 5-1 Parameters from the double-exponential model fit of correlation curves in the upper segment of the regulatory domain

Mutant	τ_{R1} (μs)	A_1	τ_{R2} (μs)	A_2
302	37.1 ± 29.3	0.21 ± 0.08	1.7 ± 1.1	0.41 ± 0.11
307	26.7 ± 9.4	0.12 ± 0.02	1.0 ± 0.2	0.18 ± 0.05
300/304	26.0 ± 8.4	0.25 ± 0.07	1.8 ± 0.9	0.48 ± 0.09

As mentioned above, ATP binding was observed to modulate the dynamics in this region of the domain. Correlation curves for CAMKII-302, CAMKII-307 and CAMKII-300/304 in the presence of ATP are shown in Fig. 5-5. For 300/304, the curve has the same profile as in the Apo case and is fit to the double-exponential quenching model. The relaxation time values determined were 24.8 μs and 1.5 μs with corresponding amplitude values of 0.25 and 0.49. Thus, FCS does not show a change in the fluctuation in this region upon ATP binding. On the other hand, for 302 and 307, the addition of ATP caused the fluorophores to become unbound from the protein. For CAMKII-307 the correlation curve in the presence of ATP shows a fast decay component and the curve is shown fit to the double exponential model (shown in magenta) as well as a two-component diffusion model (64) (shown in blue). From the analysis, it can be seen that the free TAMRA population in this case is minimal, approximately 20 %. The two-component diffusion fit for CAMKII-302, on the other hand, indicated a free TAMRA population of approximately 95%.

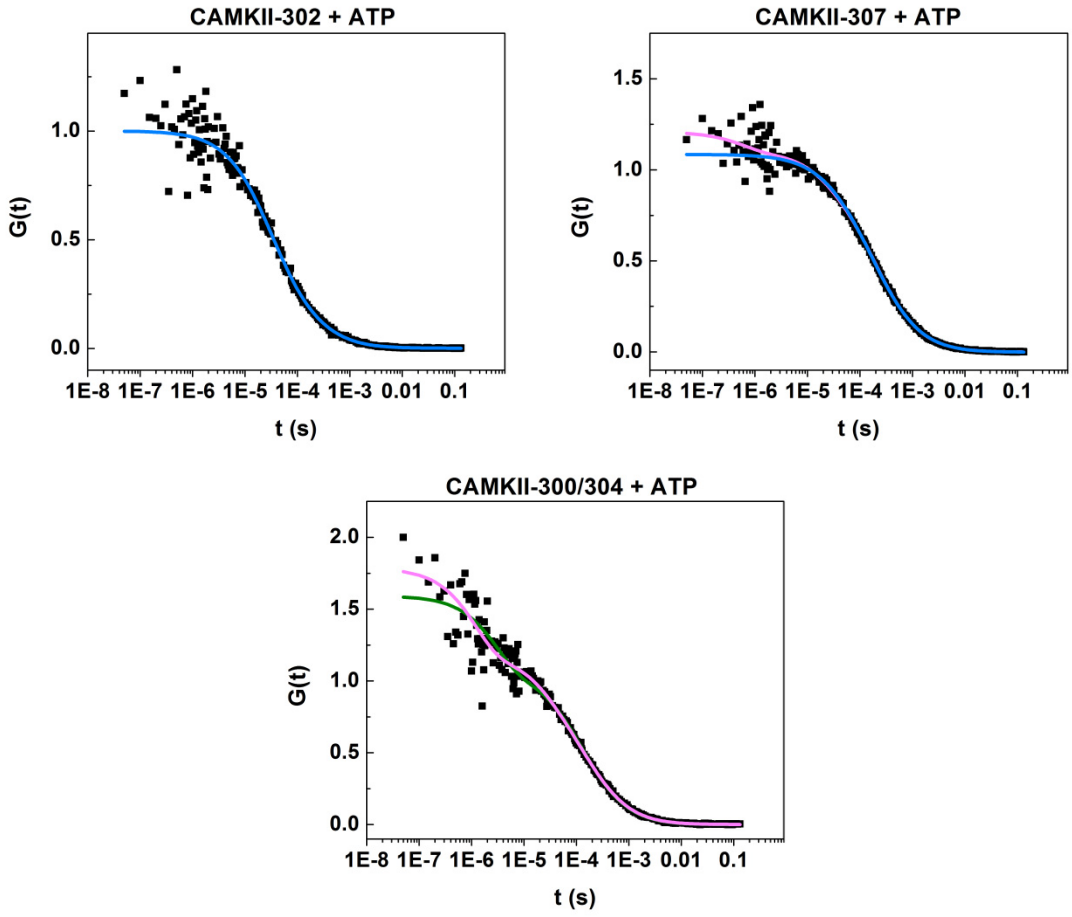


Figure 5-5. The addition of ATP causes the fluorophores to be removed from the protein for CAMKII-302 and CAMKII-307 but not for CAMKII-300/304. The correlation curves for the first two are shown fit to the two-component diffusion model (blue). In the presence of ATP, the fluctuation seen in CAMKII-300/304 is not diminished.

Correlation curves for residues in the lower segment of the domain are shown in Fig. 5-5. For residue pairs 111/285 and 111/288, the curves show fast decay components and are fit to the double-exponential diffusion/quenching model. The respective parameters (Table 5-2) show relaxation time values that are overall comparable to those found in the upper segment of the regulatory domain aside from the submicrosecond value for the second relaxation time of 111/285. In the apo state, EPR results indicated that this region is docked to the catalytic domain, in contrast to the FCS results. In addition, phosphorylation of T286 causes the lower segment of the regulatory domain to become more flexible. However, the correlation curves for residues

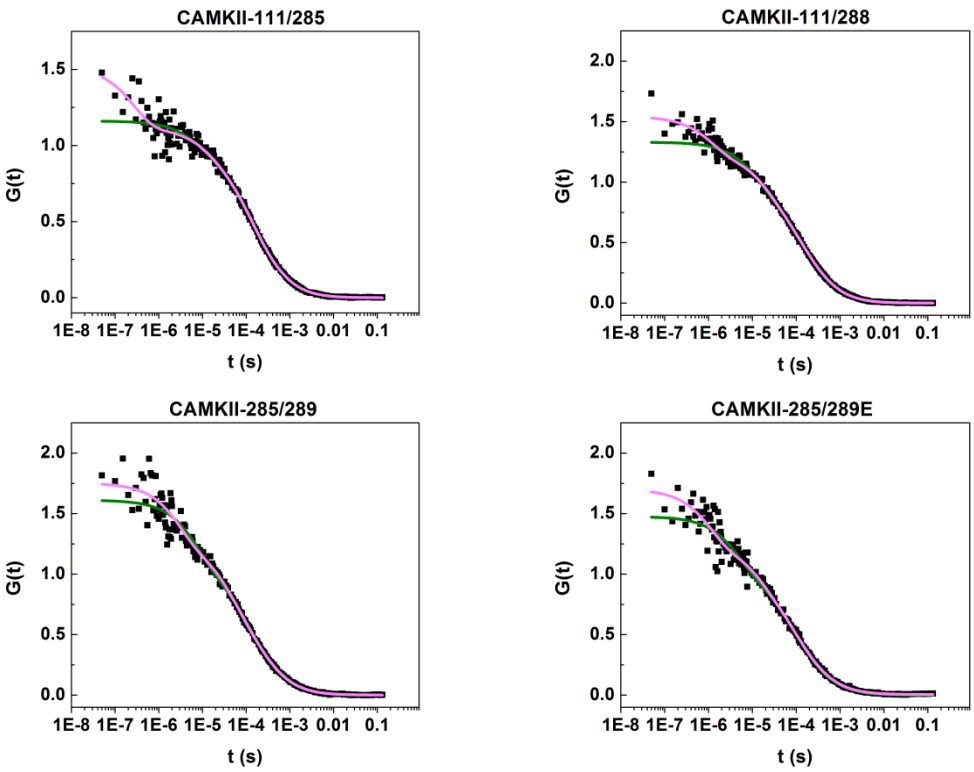


Figure 5-6. Unlike the EPR results, correlation curves representing the lower segment of the regulatory domain show fast decay components in the apo state, similar to the upper segment. All of the curves are shown fit to the single and double exponential diffusion/quenching models (green and magenta).

TABLE 5-2 Parameters from the double-exponential model fit of correlation curves in the lower segment of the regulatory domain

Mutant	τ_{R1} (μs)	A_1	τ_{R2} (μs)	A_2
111/285	25.1 ± 14.7	0.13 ± 0.04	0.8 ± 0.3	0.30 ± 0.07
111/288	29.4 ± 10.7	0.20 ± 0.06	2.0 ± 0.7	0.34 ± 0.03
285/289	27.1 ± 10.3	0.24 ± 0.04	2.9 ± 0.9	0.42 ± 0.08
285/289E	27.5 ± 11.9	0.27 ± 0.08	1.9 ± 0.6	0.45 ± 0.08

flanking this special site with and without the phosphomimic, 285/289 and 285/289E respectively, both show correlation curves with similar profiles. Under either condition, the curves indicate fluctuations at these residue sites with both curves fit to the double-exponential diffusion/quenching model. Furthermore, the parameters determined from the fits show that the fluctuations in both cases are very similar. Brightness analysis (Fig. 5-9) for this region indicates higher quenching for 285/289 and 285/289E relative to 111/285 and 111/288. Thus, the TAMRA pair is in closer proximity along the helix. The slightly lower value for 285/289E compared to 285/289 could be hinting at increased flexibility bringing the fluorophore pair closer, although the FCS analysis did not show a measurable difference.

In addition, this region of the domain was investigated using a peptide construct which binds to the catalytic domain. This in turn causes the regulatory domain to be undocked. As a result, the regulatory domain is expected to have increased flexibility. The correlation curves for 111/285 and 285/289E in the presence of peptide are shown in Fig. 5-7. In both cases, the addition of peptide resulted in the fluorophores being removed from the protein. The correlation curves were fit to the two-component diffusion model shown in blue. The analysis indicates the

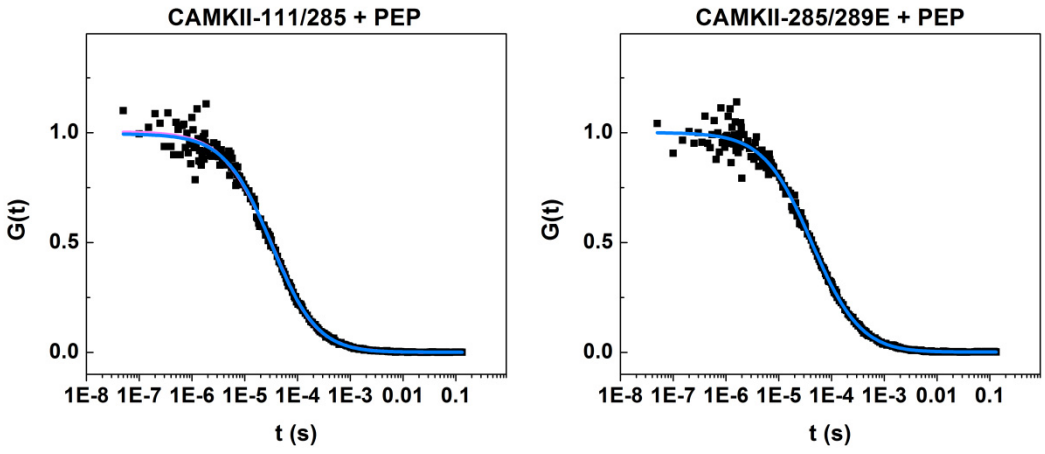


Figure 5-7. The addition of a peptide construct that binds to the catalytic domain causes the fluorophores to be removed from CAMKII. The two correlation curves are fit to the two-component diffusion model shown in blue. The presence of the peptide likely perturbs the protein enough to release the fluorophores.

presence of at least 50 % free TAMRA for CAMKII-111/285 and approximately 75 % for CAMKII-285/289E.

Finally, the lower segment was examined in the presence of CaM which, from the EPR results, causes this part of the regulatory domain to become undocked from the catalytic domain, increasing the flexibility of the region. Correlation curves for this segment were investigated at residue pairs 111/285, 285/289 and 285/289E (Fig. 5-8). In the case 285/289 and 285/289E the correlation curves show fast decay components and are fit to the diffusion/quenching model. For these mutants, the relaxation time values determined were approximately 30 - 50 µs and 4 µs with corresponding amplitudes of about 0.25 and 0.4. These values are comparable to those obtained

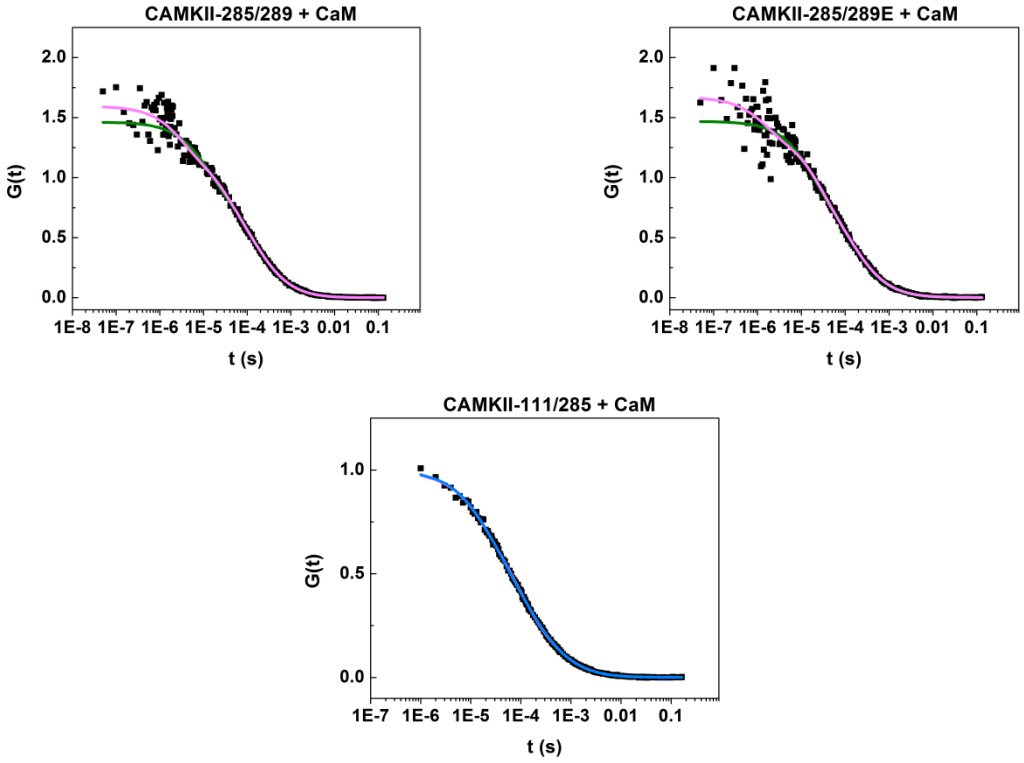


Figure 5-8. The correlation curves for 285/289 and the phosphomimic counterpart, 285/289E, both show a fast decay component that is similar to the apo state. For 111/285, CaM binding caused severe aggregation and the correlation curve is shown fit to the two-component diffusion model.

for the apo state. The similarity is evident in the brightness values as well (Fig. 5-9 *B*). For 111/285, the binding of CaM resulted in severe aggregation. The resulting free TAMRA population was 55 % and higher. Thus, binding of CaM perturbs the system leading to unfolding and aggregation. The presence of the TAMRA pair itself may be a contributing factor as well.

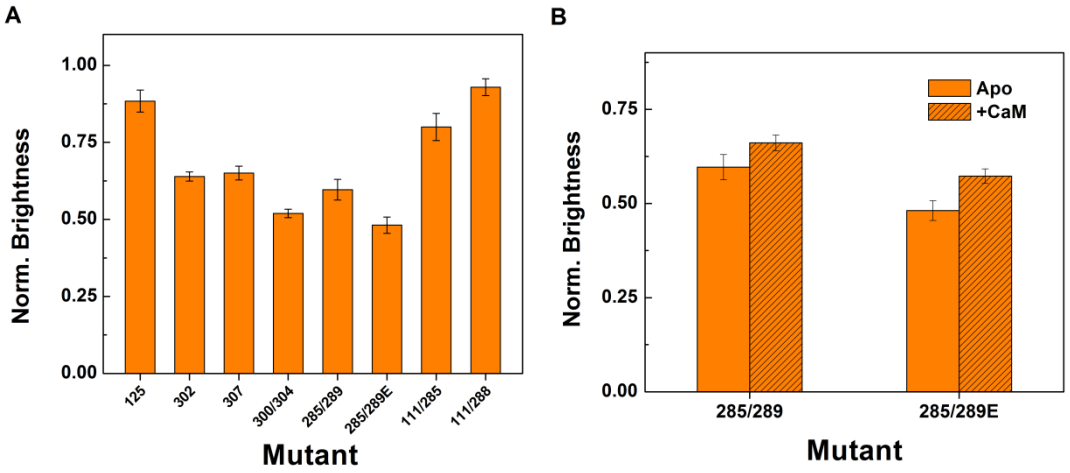


Figure 5-9. *(A)* Normalized brightness values for both segments of the regulatory domain showing quenching of fluorescence. The quenching for the single mutants 302 and 307 is likely due to interactions with nearby residues. A slight difference in the proximity of the TAMRA pairs, as reflected by the difference in brightness values, can be seen between pairs 285/289 and 285/289E as well as 111/285 and 111/288. *(B)* Brightness values show a small increase upon CaM binding.

V.2.2 MsbA

MsbA is a membrane protein belonging to the class of ATP-binding cassette (ABC) transporters. This class of proteins is involved in translocating various substrates across the membrane, powered by energy harnessed from ATP. ABC transporters have been implicated in multi-drug resistance as well as other diseases such as cystic fibrosis (95). MsbA, which is found in *E. coli*, has been extensively studied and has become a model system for structural dynamics studies (96). The crystal structures of this protein in the open and closed conformations are shown in Fig. 5-10. The structure of MsbA is a dimer and consists of two sets of transmembrane (TMD) and nucleotide-binding (NBD) domains. The transmembrane domains extend across the

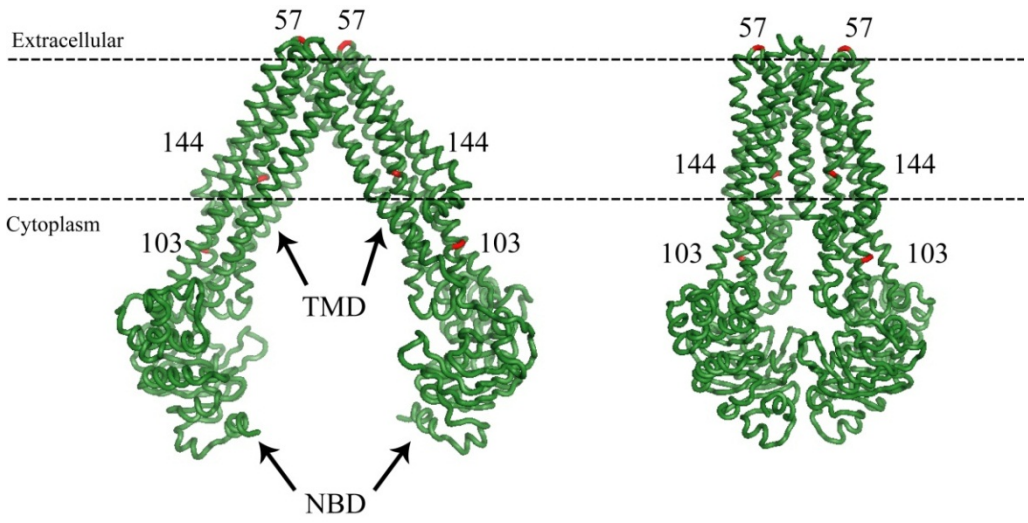


Figure 5-10. The structure of the membrane transporter MsbA is shown in the apo open (left, PDB ID 3B5W) and closed (right, PDB ID 3B5X) conformations. The structure consists of a pair of transmembrane (TMD) and nucleotide-binding (NBD) domains. The residue locations used for labeling with TAMRA are shown in red.

membrane and each domain is itself connected to a nucleotide-binding domain where binding and hydrolysis of ATP occurs. In the open conformation, the two NBDs are separated by 50 Å (96).

