Protein Diffusion in *Escherichia coli*

Kristin M. Slade

A dissertation submitted to the faculty of the University of North Carolina at
Chapel Hill in partial fulfillment of the requirements for the degree of Doctorate of
Philosophy in the Department of Chemistry

Chapel Hill
2009

Approved by:

Gary Pielak, Ph.D.

Nancy Thompson, Ph.D.

Michael Chua, Ph.D.

Mark Wightman, Ph.D.

Royce Murray, Ph.D.

Abstract

Kristin M. Slade: Protein Diffusion in *Escherichia coli*
(Under the direction of Professor Gary J. Pielak, Ph.D.)

This dissertation describes the creation of a system to provide insight about the affects of protein expression on intracellular diffusion. Fluorescence recovery after photobleaching (FRAP) is used to obtain diffusion coefficients. Nuclear magnetic resonance spectroscopy detection of ^{19}F isotopic-enriched proteins and fluorescent labeling with FlAsH and were attemped before settling on FRAP studies with green fluorecent protein (GFP). To regulate protein levels in *Escherichia coli*, several vector and promoter combinations were tested. Eventually, a single vector was created containing the structural gene for GFP under the *lac* promoter and a test protein under the *araBAD* promoter. With this vector, the test protein was expressed at varying levels and GFP was expressed at a constant level. Although chymotrypsin inhibitor 2 was originally chosen as the test protein, it was quickly replaced by α-synuclein, maltose binding protein, tau-40, and calmodulin. My most important result is that regardless of the type or amount of protein that was co-expressed, the GFP diffusion coefficient remained constant. We conclude that expression of these soluble proteins has little to no effect on the diffusion of GFP.

Several disadvantages of FRAP became apparent in the process of obtaining these data. FRAP and the other common techniques for measuring

translational diffusion, present numerous obstacles when working with structures

that are only slightly larger than optical resolution, such as *Escherichia coli* cells.

To overcome these obstacles, I also developed a new method to assess diffusion

using through-prism total internal reflection fluorescence microscopy with

continuous photobleaching. Here, the theoretical basis for this technique is

presented. I demonstrate its applicability by measuring the diffusion coefficient,

6.3 ± 1.1 µm^2/sec, of GFP in *Escherichia coli* cells.

Dedication:

To future graduate students…
you have no idea what lies ahead!

Acknowledgements

So many individuals have helped me get where I am today that my acknowledgment section could be a dissertation in itself. Thanks to all of you. I am privileged to have been influenced by such positive and motivating friends and mentors. Thanks to my teachers and professors along the way who nurtured my interest in science and encouraged my curiosity. To Raymond Dominey, my undergraduate advisor: thanks for believing in me, even when I did not. Our conversations during those early morning walks have influenced me more than you know. Gary Pielak, my advisor and mentor, thanks for always keeping me on my toes and challenging me to be a better scientist. You never failed to surprise me with your quirky comments and pranks. Thanks for constantly reminding me to appreciate the small things, even a cold rain during a bike ride to work. I don't know how many times over the past five years, your positive words of encouragement have helped me continue when my results looked bleak. To Nancy Thompson, my co-advisor and a brilliant scientist, thanks for your guidance and expertise, as well as your encouragement and moral support. I also thank the rest of my committee, Drs. Michael Chua, Royce Murray, and Mark Wightman, for their time and support.

To the Pielak and Thompson lab members, thanks for so many memories.

Andrew Miklos, I appreciate that you were always willing to bail me out of computer trouble. Lisa Charlton, thanks for your extensive advice on anything and everything I asked you about from protein expression and cloning to the stock market and cooking tips. Bridgette Steele, I'll never forget our afternoons of conversations in the dark. Jamie Pero, thanks for all of your moral support and encouragement, as well to bringing fun to any situation. Most of my craziest graduate school memories involve you. After all, who else would bring a disco ball to lab and dance on the bench. Becky Ruf, a great friend and colleague, I don't even know where to start! Thanks for a wonderfully organized lab and for keeping my fashion *faux pas* in check. On a more serious note, I appreciate that I can always rely on you in a pinch, both in and out of lab. To my undergraduate assistants: Sandy An, you never failed to cheer me up when our science wasn't working (thanks for introducing me to the lucky sombrero); Heidi Scronce, your warmth and sincerity has been touching; Naima Sharaf, your enthusiasm has been inspiring; Hao Wu, thanks for all the bizarre stories I'll have to tell about graduate school.