An EPR study (97) investigating MsbA in the apo and nucleotide-bound states revealed that in the absence of nucleotide, the cytoplasmic side is in the open conformation. With ATP binding, significant conformational changes occur that result in closing the cytoplasmic side while opening up the extracellular side, relative to the apo state. To further explore this intricate conformational dynamics, FFS experiments were performed looking at MsbA in the apo state as well as in the presence of nucleotide. Both the cytoplasmic and extracellular sides were investigated at the residue sites indicated in Fig. 5-11. At site 57 in the extracellular side, in the closed conformation the distance between these pair of sites is larger than it is in the open conformation. On the other hand, the distance between residue pairs 144 in the middle of the TMD is larger in the open conformation. Thus, the transition from the closed to the open conformation increases the distance between the 144 pair while decreasing the distance between the 57 pair. Residue site 103 was used as a control since the TAMRA pairs are not expected to come into contact in either conformation.

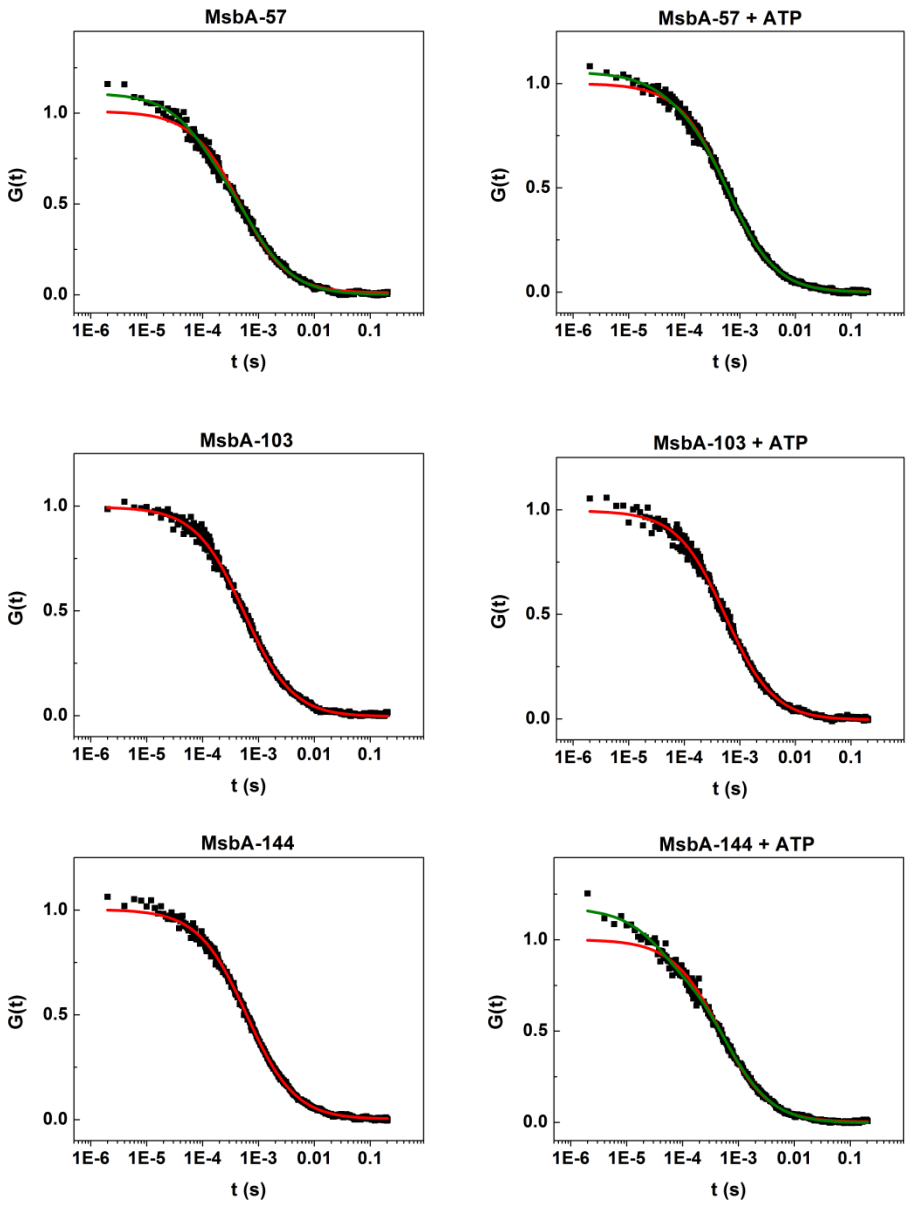


Figure 5-11. Representative correlation curves in the presence and absence of nucleotide, for the residue sites indicated in Fig. 5-10, covering the extracellular and transmembrane regions are shown. For the extracellular side, MsbA-57, the curves show a fast decay component that is slightly reduced with the addition of nucleotide. For MsbA-144, the presence of fluctuation is indicated with the addition of nucleotide. This is likely due to the TAMRA pair being in closer proximity in the closed conformation. MsbA-103 serves as a control. Fits to the diffusion-only and diffusion/quenching models are shown in red and green respectively.

The correlation curves for the apo and nucleotide-bound states at these three pairs of sites are shown in Fig. 5-11. The curves for MsbA-103 in the presence and absence of nucleotide show a single decay profile. Both curves were fit to the diffusion-only model (shown in red) giving a diffusion coefficient of approximately 35 $\mu m^2/s$. However, brightness analysis (Fig. 5-12) shows quenching which is likely due to the interaction of the fluorophore with nearby residues. The correlation curve for residue site 57 indicates the presence of fluctuation in the apo state. The diffusion/quenching model fit (shown in green) gives a relaxation time in the region of 60 μs with a non-fluorescent population of about 20%. The correlation curve for MsbA-144 does not show a fast component as seen in MsbA-57. In the presence of nucleotide, the curve for MsbA-57 still shows a fast component but with reduced amplitude corresponding to non-fluorescent population of about 10 %. On the contrary, the curve for MsbA-144 shows a fast component in the presence of nucleotide. It is possible that in this case the fast component is

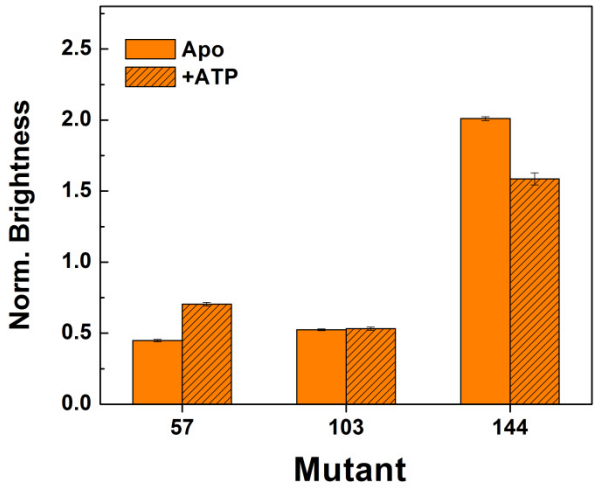


Figure 5-12. Normalized brightness values for the three MsbA mutants in the presence and absence of ATP are shown. The increase in brightness for MsbA-57, extracellular side, and the decrease for MsbA-144, in the transmembrane region, indicates an increased separation for the TAMRA pair at 57 while a decreased separation for the pair at 144. This is in agreement with the opening of the extracellular side and closing of the cytoplasmic side upon nucleotide binding.

reporting secondary structure fluctuations. A second possibility is that the fast component in the correlation curve is due to free TAMRA fluorophores.

Results from brightness analysis (Fig. 5-12) for MsbA-57 and MsbA-144 are in agreement with the conformational changes that occur with nucleotide binding. With the addition of nucleotide, the brightness value increases for MsbA-57 indicating an increase in the distance between the TAMRA pair. On the contrary, the brightness value for MsbA-144 decreases with the addition of nucleotide suggesting that the TAMRA pair are now closer, in agreement with the closure of the cytoplasmic side upon nucleotide binding.

For both MsbA and CAMKII, further experiments are required to delve deeper into the results presented above. These two systems provide valuable insights into the application of FFS as a whole to complex proteins. First, the lack of agreement between the FCS and EPR results (CAMKII) shows that what is detectable through the use of EPR may not be necessarily be detectable by FFS. The difference in the size of the fluorophore and the spin-label leading to different levels of sensitivity is one likely reason. Second, compared to T4 lysozyme, the monomer construct of CAMKII is fragile, as evidenced by the aggregation and the release of the fluorophores upon addition of substrates and nucleotides. The presence of TAMRA itself can be destabilizing to this system. For such a complex and fragile system as CAMKII, a thorough biochemical characterization is therefore necessary and crucial.

MsbA maybe an example of a system that challenges the capability of FCS to detect conformational dynamics. It is possible in this case that the timescale of fluctuation is slow enough to be comparable to the diffusion time of the protein. Under this condition, FCS will be incapable of separating the two processes. Therefore, alternative experimental approaches can be explored that allow the two processes to be temporally distinguishable. One possible approach is to attach the system to a spherical bead with large molecular weight. However, owing to the weak relationship between diffusion coefficient and molecular weight, the use of a bead may not

entirely solve the problem. A better approach would be to immobilize the system altogether through the use of biotin-avidin bound to a lipid bilayer on a glass surface (98).

V.3 Determining Molecular Species Populations by FFS

The ultimate promise and potential of single molecule fluorescence spectroscopy techniques is the ability to resolve molecular detail. In biological systems, heterogeneity exists at all levels of complexity and the use of SMFS to resolve these systems into their components, in general, is an active area of interest. Although instrumentation and computational power have improved since the introduction of the first single molecule fluorescence technique, FCS, resolving heterogeneity remains a challenge.

FCS provides the capability to resolve different components of a given system in two ways. For a system possessing structural fluctuations, such as T4 lysozyme, the theory of FCS allows the ability to resolve the system to within a two-state limit. A second type of system that can potentially be handled by FCS is one consisting of molecules of different sizes. In this case, the diffusion coefficients can be used to determine the composition of the system. However, there are limitations to this approach. First, owing to the weak relationship between molecular weight and diffusion coefficient ($D \propto (\text{molecular weight})^{-1/3}$), for the case of two species, a difference of at least 1.6 in the diffusion coefficients is required to be able to separate the species (64). Second, in a given system, more than two types of species maybe present and resolving this type of a system through diffusion coefficients is not practical. Thus, at best, FCS is able to resolve a two-species or two-state system. A version of FCS that uses higher order correlations (99) has been attempted.

A second approach to resolving systems into their components is the use of photon count statistics methods. Unlike FCS, this approach uses molecular brightness to determine the composition of a system. As discussed in chapter II, cumulant analysis and PCH have this

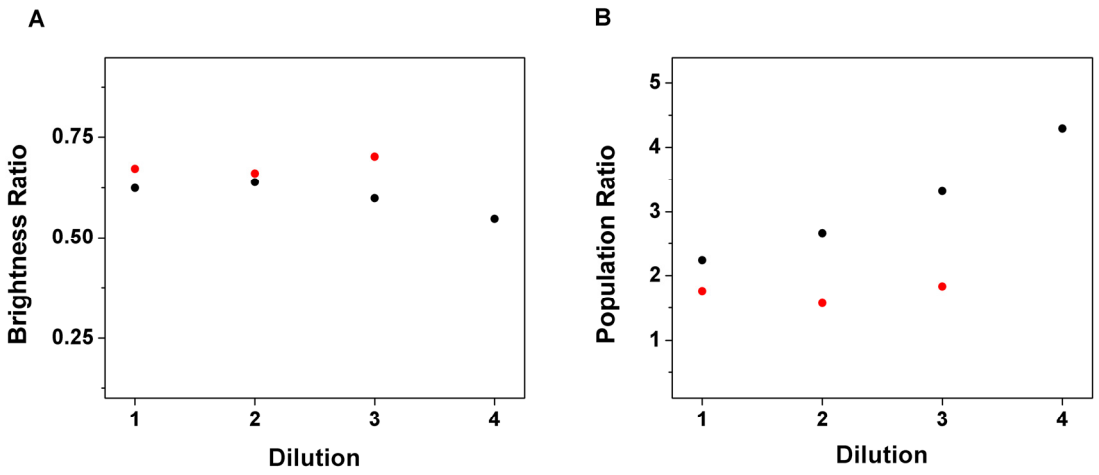


Figure 5-13. *(A)* Molecular brightness and *(B)* population ratio for a mixture system consisting of T4L-65/135-TAMRA and TAMRA. Ratios were taken as the values belonging to the species with lower brightness to that of the higher brightness. In the measurements, for each dilution used, the concentration of TAMRA was twice that of T4L-65/135 with individual concentrations reduced for successive dilutions.

capability. One advantage of cumulant analysis over PCH is in this very feature – while resolving species is possible with both, the mathematical treatment of cumulant analysis, which makes use of factorial cumulants, makes the task significantly easier, as opposed to the distribution approach of PCH (28, 48). As an example of the application of cumulant analysis to resolve species, Fig. 5-3 shows results for a system composed of two species with different molecular brightness properties. For this experiment, a mixture of TAMRA and T4L-65/135 was prepared at a ratio of 2:1, TAMRA to T4L. Furthermore, at each successive dilution, the concentration of both molecules was decreased. Cumulant analysis was applied to determine both the molecular brightness values as well as the corresponding populations. Fig. 5-3 *A* shows that the determined brightness values, given as ratios of the high to low brightness (black circles), compares well to the expected values determined from the individual molecules (red circles). However, the determined populations (Fig. 5-3 *B*, black circles) show deviations from the expected values (red circles). It can be seen that the deviation occurs as the individual populations are decreased.

Although photon count statistics, such as cumulant analysis, provides an alternative to FCS, it has multiple limitations of its own, as illustrated by the example discussed above. As discussed in chapter II, in cumulant analysis, each molecule is characterized by two parameters, the brightness and average number values. Thus, for *n* molecules, there are *2n* parameters and hence *2n* factorial cumulants are required to resolve the system. For more than two species, this task quickly becomes difficult since the higher order cumulants/moments are not statistically as reliable as the lower ones (28, 48). In addition, the application of dead-time and afterpulsing corrections to the higher moments may contain deviations as the corrections are based on approximations to the photon count distribution.

It has also been shown that the ability to successfully resolve, for example, two species is dependent on the brightness of the species. First, the ratio of the brightness of the two species needs to be different by at least a factor of two to properly resolve the system (51). However, as this factor increases, there is the possibility that the dim populations will be obscured. Second, the higher the brightness of the individual molecules, the easier it is to resolve them (51). The implication here is that even if there is at least a factor of two difference in brightness, if the individual brightness values are low, resolving the systems may not be possible. This is precisely due to the fact that it is the molecular brightness that leads to the deviation of the photon count statistics from that of Poisson statistics – the ideal case. The greater the deviation, the easier it is to resolve the system by photon count statistics (48, 51). Not only is the brightness of the species an issue, but there exists a dependence between the respective populations and their brightness ratios. For a given brightness ratio, there is an ideal ratio of populations that is best for resolving the system. Deviations from this specific ratio make it less likely that the system is resolved properly (51).

Therefore, cumulant analysis has multiple serious restrictions in being able to resolve species based on brightness. There has to be a precise balance of brightness and population ratios for the successful analysis of a given system. Furthermore, it is important to note that unlike

PCH, cumulant analysis allows the use of the characteristic diffusion time of the molecules, in being able to determine the respective populations. This can make it easier to resolve species, as in the example given above.

In biological systems, such balance maybe impossible to find. Furthermore, biological systems can often contain more than two species or that the brightness difference between two species or even two states of a single species may not be all that different. This is an important issue for proteins, the primary subjects of this work. Here, the goal has been to probe into the conformational details of the system in a way not possible by ensemble methods. Remembering that in a protein various conformations are a result of weak interactions, the structural difference between the different conformations can be small. Indeed, it has long been recognized that the difference between two conformations for a given protein, say open and closed, can be separated by a small energy difference (18). Thus, this presents a significant challenge to the use of brightness analysis for the purpose of resolving conformations in proteins.

With respect to photon count statistics, it might be best to think in terms of brightness probability distributions as opposed to searching for specific brightness values. A similar approach has been attempted with characteristic diffusion time values in FCS (100). In such an approach, a model free analysis can be utilized. As the significance of protein dynamics in various aspects of protein function continues to be revealed, and as we learn more about the 'kicking and screaming' of proteins, approaches that can give any detail at the single molecule level will be crucial. Thus, further investigations of the theory of photon count statistics approaches, as well as FCS, which can lead to novel experimental methodologies, are warranted.

APPENDIX A

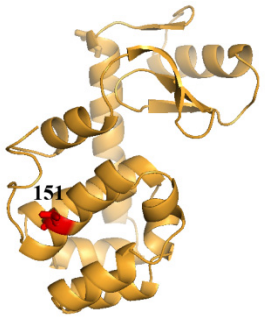
SUPPLEMENTARY MATERIAL

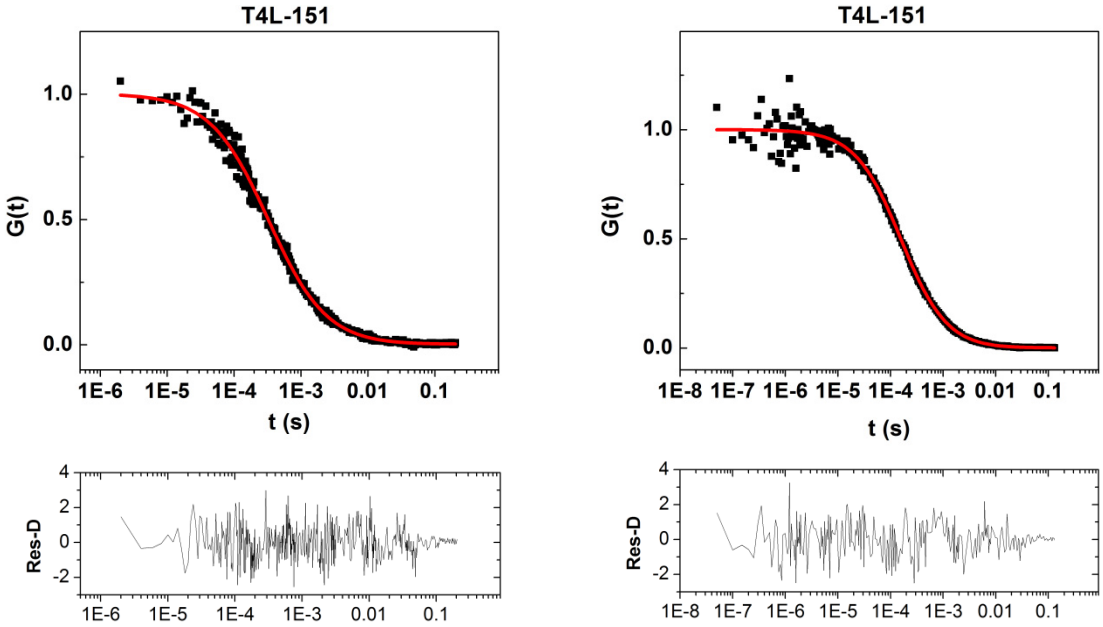
Results from analysis of FCS data for T4L mutants are given. For each mutant, the FCS curves have been fit with one or more of the FCS models. The FCS models are:

1. Diffusion-only (D), Eq. 2.18, in red
2. Diffusion/quenching (D+Q), Eq. 2.24, in green
3. Diffusion/double exponential quenching (D+2Q), Eq. 2.25, in magenta

Averages of the parameters determined for each fit along with the standard deviations are given.

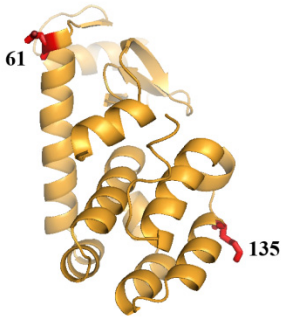
151

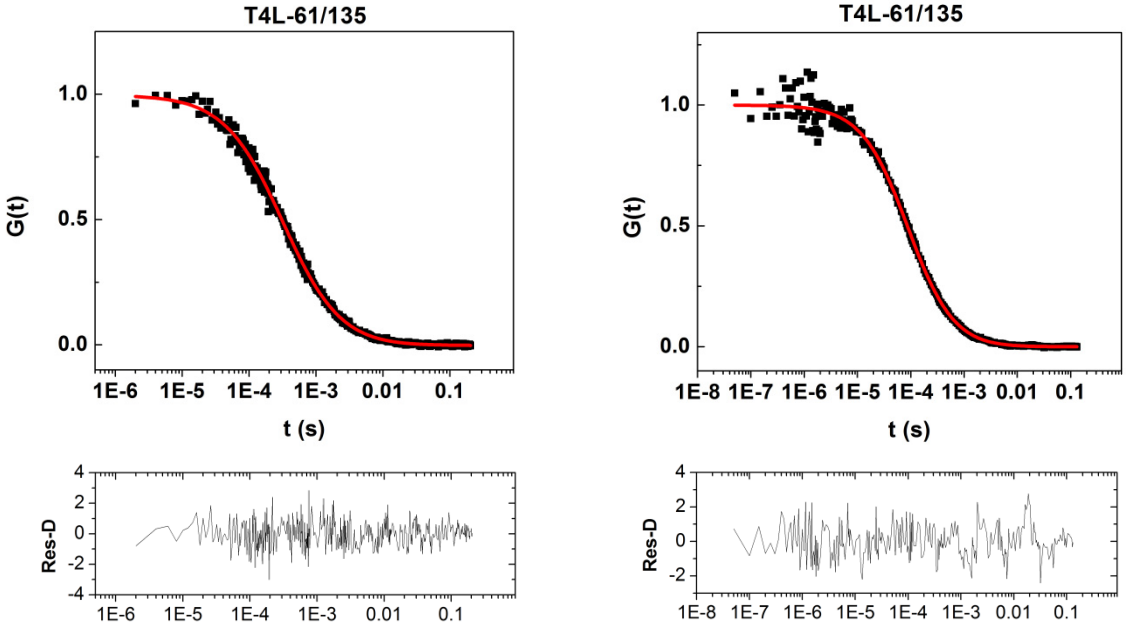




	MODEL	D ($\mu m^2/s$)	Red. χ^2
500 kHz	**D**	84.7 ± 4.4	0.72 ± 0.08
20 MHz	**D**	87.3 ± 2.6	0.79 ± 0.14

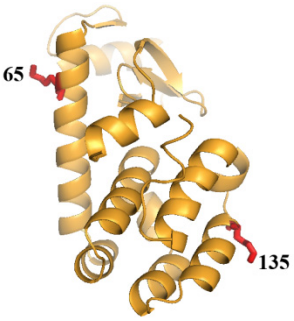
61/135

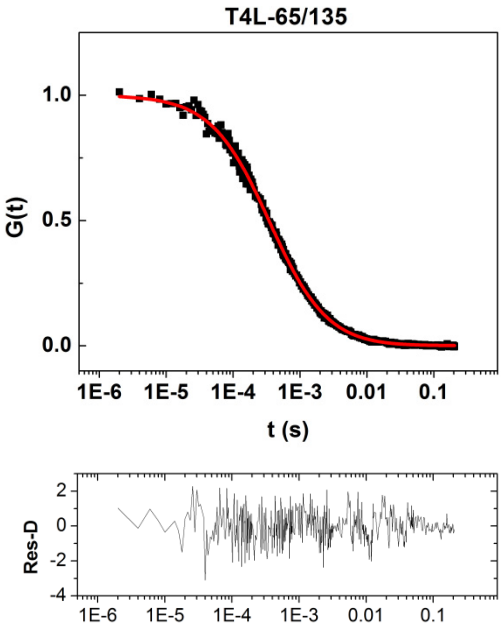




	MODEL	D ($\mu m^2/s$)	Red. χ^2
500 kHz	**D**	81.6 ± 2.9	0.71 ± 0.07
20 MHz	**D**	84.7 ± 1.3	0.85 ± 0.08

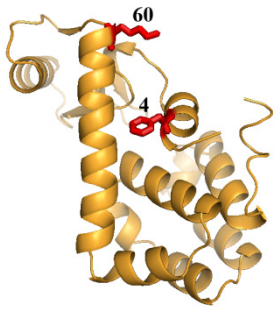
65/135

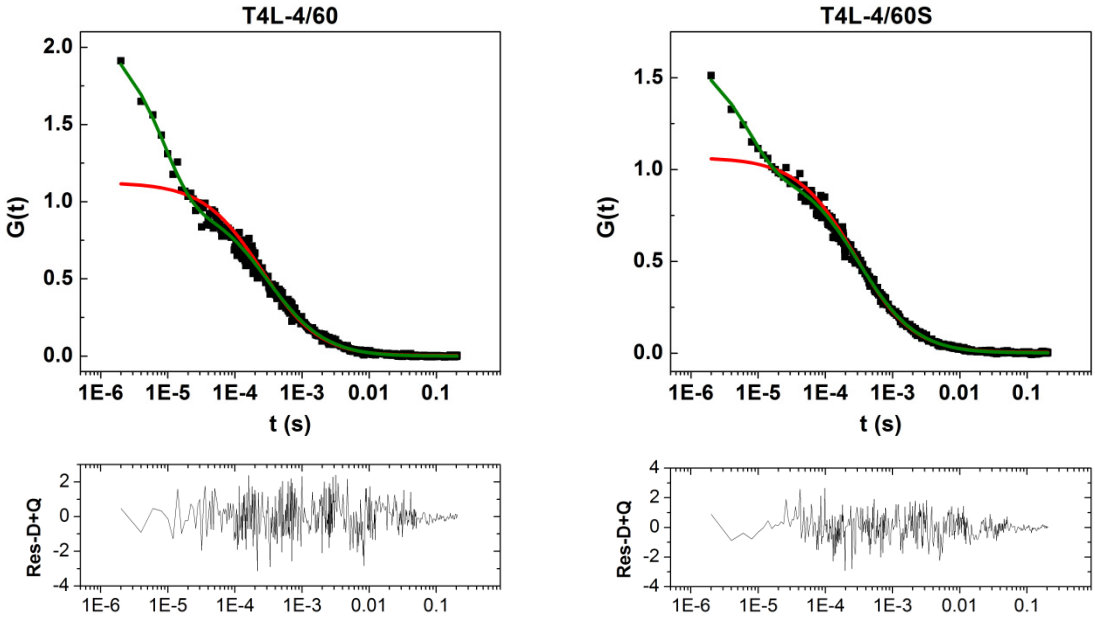




MODEL	D ($\mu m^2/s$)	Red. χ^2
D	82.1 ± 2.6	0.71 ± 0.05

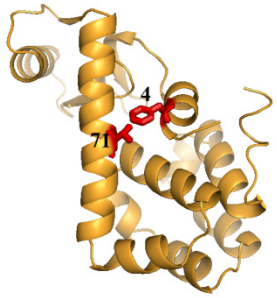
4/60(S)

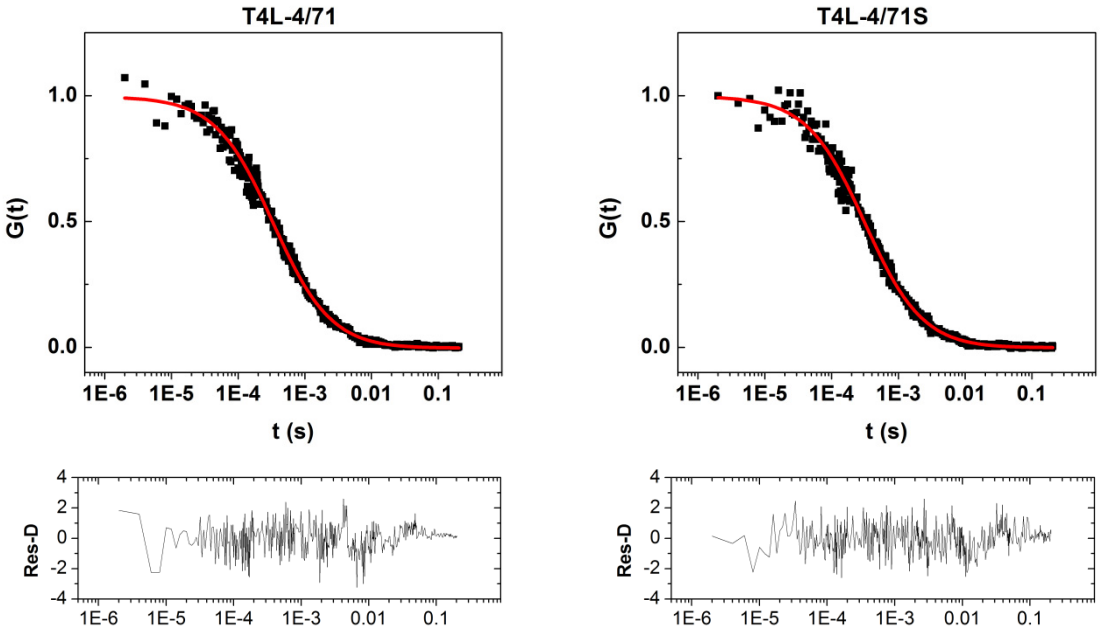




	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
4/60	**D**	114.0 ± 7.6	-	-	1.87 ± 0.40
	D+Q	93.9 ± 4.8	8.5 ± 1.1	1.20 ± 0.10 [54.4]	0.77 ± 0.07
4/60S	**D**	101.7 ± 3.8	-	-	1.55 ± 0.20
	D+Q	91.6 ± 2.5	8.6 ± 1.7	0.61 ± 0.04 [37.9]	0.76 ± 0.10

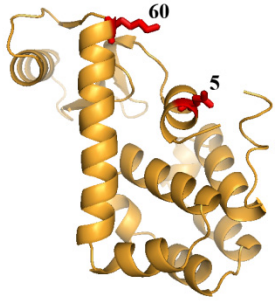
4/71(S)

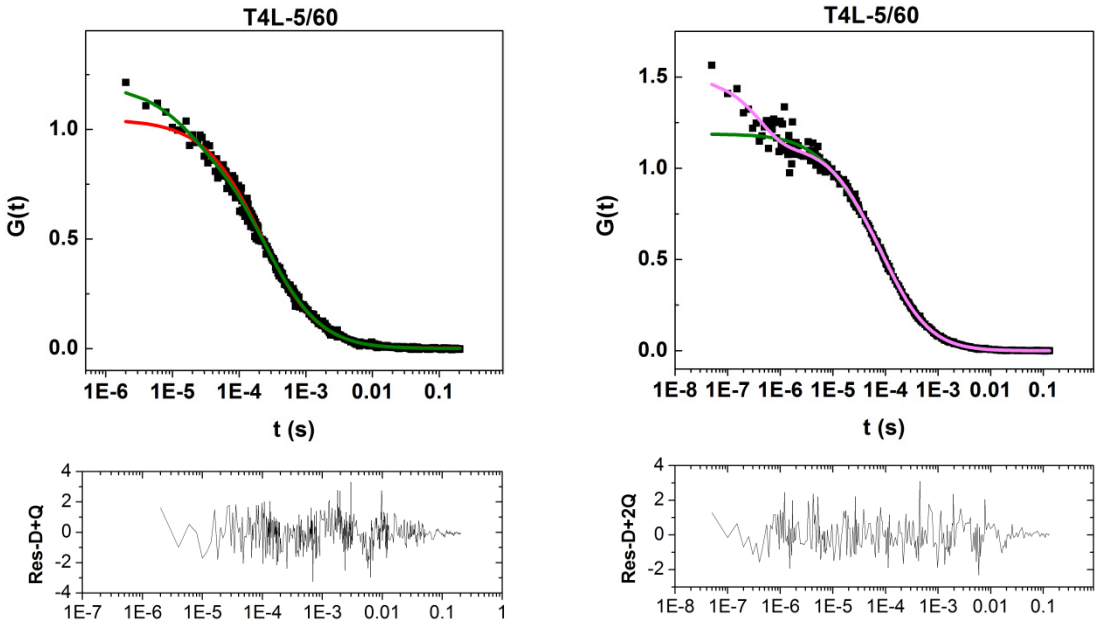




	MODEL	D ($\mu m^2/s$)	Red. χ^2
4/71	**D**	97.6 ± 2.9	0.78 ± 0.04
4/71S	**D**	107.2 ± 5.2	0.88 ± 0.04

5/60

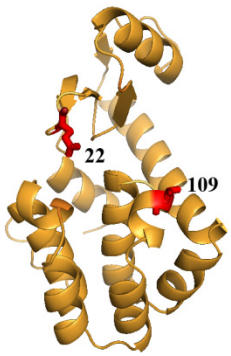


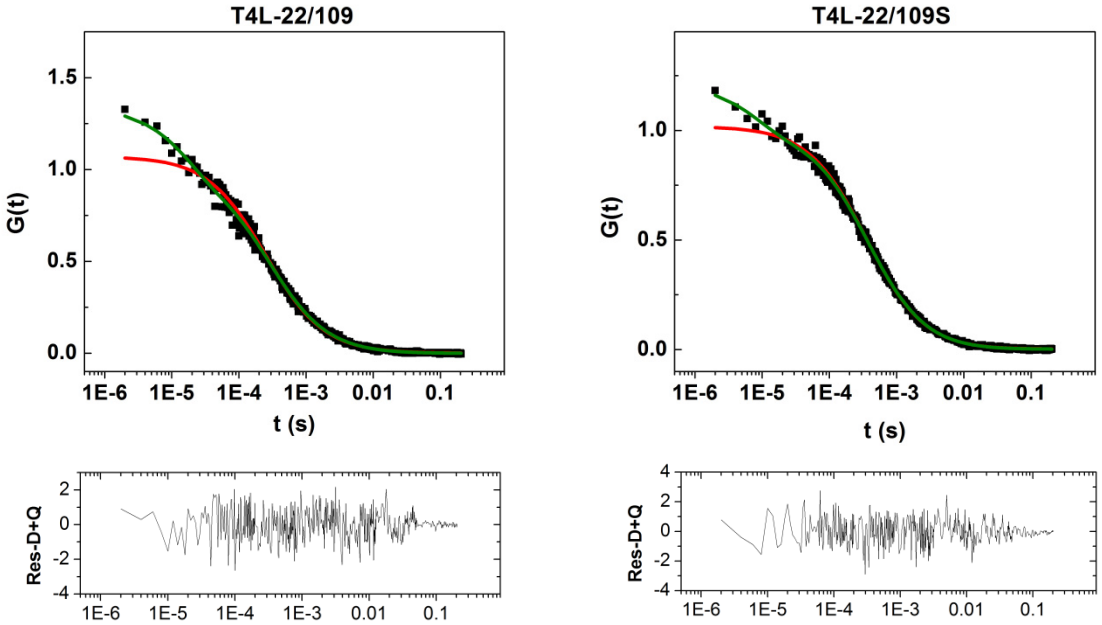


500 kHz				
MODEL	**D ($\mu m^2/s$)**	**τ_R (μs)**	**A [Non-fluor. Pop. (%)]**	**Red. χ^2**
D	107.2 ± 8.1	-	-	0.78 ± 0.20
D+Q	100.6 ± 5.9	17.3 ± 8.7	0.18 ± 0.03 [14.9]	0.69 ± 0.17

20 MHz						
MODEL	**D ($\mu m^2/s$)**	**τ_{R1} (μs)**	**A_1**	**τ_{R2} (μs)**	**A_2**	**Red. χ^2**
D+Q	116 ± 10.4	14.9 ± 6.4	0.20 ± 0.03	-	-	0.92 ± 0.06
D+2Q	113.7 ± 9.2	22.5 ± 6.0	0.18 ± 0.03	0.5 ± 0.2	0.4 ± 0.13	0.77 ± 0.08

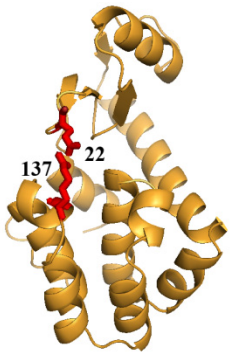
22/109(S)

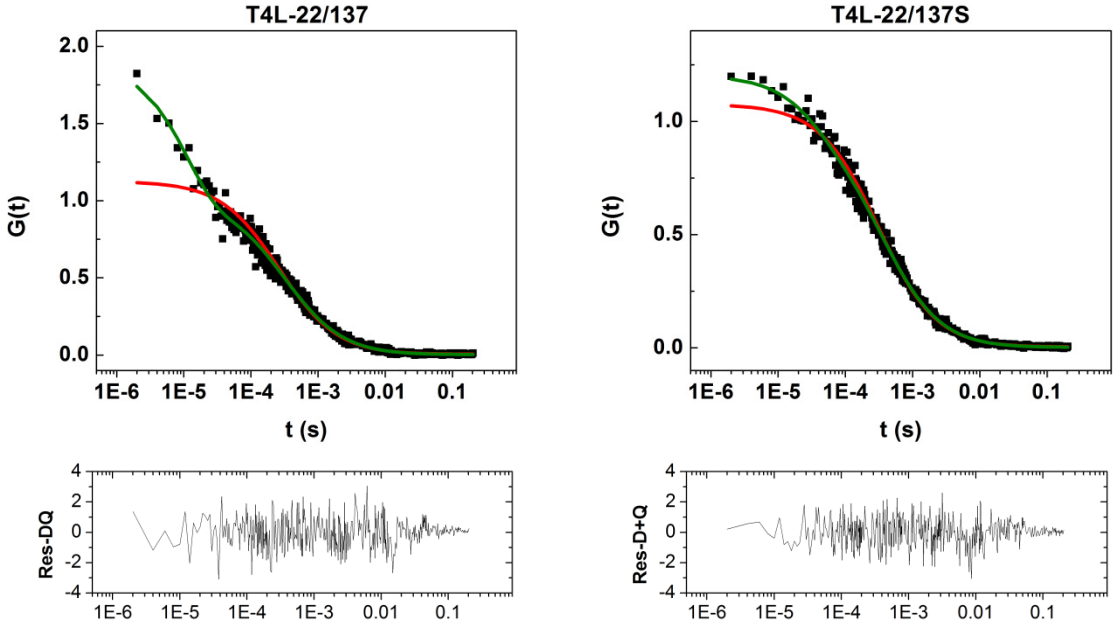




	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
22/109	**D**	113.3 ± 11.2	-	-	1.05 ± 0.12
	D+Q	99.4 ± 6.8	22.6 ± 11.1	0.33 ± 0.04 [25.0]	0.77 ± 0.07
22/109S	**D**	84.8 ± 1.8	-	-	0.86 ± 0.06
	D+Q	82.9 ± 1.5	6.56 ± 2.13	0.22 ± 0.07 [17.7]	0.75 ± 0.08