To Michael Chua, Neil Kramarcy, and Wendy Salmon, thanks not only for your help with the instrument, but also for the great laughs and wonderful chocolate. Michael, I am especially grateful for your tiffsplit program. You were always quick to help me fix any glitches and but slow to frustrate. I appreciate your patience with teaching me so many basic computer tricks.

Most importantly, I want to thank my family for their support and encouragement all of these years. Mom, I especially enjoyed all the 6 AM phone

conversations on the way into work. Thanks for always demanding my best. To Dave, my husband, I appreciate all the hours you spent listening to my verbose, detailed accounts of my day-to-day lab encounters. Thanks for being my rock and my strength through the ups and downs of graduate school.

Table of Contents

**Chapter 1. Reasons and Methods for Studying Protein
Diffusion in *Escherichia coli***

**Chapter 2. Attempts to Control the Expression of a Protein
and Measure its Diffusion Coefficient**

Chapter 3. Optimizing Expression with a Two-Protein System

Chapter 4. Effects of Recombinant Protein Expression on Green Fluorescent Protein Diffusion in *Escherichia coli*

Chapter 5. Quantifying GFP Diffusion in *Escherichia coli* by Using Continuous Photobleaching with Evanescent Illumination

List of Tables

Chapter 1

None

Chapter 2

Chapter 3

Chapter 4

Chapter 5

None

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviation and Symbols

amp	ampicillin
BAL	British Anti-Lewisite (2,3 dimercaptopropanol)
bp	base pairs
cAMP	cyclic adenosine monophosphate
C	concentration
^{o}C	degrees Celsius
CAP	cyclic AMP receptor protein
CD	circular dichroism
CI2	chymotrypsin inhibitor 2
CP	continuous photobleaching
d	evanescent wave decay
D	diffusion coefficient
D_2O	deuterium oxide
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
EDT	1,2-ethanedithiol
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
FCS	fluorescence correlation spectroscopy
FlAsH	fluorscein arsenical helix dye
FPLC	fast protein liquid chromatography
FRAP	fluorescence recovery after photobleaching

g	standard gravity
GFP	green fluorescent protein
h	hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
I_o	Initial intensity
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	Luria Broth
kan	kanamycin
kDa	kilodalton
MBP	maltose binding protein
mg	milligram
min	minute
mL	milliliter
mM	millimolar
mOsm	milliosmolality
MWCO	molecular weight cutoff
NEB	New England Biolabs
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
ori	origin of replication
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethanesulphonylfluoride

ppm	parts per million
PVP	poly(vinyl pyrrolidone)
rpm	revolutions per minute
RT	room temperature
SDS-PAG	sodium dodecyl sulfate polyacrylamide gel
sec	seconds
t	time
tfmF	L-4-trifluoromethylphenylalanine
TIR	total internal reflection
Tris	trishydrixynethylaminomethane
v/v	volume/volume
V	volts
w/v	weight/volume
z	distance from the interface

Greek-Based Symbols

κ	photobleaching propensity
τ	characteristic recovery time
μg	microgram
μL	microliter
μm	micrometer
μM	micromolar
μW	microwatt

Note:

The material in Chapter 4 is from:
Slade, KM, Baker, R, Chua, M, Thompson, NL, Pielak, GJ. 2009. Effects of Recombinant Protein Expression on Green Fluorescent Protein Diffusion in *Escherichia coli*. *Biochemistry*. In press.

The material in Chapter 5 is from:
Slade, KM, Steele, BL, Pielak, GJ, Thompson, NL. 2009. Quantifying GFP diffusion in *Escherichia coli* by using continuous photobleaching with evanescent illumination. *Journal of Physical Chemistry B*, 113(14), 4837-4845.

Chapter 1: Reasons and Methods for Studying Protein Diffusion in the Cytoplasm of *Escherichia. Coli*

1.1 Macromolecular Crowding

The interior of cells consists of a heterogeneous mixture of macromolecules that are tens to hundreds of times more concentrated than the dilute conditions used for most biophysical studies.[1] In fact, macromolecules account for more than 30% of the cell's volume.[2] Since two separate structures cannot occupy the same region of space, a soluble macromolecule will exclude volume to other macromolecules in the same solution. This steric exclusion of volume, which becomes more prevalent at the high concentrations encountered within cells, changes thermodynamic activities of molecules, alters protein chemistry, and consequently has significant ramifications for cellular function.[3]

Macromolecular crowding affects enzyme kinetics, protein folding and aggregation, diffusion, cell signaling, protein stability and a range of other conditions within the cell.[4,5] For instance, addition of a crowding agent accelerates aggregation of both apolipoprotein CII and α-synuclein.[6,7] Other studies have revealed that crowding accelerates binding and promotes macromolecular complex formation.[7-10] Furthermore, the thermal stability of several proteins increases when the proteins are confined to polyacrylamide gels.[11]   There is even evidence that crowding may cause the cytoplasm to

organize into microenvironments for substrate channeling.[12-14]