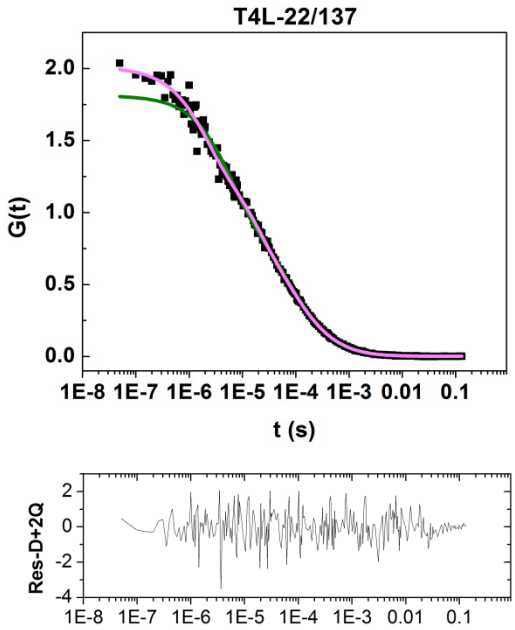
22/137(S)





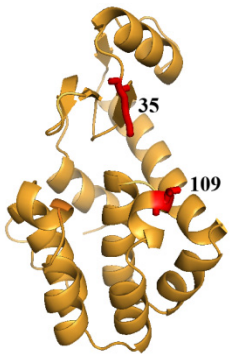
	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
22/137	**D**	125.6 ± 11.8	-	-	1.98 ± 0.5
	D+Q	100.4 ± 8.4	13.3 ± 2.7	0.78 ± 0.08 [43.7]	0.76 ± 0.10
22/137S	**D**	107.0 ± 5.6	-	-	0.84 ± 0.06
	D+Q	97.1 ± 7.2	37.0 ± 17.7	0.15 ± 0.05 [12.7]	0.75 ± 0.03

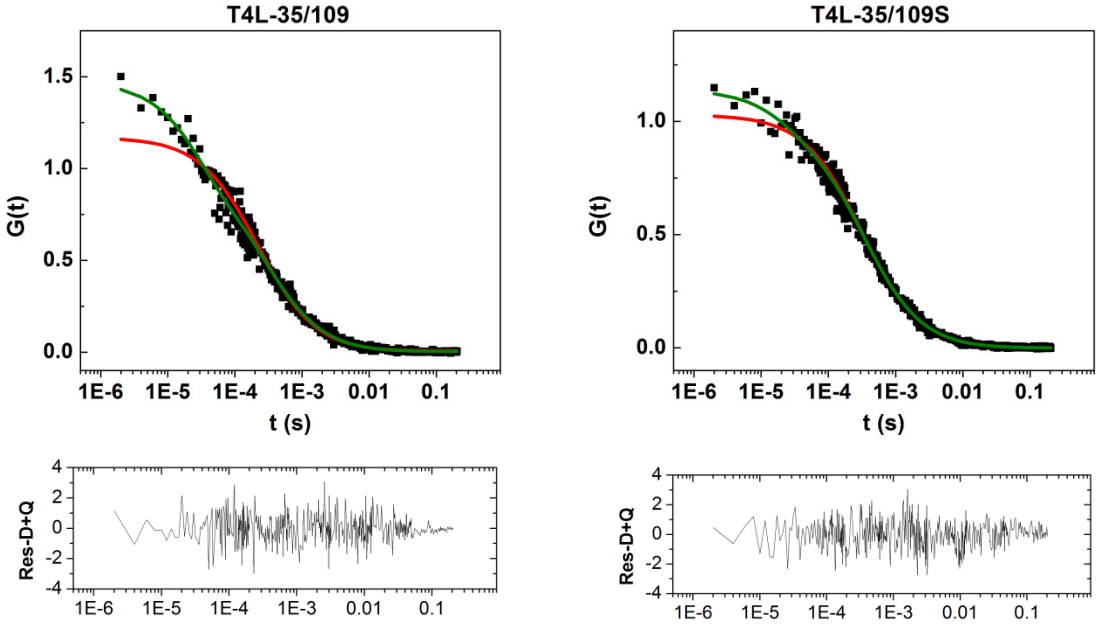
22/137



MODEL	D ($\mu m^2/s$)	τ_{R1} (μs)	A_1	τ_{R2} (μs)	A_2	Red. χ^2
D+Q	110.7 ± 9.3	5.3 ± 0.6	0.69 ± 0.04	-	-	1.06 ± 0.13
D+2Q	101.0 ± 9.0	15.1 ± 4.3	0.36 ± 0.04	2.2 ± 0.4	0.59 ± 0.07	0.80 ± 0.09

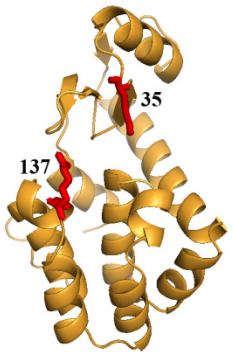
35/109(S)

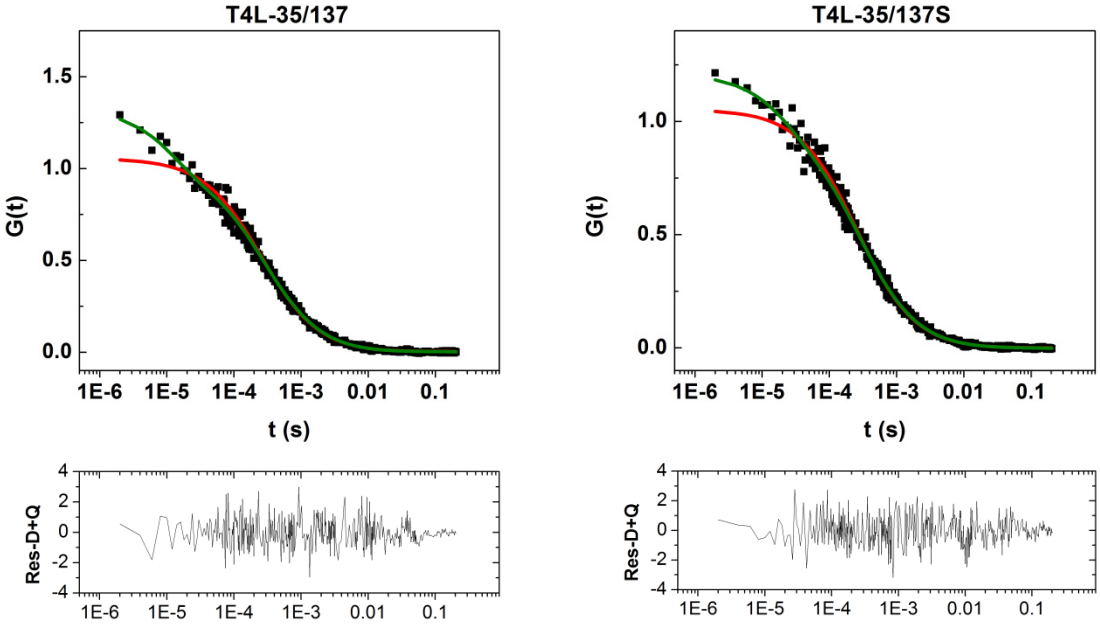




	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
35/109	**D**	134.3 ± 11.9	-	-	1.09 ± 0.08
	D+Q	100.6 ± 9.7	36.3 ± 11.1	0.54 ± 0.06 [34.9]	0.80 ± 0.08
35/109S	**D**	103.4 ± 6.0	-	-	0.80 ± 0.09
	D+Q	96.6 ± 8.2	51.4 ± 33.1	0.11 ± 0.03 [10.0]	0.76 ± 0.08

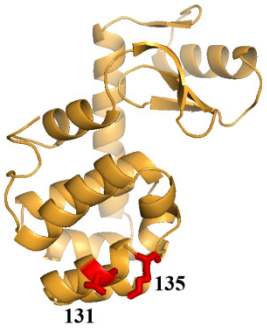
35/137(S)

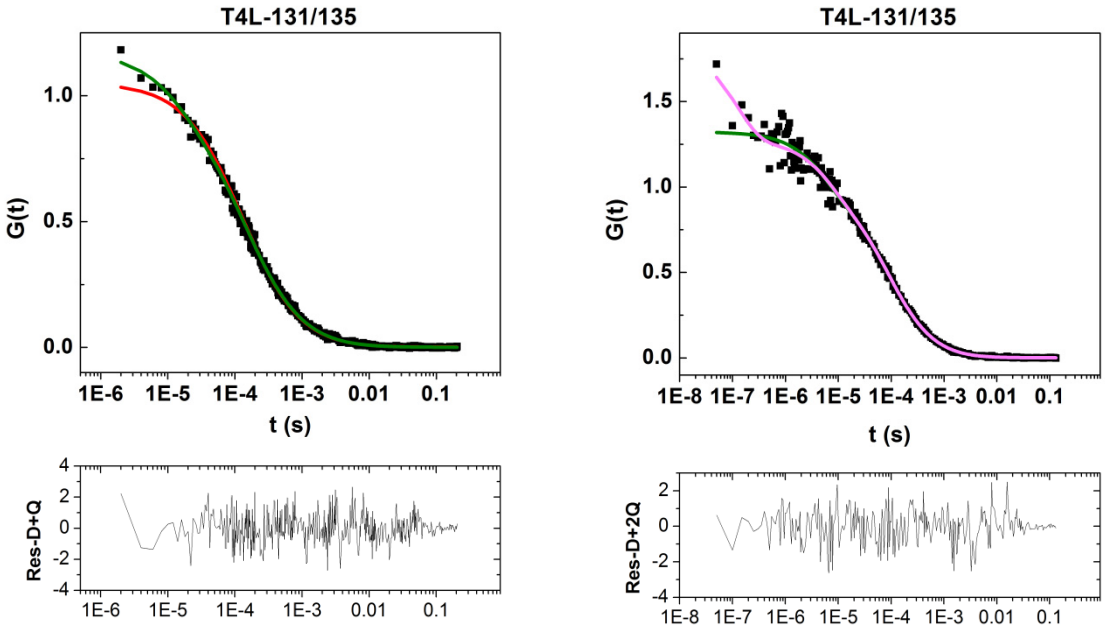




	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
35/137	**D**	104.3 ± 5.8	-	-	0.98 ± 0.09
	D+Q	96.8 ± 6.2	10.2 ± 2.1	0.37 ± 0.04 [26.7]	0.75 ± 0.09
35/137S	**D**	104.6 ± 6.2	-	-	0.83 ± 0.10
	D+Q	96.9 ± 6.3	21.3 ± 12.8	0.21 ± 0.03 [17.5]	0.72 ± 0.09

131/135

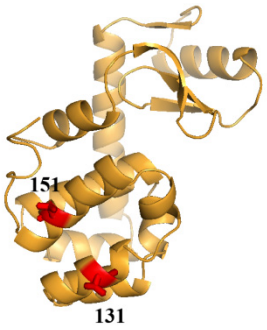


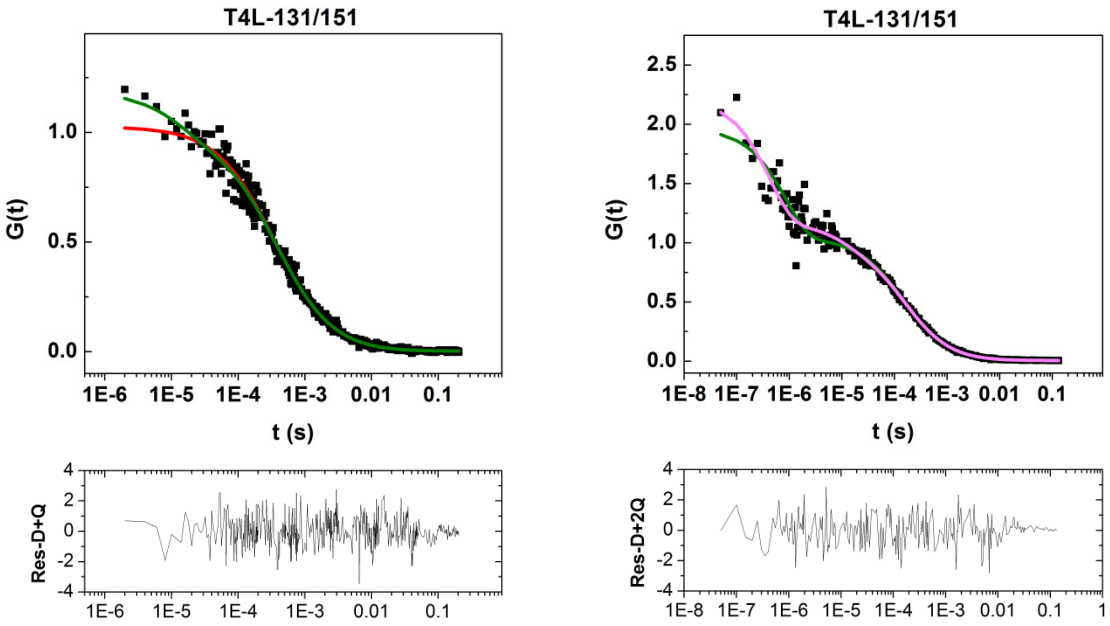


500 kHz				
MODEL	**D ($\mu m^2/s$)**	**τ_R (μs)**	**A [Non-fluor. Pop. (%)]**	**Red. χ^2**
D	119.8 ± 4.4	-	-	0.98 ± 0.08
D+Q	110.1 ± 5.1	17.0 ± 9.61	0.21 ± 0.03 [17.1]	0.79 ± 0.07

20 MHz						
MODEL	**D ($\mu m^2/s$)**	**τ_{R1} (μs)**	**A_1**	**τ_{R2} (μs)**	**A_2**	**Red. χ^2**
D+Q	111.1 ± 15.2	5.8 ± 2.8	0.27 ± 0.05	-	-	0.91 ± 0.12
D+2Q	108.1 ± 14.7	11.5 ± 5.2	0.20 ± 0.04	0.4 ± 0.2	0.49 ± 0.19	0.82 ± 0.14

131/151

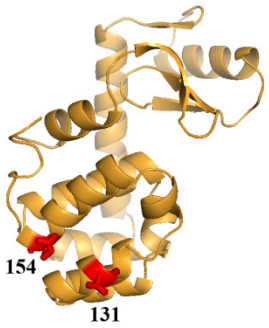


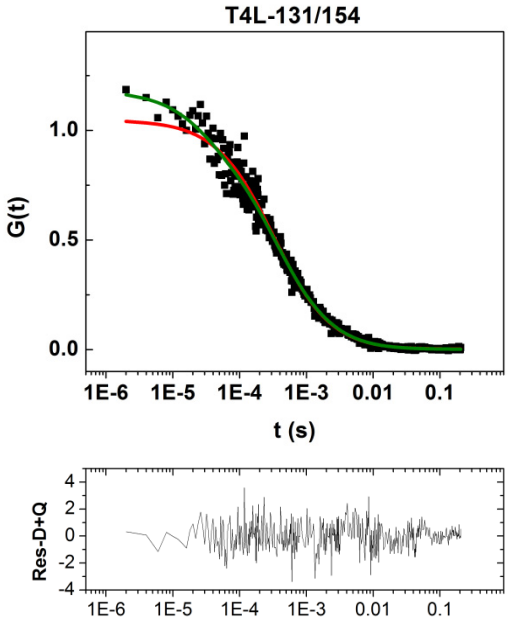


500 kHz				
MODEL	**D ($\mu m^2/s$)**	**τ_R (μs)**	**A [Non-fluor. Pop. (%)]**	**Red. χ^2**
D	106.8 ± 7.2	-	-	0.82 ± 0.09
D+Q	102.9 ± 6.8	10.1 ± 5.2	0.21 ± 0.05 [17.5]	0.77 ± 0.10

20 MHz						
MODEL	**D ($\mu m^2/s$)**	**τ_{R1} (μs)**	**A_1**	**τ_{R2} (μs)**	**A_2**	**Red. χ^2**
D+Q	99.2 ± 6.7	0.7 ± 0.2	0.81 ± 0.12	-	-	1.13 ± 0.33
D+2Q	92.6 ± 7.4	22.9 ± 13.4	0.12 ± 0.03	0.4 ± 0.1	0.91 ± 0.13	0.86 ± 0.17

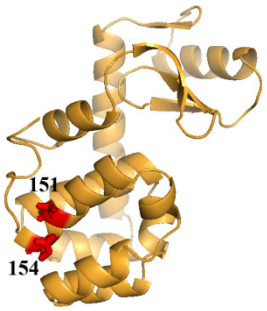
131/154

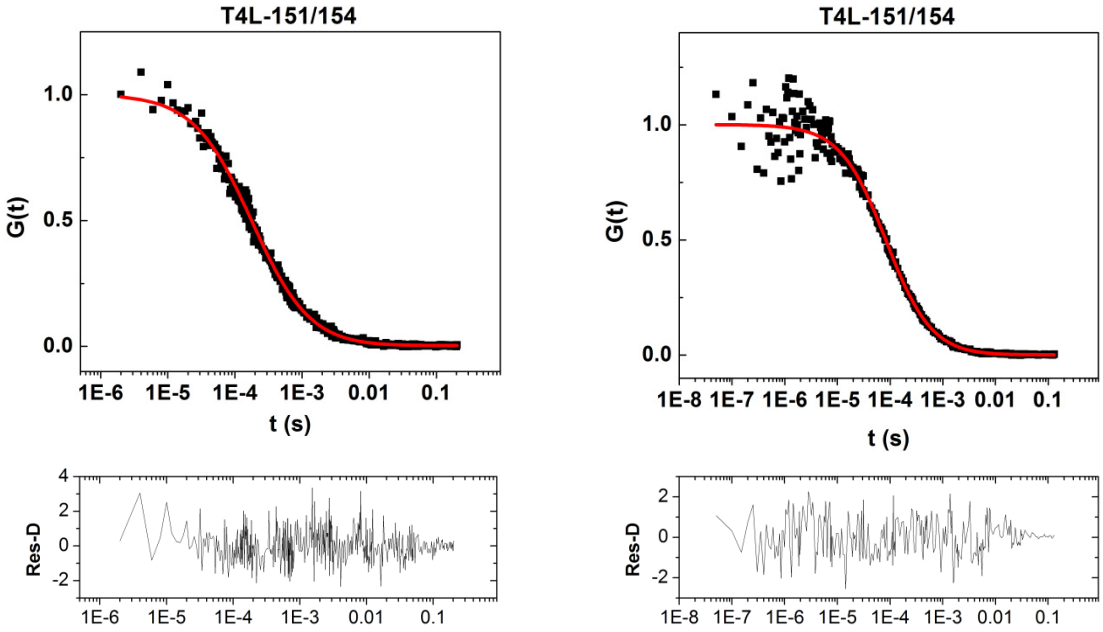




MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
D	109.8 ± 6.4	-	-	0.84 ± 0.10
D+Q	102.6 ± 8.5	32.6 ± 25.0	0.22 ± 0.14 [17.3]	0.79 ± 0.10

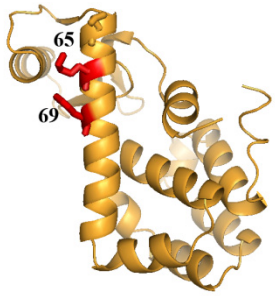
151/154

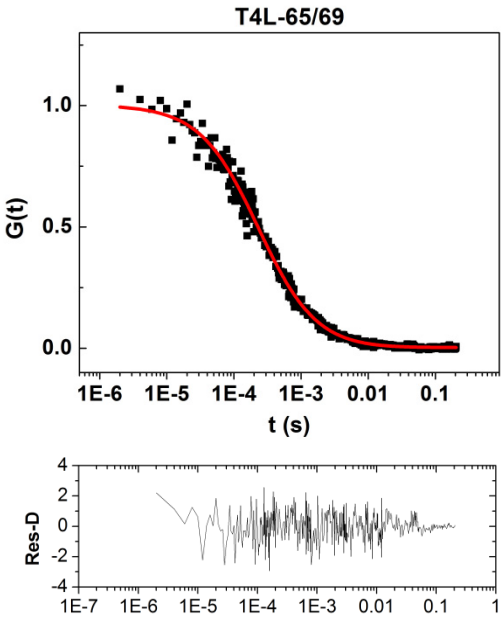




	MODEL	D ($\mu m^2/s$)	Red. χ^2
500 kHz	**D**	120.8 ± 13.4	0.79 ± 0.02
20 MHz	**D**	119.7 ± 12.3	0.88 ± 0.09

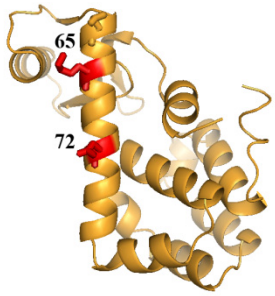
65/69

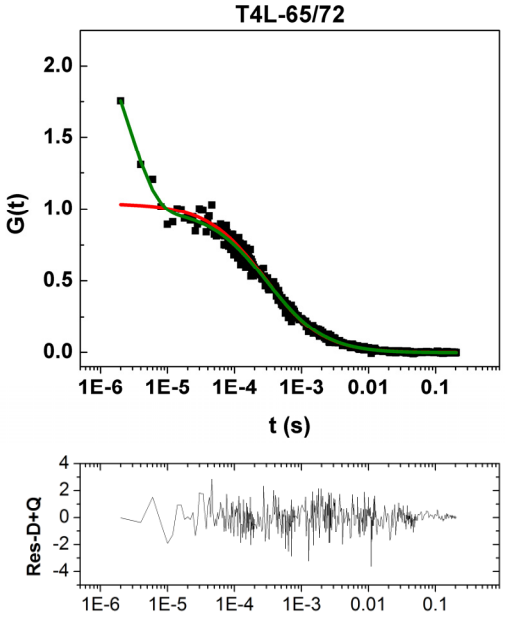




MODEL	D ($\mu m^2/s$)	Red. χ^2
D	104.5 ± 6.7	0.75 ± 0.07

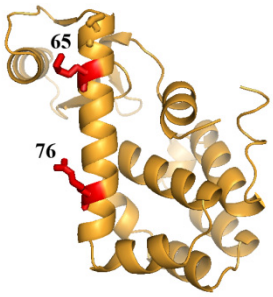
65/72

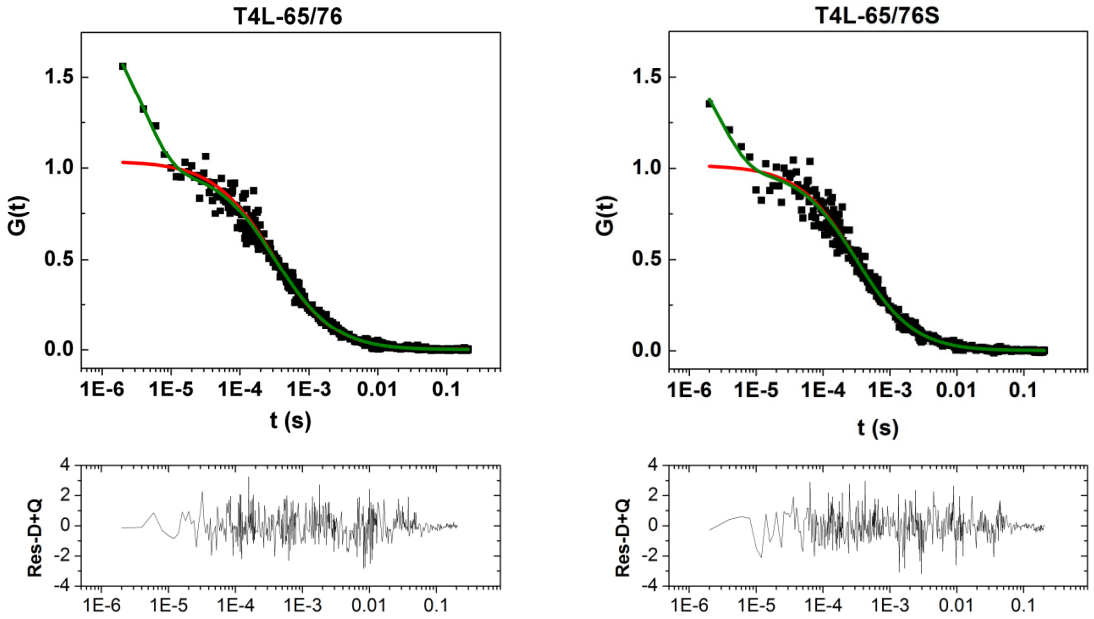




MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
D	102.2 ± 6.4	-	-	1.15 ± 0.17
D+Q	94.2 ± 5.7	3.7 ± 0.4	1.26 ± 0.19 [55.6]	0.71 ± 0.08

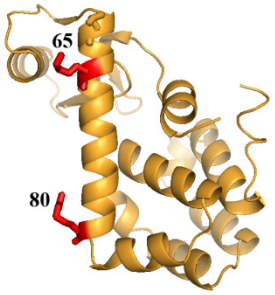
65/76

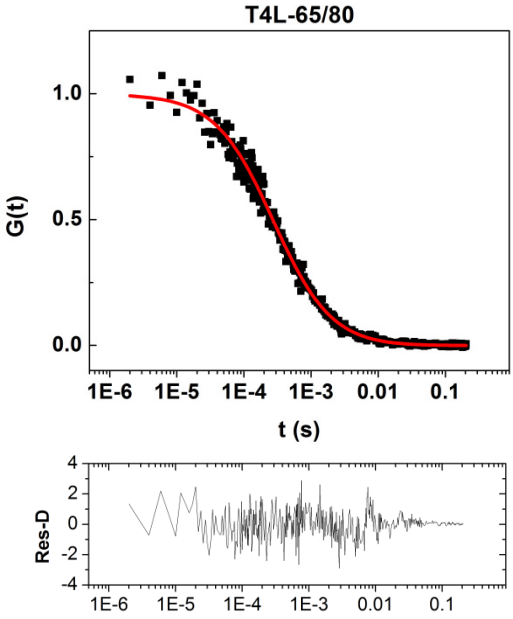




	MODEL	D ($\mu m^2/s$)	τ_R (μs)	A [Non-fluor. Pop. (%)]	Red. χ^2
65/76	**D**	99.3 ± 4.5	-	-	0.94 ± 0.08
	D+Q	92.9 ± 4.1	4.3 ± 1.0	0.89 ± 0.16 [46.8]	0.74 ± 0.06
65/76S	**D**	94.9 ± 1.3	-	-	0.79 ± 0.03
	D+Q	91.9 ± 1.8	2.3 ± 0.5	1.07 ± 0.21 [51.2]	0.70 ± 0.06

65/80





MODEL	D ($\mu m^2/s$)	Red. χ^2
D	84.4 ± 4.5	0.79 ± 0.10

APPENDIX B

SOFTWARE FOR FLUORESCENCE FLUCTUATION SPECTROSCOPY

I. FCSLab

I.1 GETTING STARTED

1. Introduction

   FCSLab is analysis software for fluorescence correlation spectroscopy (FCS). The software was entirely written in object-oriented MATLAB without any external or third-party components or packages. The design employed in the development of FCSLab was motivated by the need for analysis software that affords the highest flexibility and functionality without compromising ease of use. The data files used by FCSLab must be in ISS format.

2. The Work Area

   Starting FCSLab will open three windows or consoles: FCSLab, FCS Data and Fit Model. The first console, FCSLab (Fig. 1), is the main console of the software. This console contains the menu items as well as a file and directory browser. Closing this console either by using the close button in the top right or through the File menu will close the entire software. FCSData (Fig. 1) is used to display the currently loaded data. In addition, actions on the loaded data can be taken directly from this console by using the five buttons located at the top of the console. The third console, Fit Model (Fig. 1), lists the FCS models that can be selected by the user for analysis of FCS data. Detailed description for each of the three consoles will be given in the following sections.

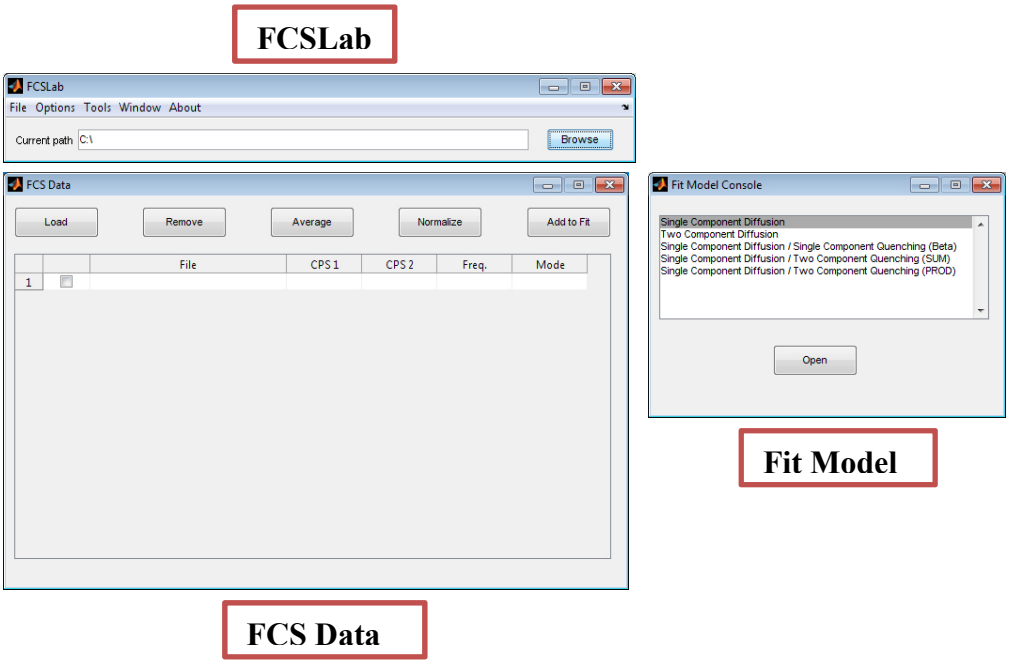


Figure 1. The FCSLab work area showing the three primary consoles.

I.2 DATA ANALYSIS

1. Loading data

   To load data, click the Load button on the FCSData console (A in Fig. 2). Then select the data file to load. Multiple data can be loaded at the same time. The loaded data will be listed in the table (B in Fig. 2) with the basic information displayed. A plot of the data will be automatically generated and displayed in a new window (Fig. 3).

   NOTE: FCSLab uses the ISS format for FCS data which are CSV files and contain information about the data in the header.

2. Fitting

   a. If the data to be fit is not selected, select by clicking on the check box in the first column of the table in the FCS Data console (C in Fig. 2). The check box in the last row will select all of the loaded data (Select-All).

b. Select the FCS model to be used for fitting from the Fit Model console (Fig. 4). The selected model will be highlighted.

c. Click the Add to Fit button on the FCS Data console (D in Fig. 2). This will open the Fit console (Fig. 5) corresponding to the selected model. The selected FCS data will be added to the table in the Fit console (A in Fig. 5). The table displays the parameters for the corresponding model in separate columns.

d. The parameter columns will contain default values (B in Fig. 5). These values will be the initial values for the fit. If different initial values are desired for one or more of the parameters, they can be changed directly in the table.

e. Above each parameter column is a check box (C in Fig. 5). Setting this check box will fix a parameter for fitting. The value displayed in the table will be used for the fixed value.

f. Once the initial values and fixed parameters are set as desired, click on the Fit button at the bottom to perform fitting (D in Fig. 5). Only the data that are selected by setting the check box in the first column of the table will be fit.

g. After the fit is complete, the final values along with the reduced chi-square will be displayed. A fit curve will be shown in the corresponding figure of the FCS data (Fig. 6).

3. Exporting

Once fitting is complete, the results can be exported by clicking on the Export button at the top right of the Fit Model console (E in Fig. 5). For a given fit, the parameters and the fit correlation curve can be exported. The exported files are in a CSV format.

4. Chi-Square Surface Search

To perform a chi-square surface search, go to the Tools menu in FCSLab and select CSS Search. In the CSS Search window, specify the folder where the input files are located or select specific input files. The input file must be a data/fit file exported by FCSLab. In

addition, the data/fit file must have been exported with parameters included in the header. Search settings can be specified by going to the Options menu and selecting General and Batch under CSS Search (see section V).

Click Start to begin the chi-square surface search. For each input file, CSS Search will produce three files whose file names end with the following strings: _Corr, _ParamOut and _Params_CorrData.

a. _Corr – gives the fit values at each iteration. The settings used for the search as well as the starting and ending parameter values are given in the header.

b. _ParamOut – gives the parameter values at each iteration. The settings used for the search as well as the starting and ending parameter values are given in the header.

c. _Params_CorrData – contains only the initial and final two fit values. The corresponding parameters are saved in the header. This file can be imported into FCSLab by going to the File menu in FCSLab console and selecting Import Data/Fit.

All of the output files will be saved to a directory with the name FCSLab_CSS_<model id>, where <model id> is an identifier string of the FCS model used.

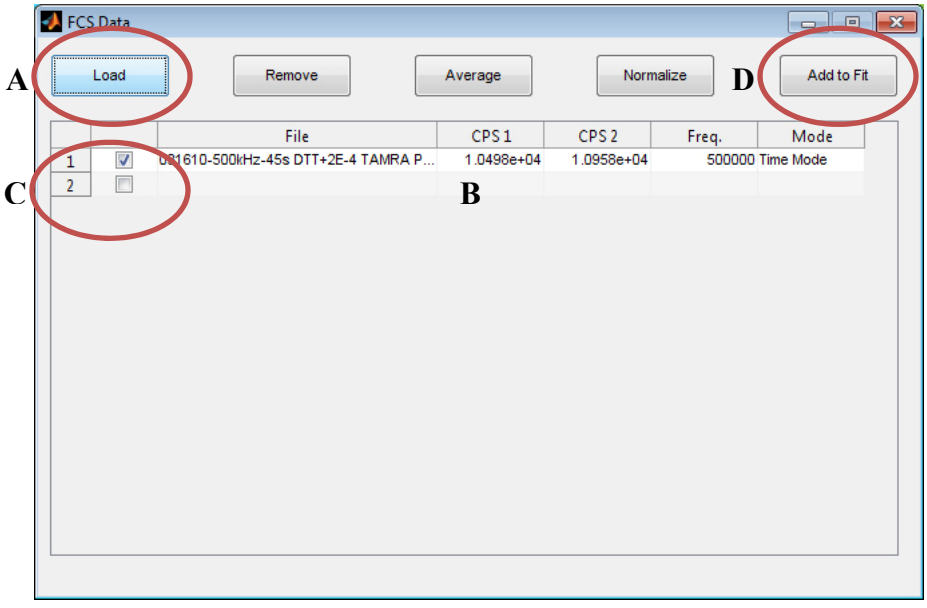


Figure 2. The FCS Data console.

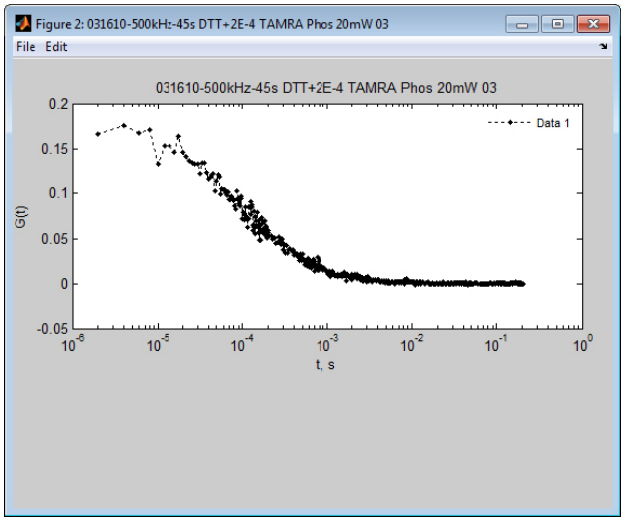


Figure 3. An example FCS data plot.

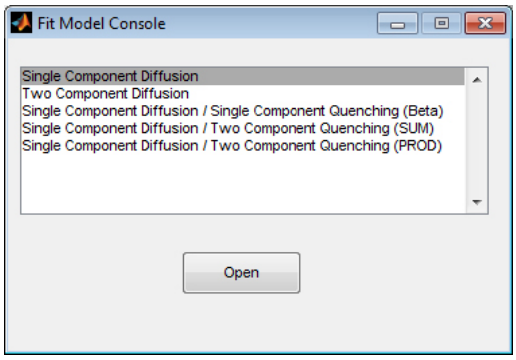


Figure 4. The Fit Model console.

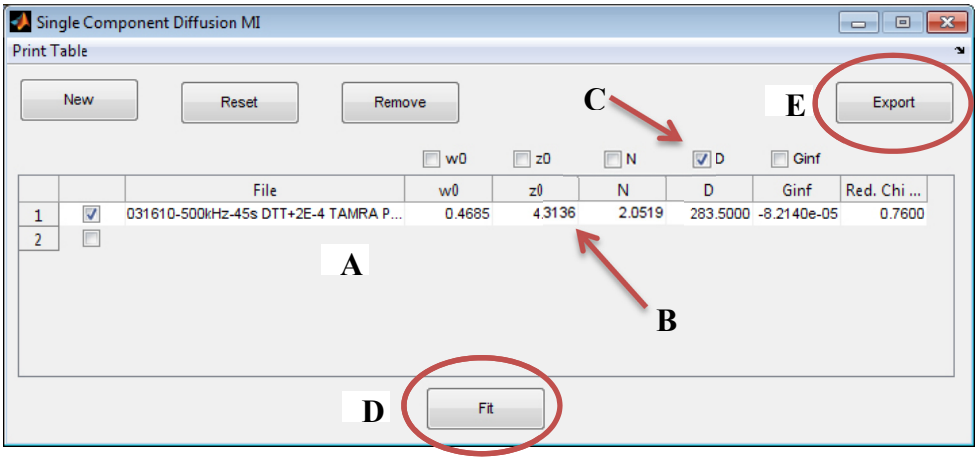


Figure 5. The single component diffusion model Fit console. The table shows one FCS data with the file name given in the second column. The five parameter values are shown in columns 3 to 7. The final column gives the current reduced chi-square value.

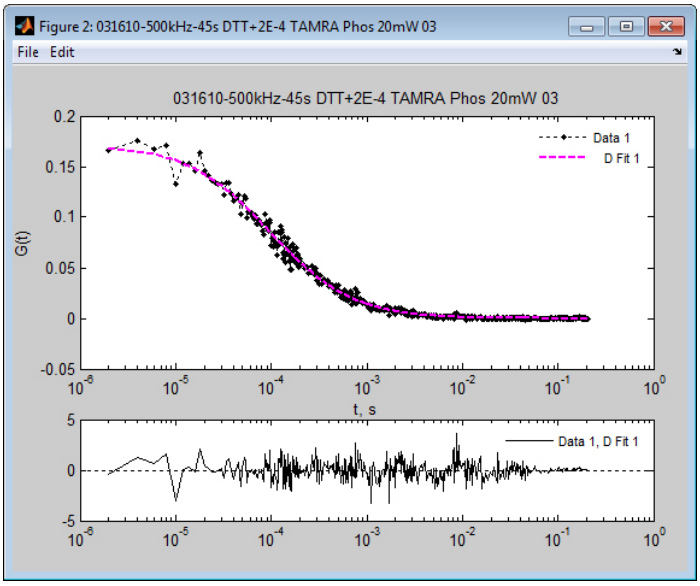


Figure 6. A figure window with data (black) and fit (magenta) correlation curves. The residual is displayed at the bottom.

I.3 DATA HANDLING OPERATIONS

1. Averaging

   A set of FCS data can be averaged using the Average feature. To perform averaging, load data and select the ones to be averaged. Then click on the Average button on the FCS Data console. The averaged data will be listed as a new data on the table in FCS Data console. The averaged curve will be plotted in the same figure window where the original data are plotted. The curve can be saved to a file by going to the File menu in the FCSLab console and selecting either Save FCS Data (ISS) or Save FCS Data. The first will save the averaged data in ISS CSV format while the second will save in a basic CSV format (see section V).

2. Normalizing

   One or more FCS data can be normalized by using the Normalize feature on FCS Data console. With the desired data selected in the FCS Data console, click on the Normalize button. The correlation curves can be normalized to a value determined by averaging a

specified number of the initial correlation points (G(0) option , Fig. 7) or to a value specified by the user (Custom option, Fig. 7). The normalized curves, one for each selected data, will be listed in the table in FCS Data console with the string NORM appended to the beginning of the file name.

The normalized curves will be plotted in the same figure window where the original data are plotted. The curves can be saved to a file by going to the File menu in the FCSLab console and selecting either Save FCS Data (ISS) or Save FCS Data. The first will save the normalized data in ISS CSV format while the second will save in a basic CSV format (see section V).

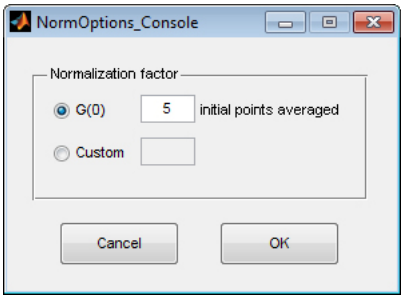


Figure 7. Normalization options.

3. Removing Duplicate Lag Times

The ISS Vista software duplicates lag time points and the corresponding correlation and standard deviation values when saving experimental FCS data. To remove these duplicated entries from FCS data files, use the Remove Duplicate Lag Times function of FCSLab. The removal of duplicated entries can be performed either on the currently loaded data or as a batch process.

a. Removing duplicates from loaded data

With the FCS data selected in FCS Data console, go to the Tools menu in FCSLab and select Remove Duplicate Lag Times. The duplicate entries will be removed from each

data and new duplicate free data will be added to the table in FCSData console. The new data will have the string DF_ appended to the beginning of the file name. Each new duplicate free data will be plotted in the same figure window where the original data are plotted. To save the duplicate free data, go to the File menu in the FCSLab console and select either Save FCS Data (ISS) or Save FCS Data. The first will save data in ISS CSV format while the second will save in a basic CSV format (see section V).

b. Removing duplicates as a batch process

Duplicate entries can be removed without first loading the data into FCSLab. To do this, go to the Tools menu in FCSLab and select Remove Duplicate Lag Times BATCH. In the Remove Duplicates (BATCH) console, use the Browse button to navigate to the appropriate folder. If subdirectories are to be included, then select the Traverse Directories check box. Finally, click on the Start button at the bottom to begin removing duplicate entries from all data files in the specified directory and subdirectories (if selected).

I.4 SOFTWARE COMPONENTS & FEATURES LIST

1. FCSLab
    a. File
        - New Session – closes the current session (all consoles and windows) and starts a new FCSLab session.
        - Load Session – loads a previously saved FCSLab session. If loading a session while in the middle of a session, the session to be loaded can be added to the current one. On the other hand, the current session can be closed and the desired session loaded.

- Save Session – saves the current FCSLab session to a text file with extension .FLB. When a session is saved, information for all of the currently loaded FCS data as well as all fit related information for each FCS data will be saved. Saving a session allows one to interrupt data analysis which then can be continued at a later time by loading the saved session.
- Load FCS Data – loads data saved in ISS file format.
- Save FCS Data (ISS) – saves the specified data in the ISS file format. This format contains a header.
- Save FCS Data – saves the specified data to a .CSV file.
- Import Data/Fit – imports data and fit correlation curves that were exported by FCSLab. The imported curves will be automatically plotted in a new figure window.
- Exit – closes FCSLab.

b. Options

- Fit Preferences
    - Algorithm – the type of algorithm to be used for fitting can be set to either Trust-Region-Reflective or Levenberg-Marquardt.
    - Display – selecting Iter opens a new window that will allow the user to monitor a fit at each step of the iteration.
    - Max Fun. Evals. – maximum number of function evaluations
    - Max Iteration – maximum number of fit iterations allowed. Fitting will not continue past the specified number of iterations.
    - Function Tol. – the amount by which the function value changes during the iteration. Fit will terminate if the change is smaller than the specified value.

- Parameter Tol. – the amount by which the parameter values change during the iteration. Fit will terminate if the change is smaller than the specified value.
- Weight – select Internal to use weights for fitting.

All of the above, with the exception of the Weight option, are MATLAB optimization options.

- Parameter Bounds – specify the lower and upper bounds for all parameters. These values will be used only when the Trust-Region-Reflective algorithm is selected.
- Plot Preferences
  - On load plotting – selecting Separately will plot each loaded data in a separate figure window. To plot data in the same window, select Together. When loading data to the same figure window, the curves are plotted with randomly generated colors. If loading to separate windows, the curves are plotted in black.
  - Plot Standard Deviation – displays the standard deviations in a separate figure window.
  - Show Error Bars – selecting this option displays the standard deviations as error bars on the correlation curve.
  - Show Legend – display figure legends.
  - Only Show Selected Data/Fits – if this option is selected, then only the data and fit that are selected (check boxes in respective Tables) will be displayed. The residual for a fit curve that is *not* selected will be displayed in light grey.

- Averaging Options – if Internal is selected, standard deviation values will be calculated when averaging FCS data. These values are calculated by propagating the standard deviation values of the original data.
- CSS Search
    - General
        - Termination (Lower/Upper) – the values specified for the parameters will be used for termination of a chi-square surface search.
        - Chi Sq. – search will terminate if the chi-square value is greater than this value.
        - Steps – the number of steps or iterations to take during a chi-square surface search.
        - Search Resolution – during a chi-square surface search, the axis parameter (the parameter held fixed) will be incremented by these values.
        - Scale – this value will be applied to the Search Resolution values when an order of magnitude change is encountered between two successive axis parameter values. The scale value adjusts the increment values accordingly. If no adjustment is desired, set this value to 1.
    - Batch – the selected FCS model and parameter are used for chi-square surface search. The selected parameter is the axis parameter – the parameter that will be held fixed and incremented iteratively. The selected model must match the input file that will be used for the search, i.e., the input file must have been a data/fit export of the same model.

See Beechem et al. (reference 89) for implementation details.

c. Tools

- Shift Curves – shift correlation curves vertically by a specified amount. The shift amount can be determined by averaging a specified number of the final correlation points. A specific value from the user can also be used.
- Remove Duplicate Lag Times – for currently loaded data, this tool removes duplicate lag time, correlation and standard deviation values (see section IV).
- Remove Duplicate Lag Times BATCH – this tool can be used to remove duplicate entries from FCS data files (ISS) without having to load the data into FCSLab (see section IV).
- CSS Search – a chi-square surface search can be performed using this tool (see section III).

d. Window

- FCS Models – opens the Fit Model console.
- Standard Deviation Plots – opens figure windows with standard deviation plots, if any.

e. About – credits and current version of FCSLab.

2. FCS Data

a. Load – loads data into FCSLab. Multiple data can be loaded simultaneously.

b. Remove – the selected data are removed from FCSLab. If the data have been fit, then the fits will also be removed.

c. Average – generates an average curve from the selected data (see section IV).

d. Normalize – normalizes the selected data (see section IV).

e. Add to Fit – the selected data are added to the Fit console belonging to the model selected in Fit Model console.

3. Fit Model

    a. Models list – select the FCS model to be used for fitting from this list.

    b. Open – displays the Fit console corresponding to the selected model.

    Two-photon excitation is assumed. Descriptions of the models can be found in chapter II as well as in the literature (references 27, 34, 37, 64, 65).

4. Fit Console

    a. New – duplicates the selected fits.

    b. Reset – the parameters for the selected fits are set to the default values. The chi-square value and the fit curve (figure window) are updated.

    c. Remove – removes the selected fit from the Fit console. Note: this does not remove the data, only the fit to the data.

    d. Export – parameters and curves for the selected fits can be exported to a text file in a CSV format. The export options are:

        - Export data with fits – select to export the fit and experimental correlation curves along with the residuals and standard deviation values to a CSV file. If Include Fit Parameters is selected, the corresponding parameter values for the data/fit being exported will be appended in the header. If multiple data/fit are being exported, selecting Export To The Same File will save all in the same CSV file. If a custom file name is desired, deselect Use Default File Name and enter a file name in the text box. Otherwise, the data file name will be used for the export file name (default). Note: if the file already exists, it will be overwritten.
        - Export fit parameters – select to export the parameter values as well as the reduced chi-square value to a CSV file. If a custom file name is desired, deselect Use Default File Name and enter a file name in the text box below.

Otherwise the data file name will be used for the export file name (default).

Note: if the export file already exists, it will be appended to.

In either case, the exported file will be saved to a directory called FCSLab_<model name>, where <model name> is an identifier string corresponding to the FCS model being used. The model identifier string will also be appended to the file names. In addition, for the parameter file, the string _ParamOut will be appended to the end of the file name.

e. Fit – performs fit for the selected fit instances. The fit is performed using current parameter values as initial values for the free parameters and as fixed values for the fixed parameters. The current Fit Options will be used.

NOTE: closing a Fit console will remove all fits.

5. Plot Figure

a. File

- Load Data – use this option to load FCS data that will be plotted in the current figure window.
- Print – the figure window with the correlation/fit curves as well as the residuals can be printed as displayed.
- Export – the figure window can be exported to image files and as a PDF. The available image file formats are:
  - Enhanced meta file (*.emf)
  - Windows bitmap (*.bmp)
  - JPEG image (*.jpg)
  - Portable Network Graphics (*.png)
  - Tiff image (*.tif)

b. Edit

- Curve Appearance – change the color, marker type and line style of correlation and fit curves as well as the residual.
- Update Title – display a custom figure title.

NOTE: closing a figure window will remove all data and fits plotted in the figure from FCSLab.

II. FCSSim

II.1 GETTING STARTED

1. Introduction

   FCSSim is a software for generating fluorescence correlation spectroscopy (FCS) curves. The software was entirely written in object-oriented MATLAB without any external or third-party components or packages. The design of FCSSim mirrors the approach taken in the design of the FCS analysis software, FCSLab. Although these two are independent applications, they share similar design features. The software has the capability to generate correlation curves for various models. The generated curves can be exported to a file, printed or exported to an image file.

2. The Work Area

   When FCSSim is started, two windows or consoles will be displayed. FCSSim (Fig. 1) is the main console of the application and contains menu items. The FCS Models (Fig. 1) console lists the various FCS models that can be used to generate curves.

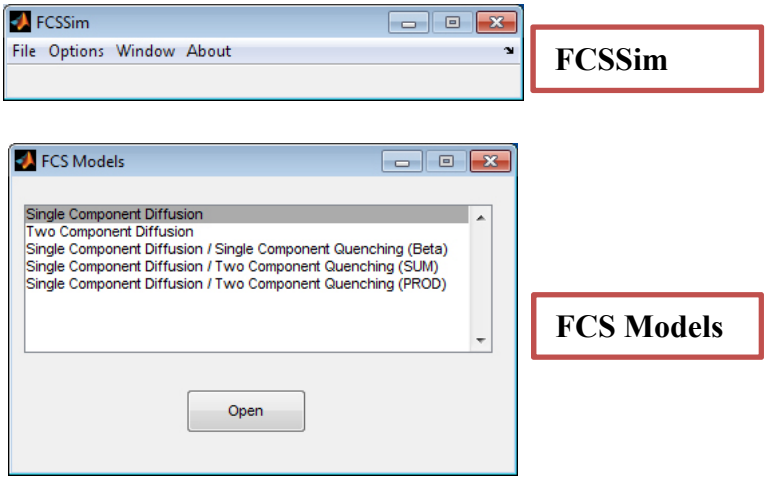


Figure 1. The FCSSim work area with the two primary consoles.

II.2 CREATING FCS CURVES

1. Select a model

   Before a curve can be generated, an FCS model must be selected. In the FCS Models console, click on the desired model from the list and click the Open button. This will open two new windows: an FCS Curves console and a corresponding Controls console (Fig. 2).

2. Create a curve

   In the Curves console, click the New button (A, in Fig. 2) to create an FCS curve. In generating the curve, the lag time points are determined based on the current settings (see section III). In addition, default parameters of the selected FCS model will be used to calculate the curve. The parameters used to calculate the curve will be displayed in the table (B, in Fig. 2). A plot of the curve will be generated and displayed in a figure window (Fig. 3). Curves are plotted with randomly generated colors.

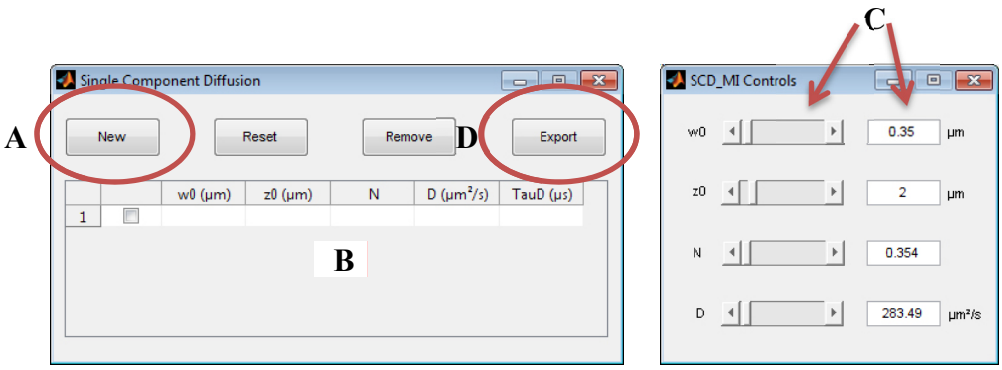


Figure 2. The FCS Curves and Controls consoles for the single component diffusion model.

3. Modifying a curve

   FCS curves can be modified by changing the respective parameters. This can be done in two ways: by directly changing a parameter value in the table or using the Controls console (Fig. 2). When a parameter is changed, the plot of the correlation curve will be automatically updated to reflect the change. Changing a parameter in the table will update the matching scroll-bar and text box (C in Fig. 2) in the Controls console. Likewise, changing a parameter through the Controls console using either the scroll-bars or text boxes will update the matching value in the table. Furthermore, updating a parameter in the table will update all of the curves in the table if the Select-All check box (the check box in the last row) is set. Updating through the Controls console will only apply to the curves that are selected in the table. If the Select-All check box (check box in last row) is set, changes will apply to all curves.

4. Exporting

   FCS curves can be exported by clicking on the Export button (D in Fig. 2). Both the parameter values and the generated curves will be exported.

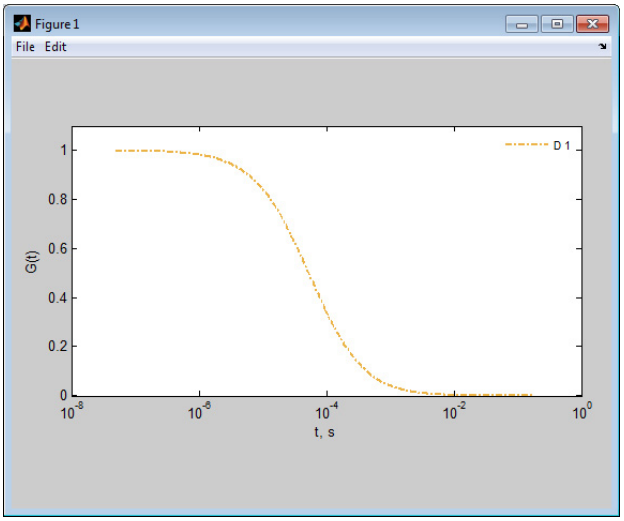


Figure 3. An example curve corresponding to the single component diffusion model.

II.3 SOFTWARE COMPONENTS & FEATURES LIST

1. FCSSim

    a. File

- New Figure – opens a new figure window. This new figure window becomes the active figure and all subsequent curves will be plotted here. To make another figure the active one, go to the File menu in the figure and click on Set Active. This will automatically deactivate all other figures.
- New Session – closes the current session (all consoles and windows) and starts a new FCSSim session.
- Load Session – loads a previously saved FCSSim session. If loading a session while in the middle of a session, the session to be loaded can be added to the current session.

- Save Session – saves the current FCSSim session to a text file with extension .SIM. When a session is saved, information for all of the currently generated FCS curves will saved. This saved session can be loaded using Load Session.
- Export All – will export all FCS curves. This is equivalent to exporting curves using the Export feature of the Curves consoles.
- Exit – closes FCSSim.

b. Options

- Plot Preferences
    - Plotting new simulation – select Separately will plot new curves in separate figure windows. To plot curves in the same window, select Together. For either case, the curves are plotted with randomly generated colors.
    - Show Legend – display figure legends.
    - Only Show Selected Data/Fits – if this option is selected only the curves that are selected (check box in respective Tables) will be displayed.
- Curve Properties – these values are used to determine lag time points. The calculation of lag time points is based on the Multi-Tau scheme (see ISS Vista manual and reference 65).
    - Multi Tau Base – base value.
    - Number of Sections – the quantity of sections that make up a correlation curve.

- Points per Section – the number of points in each section of a correlation curve. The total number of points is the product of points per section and number of sections.
- Sampling Time – correlation curves are calculated based on a specific sampling time value.

- Parameter Bounds – specify the lower and upper bounds for all parameters. When specifying parameter values for a given curve through the Curves console or the Controls console, the value must stay within the bounds specified here.

c. Window

- FCS Models – opens the FCS Models console.
- Model Controls – opens the Controls consoles corresponding to all of the currently active Curves consoles.

d. About – credits and current version of FCSSim.

2. FCS Models

a. Models list – select the FCS model to be used for fitting from this list.

b. Open – displays the Fit Console corresponding to the selected model.

Descriptions of the models can be found in chapter II as well as in the literature (references 27, 34, 37, 64, 65).

3. Curves Console

a. New – creates a new curve using default parameter values and the current Curve Properties settings for lag time points.

b. Reset – the parameters for the selected curves are set to the default values. Plots are updated.

c. Remove – removes the selected curves from the Curves console.

d. Export – parameters and correlation values for the selected curves are exported to a text file.

NOTE: closing a Curves console will remove the generated curves from FCSSim.

4. Controls Console
    a. Scroll Bars – use the scroll-bars to easily change a parameter's value. The parameter value in the table will be automatically updated. In addition, the plot for all of the selected curves will be updated.
    b. Text Boxes – in addition to the scroll-bars, a parameter's value can be changed using the text box. The corresponding scroll-bar as well as the parameter value in the table will be automatically updated. Plots for all of the selected curves will be updated as well.
5. Plot Figure
    a. File
        - Set Active – if selected, the current figure is the active one and new curves will be plotted in this figure.
        - Print – the figure window with the curves can be printed as displayed.
        - Export – the figure window can be exported to image files and as a PDF. The available image file formats are:
            - Enhanced meta file (*.emf)
            - Windows bitmap (*.bmp)
            - JPEG image (*.jpg)
            - Portable Network Graphics (*.png)
            - Tiff image (*.tif)
    b. Edit

- Curve Appearance – change the color, marker type and line style of curves in the figure window.
- Update Title – display a custom figure title.

NOTE: closing a figure will remove the generated curves from FCSSim.

III. BRIGHTNESS ANALYSIS STOOL

III.1 GETTING STARTED

1. Introduction

   Brightness Analysis Tool (BAT) is analysis software for photon count data within the scope of cumulant analysis (chapter II). The software was entirely written in object-oriented MATLAB without any external or third-party components or packages. BAT is capable of performing cumulant analysis through the calculation or fit approach (references 51 and 53). The latter allows for two component analysis. The data to be analyzed must be ISS histogram files.

2. The Work Area

   When BAT is started, two windows or consoles will be displayed. BAT (Fig. 1) is the main console of the application and contains menu items. The PCData console (Fig. 1) will be used to display the currently loaded photon count (histogram) data. In addition, certain actions on the loaded photon count data can be taken directly from this console by using the five buttons located at the top of the console.

III.2 DATA ANALYSIS

Cumulant analysis can be performed in two ways: calculation or fitting (references 51 and 53). These two approaches differ significantly in their technical detail and thus are treated separately.

The calculation approach is used for single molecular species while fitting allows the treatment of two molecular species. Different types of input files are required for the two approaches.

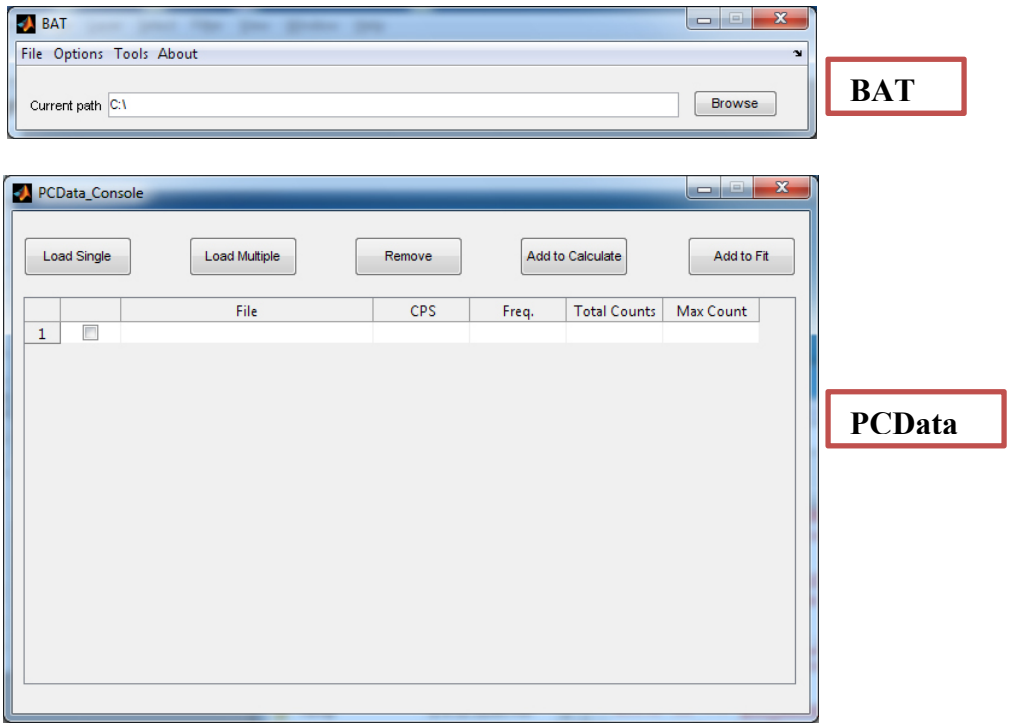


Figure 1. The Brightness Analysis Tool work area with the two primary consoles.

3. Data Analysis by Calculation

   a. Loading Data

      To load data for calculation, click the Load Single button on the PCData console (A in Fig. 2). Then select the data file to load. The file to be loaded must be in ISS format. Multiple data can be loaded at the same time. Once the data is loaded, it will be listed in the table (B in Fig. 2).

   b. Performing Calculation

      i. If the data to be used is not selected, select it by setting the check box in the first column of the table in PCData console (C in Fig. 2). Click on the Add

to Calculate button (D in Fig. 2) to add the selected data to the Calculate console (Fig. 3). The added data is displayed in the table in Calculate console.

ii. Corrections for detector dead-time and afterpulsing as well as finite sampling time will be calculated only if set in Correction Options found in the Options menu of the BAT console (see section III). The characteristic diffusion time and aspect ratio values given in columns 3 and 4 of the table (A in Fig. 3) will be used to calculate corrections for finite sampling time.

iii. Click on Calculate (B in Fig. 3) to perform calculation and obtain molecular brightness (lambda) and average population values. The calculated values will be displayed in the table in their respective columns.

d. Exporting

The calculation results can be exported by clicking on the Export button (C in Fig. 3) in the Calculate console. The results will be exported to a .CSV file.

e. Plotting

The calculation results can be plotted by clicking on the Plot button (D in Fig. 3). The figure will display the brightness and average population values for each data on a double y-axis plot with brightness values (lambda) on the left axis and population values (N) on the right axis (Fig. 4). The x-axis values correspond to the order of the data in the table of Calculate console.

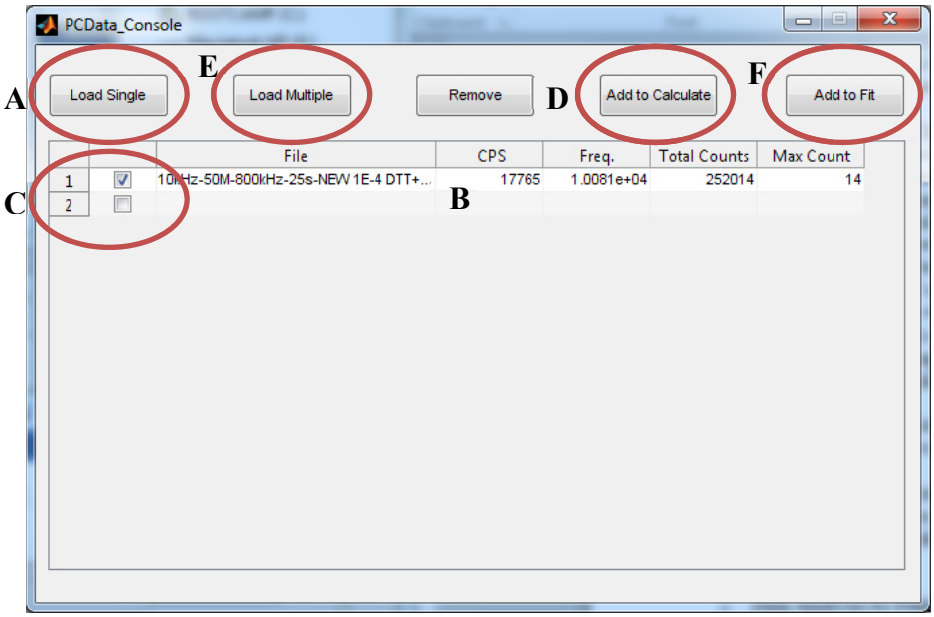


Figure 2. The PCData console is shown with one loaded data displayed in the table.

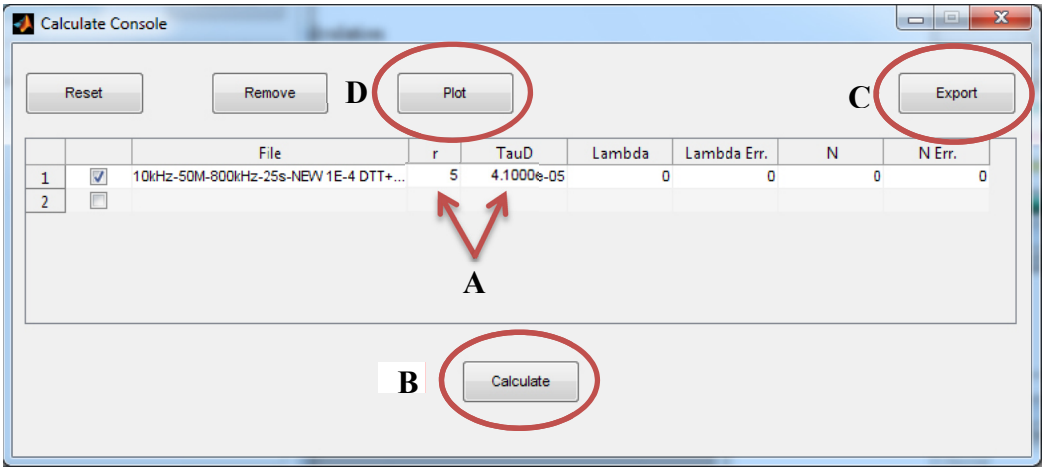


Figure 3. The Calculate console with data displayed in the table.

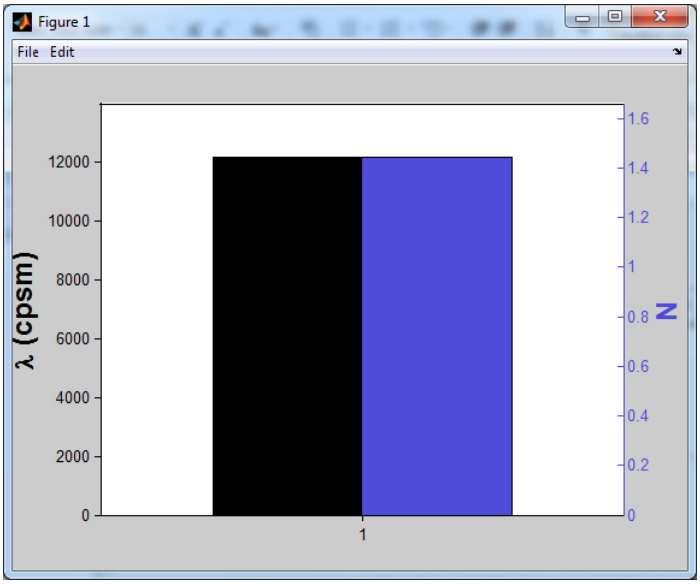


Figure 4. Calculated brightness (black) and average population (blue) values for data 1.

4. Data Analysis by Fitting

    a. Loading Data

    To load data for fitting, click the Load Multiple button (E in Fig. 2) on the PCData console. Then select the data files to load. For fitting cumulants, at least four sets of photon count data or histograms must be provided. These data can be in separate files in or in the same file. In either case, the data must have been saved in ISS format. Multiple data can be loaded at the same time. Once a data set is loaded, it will be listed in the table. The key word MULT will be displayed in the table columns to signify that multiple photon count data have been loaded.

    b. Performing Fit

        i. If the data to be used is not selected, select it by setting the check box in the first column of the table in PCData console. Click on the Add to Fit button

(F in Fig. 2) to add the selected data to the Fit console (Fig. 4). The added data is displayed in the table.

ii. Default values for the two brightness and average population values will be displayed in the columns of the table. These will be the initial values for the fit. If different initial values are desired for one or more of the parameters, they can be changed in the table.

iii. One or more of the parameters can be fixed for the fit by setting the check box (A in Fig. 5) found above the parameter columns. The value displayed in the table will be the fixed value for the parameter.

iv. Corrections for detector afterpulsing and finite sampling time will be calculated only if set in Correction Options found in the Options menu of the BAT console. The two characteristic diffusion time values for the two molecular species (TauD 1, TauD 2 in columns 4 and 5 respectively) and the aspect ratio values (r, in column 3) given in the table will be used to calculate corrections for finite sampling time.

v. Once the initial values, fixed parameters and correction options are set as desired, click on the Fit button (B in Fig. 5) at the bottom to fit the data. Only the data that are selected by setting the check box in the first column of the table will be fit.

vi. After the fit is complete, the final parameter and reduced chi square values will be displayed in the table.

c. Exporting

Fit results can be exported by clicking on the Export button (C in Fig. 5) in the Fit console. For a given fit, the final cumulant values and the fit parameters can be exported. The exported files are in a CSV format.

d. Plotting

The fit results can be plotted by clicking on the Plot button (D in Fig. 5). This will generate two figures for the two species (Fig. 6). Each figure will display the brightness and average population values for each data on a double y-axis plot with brightness values (lambda 1 and lambda 2) on the left axis and population values (N 1 and N 2) on the right axis. The x-axis values correspond to the order of the data in the table of Fit console.

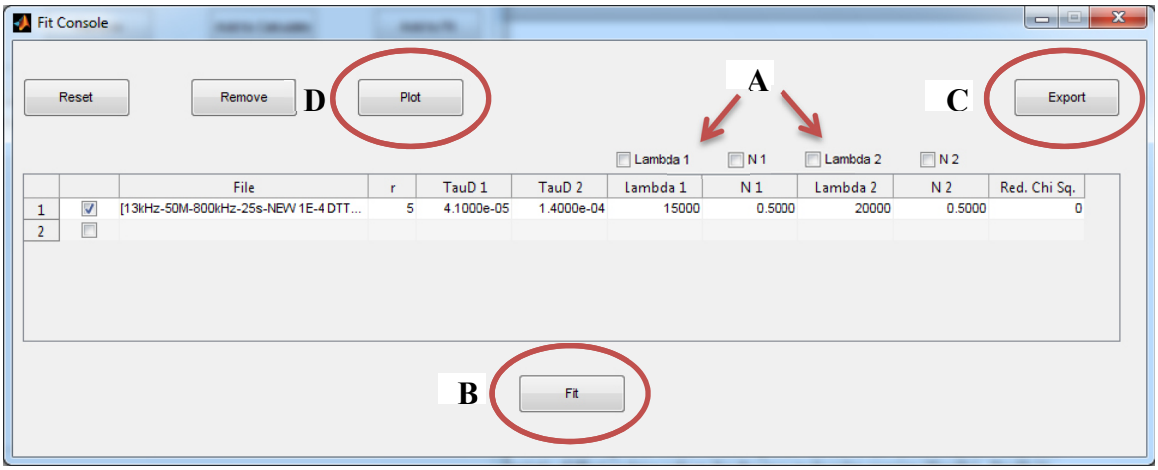


Figure 5. The Fit console is shown with a single data displayed in the table.

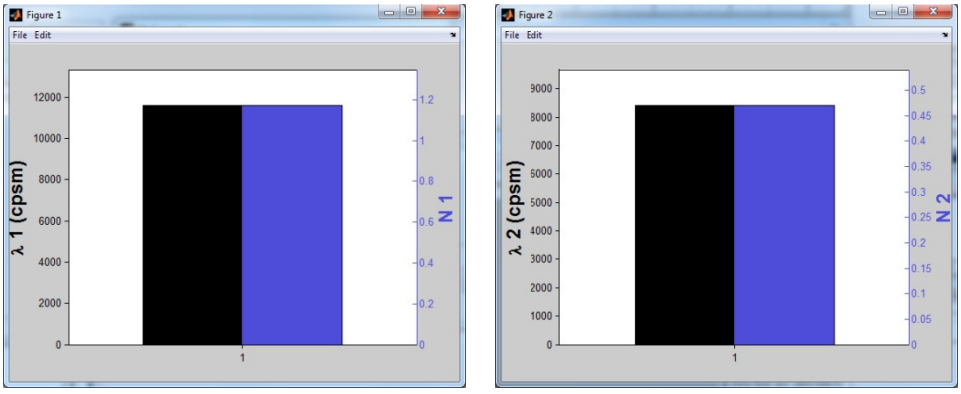


Figure 6. Brightness and average population values for the two molecular species obtained from the fit.

III.3 SOFTWARE COMPONENTS & FEATURES LIST

1. BAT

    a. File

        - Load Single – loads photon count (histogram) data for calculation. The data to be loaded must be in ISS histogram format.
        - Load Multiple – loads photon count (histogram) data for fitting. A minimum of four photon count data is required for fitting. The data can be in separate files or a single file and must have been saved in ISS histogram format.
        - Exit – closes BAT.

    b. Options

        - Corrections

          Correction values for detector dead-time and afterpulsing as well as finite sampling time can be calculated if selected (Fig. 7) (see references 53 and 54). For the calculation approach, all three can be applied while for the fitting approach afterpulsing and finite sampling time corrections can be applied. The detector dead-time and afterpulsing corrections are calculated based on previously determined values for the specific detectors (avalanche photodiodes). Finite sampling time (undersampling) corrections are determined based on the characteristic diffusion of the molecules (TauD) and the aspect ratio of the observation volume (r).

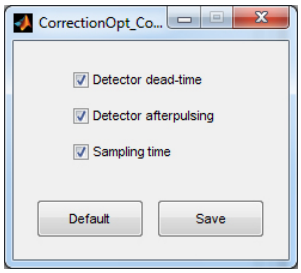


Figure 7. Correction options.

- Fit Preferences
    - Algorithm – the type of algorithm to be used for fitting can be set to either Trust-Region-Reflective or Levenberg-Marquardt.
    - Max Fun. Evals. – maximum number of function evaluations
    - Max Iteration – maximum number of fit iterations allowed. Fitting will not continue past the specified number of iterations.
    - Function Tol. – the amount by which the function value changes during the iteration. Fit will terminate if the change is smaller than the specified value.
    - Parameter Tol. – the amount by which the parameter values change during iteration. Fit will terminate if the change is smaller than the specified value.
- Parameter Bounds – specify the lower and upper bounds for all parameters. These values will be used only when the Trust-Region-Reflective algorithm is selected.

c. Tools

- Generate Histogram

  This feature allows the user to create histograms from raw photon data. The ISS Vista software outputs raw photon count data to a file which is then used

to obtain photon count data commonly referred to as histograms. The raw ISS file is saved in a proprietary binary format (see ISS Vista Manual) which is accessed by the Generate Histogram tool.

Selecting this tool in the menu opens the Generate Histogram console (Fig. 8). The raw ISS files can be selected by clicking on Select Files in the Generate Histogram console. If specific files are not provided, the tool will search for and process raw ISS files in the current path. To start generating histograms, click on the Start button at the bottom of the Generate Histogram console.

The generated histograms can be exported to a single file by setting the Export to Single File check box. Otherwise, each histogram will be saved to a separate file with the sampling frequency indicated in the file name. The export file name from the user, if provided, will be used. If an export file name is not given, the file name for the raw data will be used. In either case, the string – "-BAT" will be appended at the end of the file name.

Histograms are exported in ISS format.

NOTE: histograms can be generated by using ISS Vista. However, this software does not allow multiple histograms to be saved to the same file.

This type of file produced by Generate Histogram can be used as an input file for fit analysis by BAT.

d. About – credits and current version of BAT.

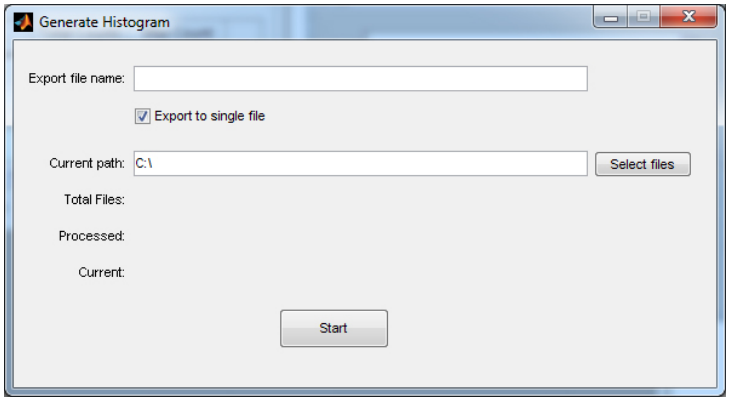


Figure 8. The Generate Histogram console.

2. PCData
    a. Load Single – loads photon count (histogram) data for calculation. The data to be loaded must be in ISS histogram format.
    b. Load Multiple – loads photon count (histogram) data for fitting. A minimum of four photon count data is required for fitting. The data can be in separate files or a single file and must have been saved in ISS histogram format.
    c. Remove – the selected photon count data are removed from BAT. If any of the data have been used to calculate or fit, the results will also be removed from the Calculate and Fit consoles.
    d. Add to Calculate – the selected photon count data are added to the Calculate console.
    e. Add to Fit – the selected photon count data are added to the Fit console.
3. Calculate Console
    a. Reset – resets the calculated brightness and average population values to the default values.
    b. Remove – removes the calculation results from Calculate console. If any of the results have been plotted, the plots will also be removed from the figures.

c. Plot – the brightness and average population values for the selected calculations will be plotted. Clicking plot opens a new figure window. A calculation result can be plotted multiple times.

d. Export – the calculated brightness and average population values can be exported to a text file in a CSV format. In addition, parameter error values and the relative error for the second cumulant will be exported. The exported file will be in a directory with the name BAT_EXP.

e. Calculate – determine brightness (lambda) and average population (N) values for the selected data. If the selected instances have previously been plotted, the plots will be updated with the new brightness and population values.

4. Fit Console

a. Reset – resets the brightness and average population as well as the characteristic diffusion time and aspect ratio values to the default values.

b. Remove – removes the fit results from Fit console. If any of the results have been plotted, the plots will also be removed from the figures.

c. Plot – the brightness and average population values for the selected fits will be plotted on two figures for the two molecular species. Clicking plot opens a new figure window. A fit result can be plotted multiple times.

d. Export – fit results can be exported to text files in a CSV format. The fit parameters and the corresponding cumulant values can be exported, each in a separate file. The exported files will be in a directory with the name BAT_EXP.

e. Fit – performs fit for the selected fit instances. The fit is performed using current parameter values as initial values for the free parameters and as fixed values for the fixed parameters. The current Correction (finite sampling time and afterpulsing) and Fit Options will be used.

5. Plot Figure

a. File

- Print – the figure window with the parameters can be printed as displayed.
- Export – the figure window can be exported to image files and as a PDF.

  The available image file formats are:

  - Enhanced meta file (*.emf)
  - Windows bitmap (*.bmp)
  - JPEG image (*.jpg)
  - Portable Network Graphics (*.png)
  - Tiff image (*.tif)

b. Edit

- Curve Appearance – change the color of the bar plots for molecular brightness and average population.
- Update Title – display a custom figure title.

REFERENCES

1. Tanford, C., and J. Reynolds. 2001. Nature's Robots A history of protiens. Oxford University Press, Oxford.

2. Stryer, L. 1995. Biochemistry. W.H. Freeman and Company, USA.

3. Berg, J. M., J. L. Tymockzo, and L. Stryer. 2002. Biochemistry. W. H. Freeman, New York.

4. Drenth, J. 2007. Principles of Protein X-Ray Crystallography. Springer, The Netherlands.

5. Petsko, G. A., and D. RInge. 2004. Protein Structure and Function. New Science Press Ltd, London.

6. Weber, G. 1975. Energetics of ligand binding to proteins. In Mechanisms and Pathways of Heterotrimeric G Protein Signaling. C. Anfinsen, J. Edsall, and F. Richards, editors. 1-83.

7. III, C. L. B., M. Karplus, and B. M. Pettitt. 1988. Proteins: A Theoretical Perspecitve of Dynamics, Structure, and Thermodynamics. John Wiley & Sons, New York.

8. McCammon, J. A., and S. C. Harvey. 1987. Dynamics of proteins and nucleic acids. Cambridge University Press, New York.

9. Henzler-Wildman, K., and D. Kern. 2007. Dynamic personalities of proteins. Nature 450:964-972.

10. Henzler-Wildman, K. A., M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern. 2007. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature 450:913-916.

11. Hilser, V. J. 2010. Biochemistry. An ensemble view of allostery. Science 327:653-654.

12. Koshland, D. E., Jr., G. Nemethy, and D. Filmer. 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5:365-385.

13. Monod, J., J. P. Changeux, and F. Jacob. 1963. Allosteric proteins and cellular control systems. Journal of molecular biology 6:306-329.

14. Bai, F., R. W. Branch, D. V. Nicolau, Jr., T. Pilizota, B. C. Steel, P. K. Maini, and R. M. Berry. 2010. Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch. Science 327:685-689.

15. Smock, R. G., and L. M. Gierasch. 2009. Sending signals dynamically. Science 324:198-203.

16. Swain, J. F., and L. M. Gierasch. 2006. The changing landscape of protein allostery. Current opinion in structural biology 16:102-108.

17. Careri, G., P. Fasella, and E. Gratton. 1979. Enzyme dynamics: the statistical physics approach. Annual review of biophysics and bioengineering 8:69-97.

18. Gerstein, M., A. M. Lesk, and C. Chothia. 1994. Structural mechanisms for domain movements in proteins. Biochemistry 33:6739-6749.

19. Doucet, N., E. D. Watt, and J. P. Loria. 2009. The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. Biochemistry 48:7160-7168.

20. Eisenmesser, E. Z., O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern. 2005. Intrinsic dynamics of an enzyme underlies catalysis. Nature 438:117-121.

21. Wolf-Watz, M., V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, and D. Kern. 2004. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nature structural & molecular biology 11:945-949.

22. Bhabha, G., J. Lee, D. C. Ekiert, J. Gam, I. A. Wilson, H. J. Dyson, S. J. Benkovic, and P. E. Wright. 2011. A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. Science 332:234-238.

23. Loveridge, E. J., L. H. Tey, E. M. Behiry, W. M. Dawson, R. M. Evans, S. B. Whittaker, U. L. Gunther, C. Williams, M. P. Crump, and R. K. Allemann. 2011. The role of large-scale motions in catalysis by dihydrofolate reductase. Journal of the American Chemical Society 133:20561-20570.

24. Daniel, R. M., R. V. Dunn, J. L. Finney, and J. C. Smith. 2003. The role of dynamics in enzyme activity. Annual review of biophysics and biomolecular structure 32:69-92.

25. Kitao, A. 2011. Transform and relax sampling for highly anisotropic systems: application to protein domain motion and folding. The Journal of chemical physics 135:045101.

26. Fitzpatrick, J. A., and B. F. Lillemeier. 2011. Fluorescence correlation spectroscopy: linking molecular dynamics to biological function in vitro and in situ. Current opinion in structural biology 21:650-660.

27. Krichevsky, O., and G. Bonnet. 2002. Fluorescence correlation spectroscopy: the technique and its applications. Rep Prog Phys 65:251-297.

28. Muller, J. D. 2004. Cumulant analysis in fluorescence fluctuation spectroscopy. Biophysical journal 86:3981-3992.

29. Magde, D., W. W. Webb, and E. Elson. 1972. Thermodynamic Fluctuations in a Reacting System - Measurement by Fluorescence Correlation Spectroscopy. Phys Rev Lett 29:705-&.

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

31. Van Orden, A., K. Fogarty, and J. Jung. 2004. Fluorescence fluctuation spectroscopy: a coming of age story. Appl Spectrosc 58:122A-137A.

32. Xu, C., and W. W. Webb. 1996. Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm. JOURNAL OF THE OPTICAL SOCIETY OF AMERICA B-OPTICAL PHYSICS 13:481-491.

33. Schwille, P., U. Haupts, S. Maiti, and W. W. Webb. 1999. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. Biophysical journal 77:2251-2265.

34. Nagy, A., J. Wu, and K. M. Berland. 2005. Observation volumes and {gamma}-factors in two-photon fluorescence fluctuation spectroscopy. Biophysical journal 89:2077-2090.

35. Amundson, N. R. 1966. Mathematical Methods in Chemical Engineering Matrices and Their Applications. Prentice-Hall, New Jersey.

36. Gurunathan, K., and M. Levitus. 2010. FRET fluctuation spectroscopy of diffusing biopolymers: contributions of conformational dynamics and translational diffusion. The journal of physical chemistry. B 114:980-986.

37. Orden, A. V., and J. Jung. 2008. Review fluorescence correlation spectroscopy for probing the kinetics and mechanisms of DNA hairpin formation. Biopolymers 89:1-16.

38. Chen, H. M., E. Rhoades, J. S. Butler, S. N. Loh, and W. W. Webb. 2007. Dynamics of equilibrium structural fluctuations of apomyoglobin measured by fluorescence correlation spectroscopy. P Natl Acad Sci USA 104:10459-10464.

39. Haupts, U., S. Maiti, P. Schwille, and W. W. Webb. 1998. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. P Natl Acad Sci USA 95:13573-13578.

40. Neuweiler, H., W. Banachewicz, and A. R. Fersht. 2010. Kinetics of chain motions within a protein-folding intermediate. P Natl Acad Sci USA 107:22106-22110.

41. Sherman, E., and G. Haran. 2011. Fluorescence Correlation Spectroscopy of Fast Chain Dynamics within Denatured Protein L. Chemphyschem 12:696-703.

42. Jung, J., and A. Van Orden. 2006. A three-state mechanism for DNA hairpin folding characterized by multiparameter fluorescence fluctuation spectroscopy. Journal of the American Chemical Society 128:1240-1249.

43. Hendrix, J., C. Flors, P. Dedecker, J. Hofkens, and Y. Engelborghs. 2008. Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. Biophysical journal 94:4103-4113.

44. Malvezzi-Campeggi, F., M. Jahnz, K. G. Heinze, P. Dittrich, and P. Schwille. 2001. Light-induced flickering of DsRed provides evidence for distinct and interconvertible fluorescent states. Biophysical journal 81:1776-1785.

45. Schwille, P., S. Kummer, A. A. Heikal, W. E. Moerner, and W. W. Webb. 2000. Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins. Proc Natl Acad Sci U S A 97:151-156.

46. Werner, J. H., R. Joggerst, R. B. Dyer, and P. M. Goodwin. 2006. A two-dimensional view of the folding energy landscape of cytochrome c. P Natl Acad Sci USA 103:11130-11135.

47. Qian, H., and E. L. Elson. 1990. On the analysis of high order moments of fluorescence fluctuations. Biophysical journal 57:375-380.

48. Chen, Y., J. D. Muller, P. T. So, and E. Gratton. 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. Biophysical journal 77:553-567.

49. Kask, P., K. Palo, D. Ullmann, and K. Gall. 1999. Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. Proc Natl Acad Sci U S A 96:13756-13761.

50. Bedard, G. 1967. DEAD-TIME CORRECTIONS TO STATISTICAL DISTRIBUTION OF PHOTOELECTRONS. PROCEEDINGS OF THE PHYSICAL SOCIETY OF LONDON 90.

51. Muller, J. D., Y. Chen, and E. Gratton. 2000. Resolving heterogeneity on the single molecular level with the photon-counting histogram. Biophysical journal 78:474-486.

52. Saleh, B. 1978. Photoelectron Statistics with Applications to Spectroscopy and Optical Communication. Springer-Verlag, Berlin, New York.

53. Wu, B., and J. D. Muller. 2005. Time-integrated fluorescence cumulant analysis in fluorescence fluctuation spectroscopy. Biophysical journal 89:2721-2735.

54. Hillesheim, L. N., and J. D. Muller. 2003. The photon counting histogram in fluorescence fluctuation spectroscopy with non-ideal photodetectors. Biophysical journal 85:1948-1958.

55. Hobel, M., and J. Ricka. 1994. DEAD-TIME AND AFTERPULSING CORRECTION IN MULTIPHOTON TIMING WITH NONIDEAL DETECTORS. REVIEW OF SCIENTIFIC INSTRUMENT 65:2326-2336.

56. Finn, M., G. Greenless, and H. TW. 1988. REAL-TIME ELIMINATION OF DEAD TIME AND AFTERPULSING IN COUNTING SYSTEMS. REVIEW OF SCIENTIFIC INSTRUMENTS 59:2457-2459.

57. Johnson, F., R. Jones, M. TP, and P. ER. 1966. DEAD-TIME CORRECTIONS TO PHOTON COUNTING DISTRIBUTIONS. Phys Rev Lett 16:589.

58. O'Donnell, K. 1986. CORRECTION OF DEAD-TIME EFFECTS IN PHOTOELECTRIC-COUNTING DISTRIBUTIONS. JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A-OPTICS IMAGE SCIENCE AND VISION 3:113-115.

59. Campbell, L. 1992. AFTERPULSE MEASUREMENT AND CORRECTION. REVIEW OF SCIENTIFIC INSTRUMENTS 63:5749.

60. Caccia, M., E. Camozzi, M. Collini, and C. Giuzeppe. 2005. Photon moment analysis in cells in the presence of photo-bleaching. APPLIED SPECTROSCOPY 59:227-236.

61. Eid, J., M. JD, and G. E. 2000. Data acquisition card for fluctuation correlation spectroscopy allowing full access to the detected photon sequence. REVIEW OF SCIENTIFIC INSTRUMENTS 71:361-368.

62. Pabit, S., and K. Berland. 2008. Two-Photon Fluorescence Correlation Spectroscopy. In Handbook of BIomedical Nonlinear Optical Microscopy. B. R. Masters, and P. T. C. So, editors. Oxford University Press, Oxford. 644-672.

63. Berland, K. M. 2004. Detection of specific DNA sequences using dual-color two-photon fluorescence correlation spectroscopy. Journal of biotechnology 108:127-136.

64. Meseth, U., T. Wohland, R. Rigler, and H. Vogel. 1999. Resolution of fluorescence correlation measurements. Biophysical journal 76:1619-1631.

65. Wohland, T., R. Rigler, and H. Vogel. 2001. The standard deviation in fluorescence correlation spectroscopy. Biophysical journal 80:2987-2999.

66. Faber, H. R., and B. W. Matthews. 1990. A mutant T4 lysozyme displays five different crystal conformations. Nature 348:263-266.

67. Dixon, M. M., H. Nicholson, L. Shewchuk, W. A. Baase, and B. W. Matthews. 1992. Structure of a hinge-bending bacteriophage T4 lysozyme mutant, Ile3-->Pro. Journal of molecular biology 227:917-933.

68. Zhang, X. J., J. A. Wozniak, and B. W. Matthews. 1995. Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. Journal of molecular biology 250:527-552.

69. Mchaourab, H. S., K. J. Oh, C. J. Fang, and W. L. Hubbell. 1997. Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling. Biochemistry 36:307-316.

70. Kuroki, R., L. H. Weaver, and B. W. Matthews. 1993. A Covalent Enzyme-Substrate Intermediate with Saccharide Distortion in a Mutant T4 Lysozyme. Science 262:2030-2033.

71. Goto, N. K., N. R. Skrynnikov, F. W. Dahlquist, and L. E. Kay. 2001. What is the average conformation of bacteriophage T4 lysozyme in solution? A domain orientation study using dipolar couplings measured by solution NMR. Journal of molecular biology 308:745-764.

72. Chen, Y., D. H. Hu, E. R. Vorpagel, and H. P. Lu. 2003. Probing single-molecule T4 lysozyme conformational dynamics by intramolecular fluorescence energy transfer. Journal of Physical Chemistry B 107:7947-7956.

73. Wang, Y., and H. P. Lu. 2010. Bunching effect in single-molecule T4 lysozyme nonequilibrium conformational dynamics under enzymatic reactions. The journal of physical chemistry. B 114:6669-6674.

74. de Groot, B. L., S. Hayward, D. M. van Aalten, A. Amadei, and H. J. Berendsen. 1998. Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. Proteins 31:116-127.

75. Lieto, A. M., R. C. Cush, and N. L. Thompson. 2003. Ligand-receptor kinetics measured by total internal reflection with fluorescence correlation spectroscopy. Biophysical journal 85:3294-3302.

76. Schwille, P., F. Oehlenschlager, and N. G. Walter. 1996. Quantitative hybridization kinetics of DNA probes to RNA in solution followed by diffusional fluorescence correlation analysis. Biochemistry 35:10182-10193.

77. Chattopadhyay, K., E. L. Elson, and C. Frieden. 2005. The kinetics of conformational fluctuations in an unfolded protein measured by fluorescence methods. P Natl Acad Sci USA 102:2385-2389.

78. Chattopadhyay, K., S. Saffarian, E. L. Elson, and C. Frieden. 2002. Measurement of microsecond dynamic motion in the intestinal fatty acid binding protein by using fluorescence correlation spectroscopy. Proc Natl Acad Sci U S A 99:14171-14176.

79. Price, E. S., M. Aleksiejew, and C. K. Johnson. 2011. FRET-FCS detection of intralobe dynamics in calmodulin. The journal of physical chemistry. B 115:9320-9326.

80. Price, E. S., M. S. DeVore, and C. K. Johnson. 2010. Detecting intramolecular dynamics and multiple Forster resonance energy transfer states by fluorescence correlation spectroscopy. The journal of physical chemistry. B 114:5895-5902.

81. Mchaourab, H. S., M. A. Lietzow, K. Hideg, and W. L. Hubbell. 1996. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. Biochemistry 35:7692-7704.

82. Arbeloa, F. L., P. R. Ojeda, and I. L. Arbeloa. 1989. Fluorescence Self-Quenching of the Molecular-Forms of Rhodamine-B in Aqueous and Ethanolic Solutions. J Lumin 44:105-112.

83. Arbeloa, I. L., and P. R. Ojeda. 1982. Dimeric States of Rhodamine-B. Chem Phys Lett 87:556-560.

84. Ilich, P., P. K. Mishra, S. Macura, and T. P. Burghardt. 1996. Direct observation of rhodamine dimer structures in water. Spectrochim Acta A 52:1323-1330.

85. Packard, B. Z., A. Komoriya, D. D. Toptygin, and L. Brand. 1997. Structural characteristics of fluorophores that form intramolecular H-type dimers in a protease substrate. Journal of Physical Chemistry B 101:5070-5074.

86. Christie, R. J., C. J. Tadiello, L. M. Chamberlain, and D. W. Grainger. 2009. Optical Properties and Application of a Reactive and Bioreducible Thiol-Containing Tetramethylrhodamine Dimer. Bioconjugate Chem 20:476-480.

87. Packard, B. Z., D. D. Toptygin, A. Komoriya, and L. Brand. 1996. Profluorescent protease substrates: Intramolecular dimers described by the exciton model. P Natl Acad Sci USA 93:11640-11645.

88. Widengren, J., U. Mets, and R. Rigler. 1995. Fluorescence Correlation Spectroscopy of Triplet-States in Solution - a Theoretical and Experimental-Study. J Phys Chem-Us 99:13368-13379.

89. Beechem, J. M., E. Gratton, M. Ameloot, J. R. Knutson, and L. Brand. 2002. The global analysis of fluorescence intensity and anisotropy decay data: second-generation theory and programs. In Topics in fluorescence spectroscopy. J. R. Lakowicz, editor. 241-305.

90. Choi, Y., I. S. Moody, P. C. Sims, S. R. Hunt, B. L. Corso, I. Perez, G. A. Weiss, and P. G. Collins. 2012. Single-molecule lysozyme dynamics monitored by an electronic circuit. Science 335:319-324.

91. Matthews, B. W., and S. J. Remington. 1974. The three dimensional structure of the lysozyme from bacteriophage T4. Proc Natl Acad Sci U S A 71:4178-4182.

92. Hayward, S. 1999. Structural principles governing domain motions in proteins. Proteins 36:425-435.

93. Hoffman, L., R. A. Stein, R. J. Colbran, and H. S. McHaourab. 2011. Conformational changes underlying calcium/calmodulin-dependent protein kinase II activation. The EMBO journal 30:1251-1262.

94. Kolodziej, S. J., A. Hudmon, M. N. Waxham, and J. K. Stoops. 2000. Three-dimensional reconstructions of calcium/calmodulin-dependent (CaM) kinase IIalpha and truncated CaM kinase IIalpha reveal a unique organization for its structural core and functional domains. The Journal of biological chemistry 275:14354-14359.

95. Jones, P. M., M. L. O'Mara, and A. M. George. 2009. ABC transporters: a riddle wrapped in a mystery inside an enigma. Trends in biochemical sciences 34:520-531.

96. Mchaourab, H. S., P. R. Steed, and K. Kazmier. 2011. Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy. Structure 19:1549-1561.

97. Dong, J., G. Yang, and H. S. McHaourab. 2005. Structural basis of energy transduction in the transport cycle of MsbA. Science 308:1023-1028.

98. Boukobza, E., A. Sonnenfeld, and G. Haran. 2001. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. JOURNAL OF PHYSICAL CHEMISTRY B 105:12165-12170.

99. Palmer, A. G., and N. L. Thompson. 1987. MOLECULAR AGGREGATION CHARACTERIZED BY HIGH-ORDER AUTO-CORRELATION IN FLUORESCENCE CORRELATION SPECTROSCOPY. Biophysical journal 52:257-270.

100. Rigler, R., A. Pramanik, P. Jonasson, G. Kratz, O. T. Jansson, P. Nygren, S. Stahl, K. Ekberg, B. Johansson, S. Uhlen, M. Uhlen, H. Jornvall, and J. Wahren. 1999. Specific binding of proinsulin C-peptide to human cell membranes. Proc Natl Acad Sci U S A 96:13318-13323.