Predicting the effects of crowding are complicated and require the consideration of numerous factors, including the size and shape of both the crowding agent and the protein of interest. Natively-disordered proteins, for example, can gain structure inside cells, but other proteins have been shown to lose their structure in cells compared to dilute solution.[15,16] Furthermore, crowding elevates enzymatic activity of several proteins, but decreases the activity of others.[17-21] Although the rate of diffusion-limited reactions is slowed by high macromolecular concentrations, this condition serves to increase the rate of transition-state limited reactions.[22,23]

Most of the aforementioned evidence was obtained with theoretical modeling or *in vitro* experiments. Simple statistical-thermodynamic models, such as scaled particle theory and excluded volume theory, have been used to predict the qualitative effects of crowding. For *in vitro* experiments, inert macromolecules, such as dextrans, Ficolls, poly(ethylene glycol) (PEG), or poly(vinylpyrrolidone) (PVP) often serve as crowding agents, yet biological systems are much more complex.[24,25] Both theoretical and *in vitro* studies often fail to account for the heterogeneous environment and non-specific interactions encountered within cells. Consequently, the extent to which crowding affects the cellular environment still is controversial and not well understood. To gain further insight, experiments must be conducted on protein in their native environment: the cell. Such experiments will require further development of bioanalytical techniques.

1.2 In-cell NMR

Currently, in-cell NMR is one of the only non-invasive techniques with atomic-resolution that is capable of monitoring the complex environment inside living cells. The technique was first developed by Brindle and colleagues when they incorporated ^{19}F into several glycolytic enzymes using fluorinated amino acids and studied ligand binding, protein mobility and protein-protein interactions in *Saccharomyces cervisiae* with ^{19}F NMR.[26-28] Several years later, Serber and Dötsch pioneered a way to obtain high-resolution heteronuclear single quantum coherence (HSQC) spectra of protein in living *E. coli* cells.[29] To conduct an in-cell NMR experiment, the protein of interest is enriched with an NMR-active isotope, such as ^{15}N or ^{13}C, to differentiate it from other cellular components. The Dötsch group grew *E. coli* cells in rich media, but transferred the cells to media containing ^{15}N as the sole nitrogen source just before induction, so that the backbone and side chain amide nitrogens were uniformly enriched in ^{15}N.[30] Following this breakthrough, other groups have further developed in-cell NMR to provide insight on protein structure, stability, and dynamics.

Obstacles arise because NMR is insensitive.[31] Sensitivity can be increased with longer time periods for data collection and higher cells densities. These options, however, deprive cells of essential oxygen and nutrients which may ultimately decrease cell viability. Consequently, to obtain high-quality spectra for in-cell NMR, the protein of interest must be expressed at higher concentrations than other cellular components, which raises concerns about the relevancy of the observations from in-cell NMR. One might expect protein

3

production to cause further crowding, yet Dedmon *et al.* showed that induced and uninduced *E. coli* cells have the same amount of total protein.[15] This observation suggests that cells compensate for overexpression by decreasing the level of other components. Methods for increasing protein expression are known, but it is not known how overexpression affects cellular dynamics.[32]

1.3 Diffusion

Intracellular protein diffusion should provide a way to quantify the effects of protein overexpression on the cellular environment. For simple diffusion, the Stokes-Einstein relationship shows that diffusion is inversely related to viscosity, which in turn depends on the concentration of macromolecules in cells.[33] Thus, crowding should slow diffusion. Yet, the Stokes-Einstein equation assumes that the solute is diffusing through a continuous, hydrodynamic fluid with a constant viscosity, which is not true in the complex, heterogeneous environment of the cell.[34] The large size of intracellular macromolecules compared to the negligible-sized solvents of dilute solution often results in a range of micro viscosities in the cytoplasm which can affect diffusion.[5,35] Diffusion that does not follow the Einstein's equations is referred to as anomalous diffusion.[33] In anomalous diffusion, the mean squared displacement is not linearly related to time. Such diffusion is suspected to result from macromolecular crowding and protein-protein interactions. Although there is evidence of anomalous diffusion in both eukaryotes and bacteria, the extent to which intracellular diffusion follows Brownian motion or is anomalous has been heavily debated.[36] Furthermore,

insight about intracellular diffusion is important in its own right because diffusion is crucial for metabolism, gene transcription, protein assembly, macromolecular interactions, signaling, and other regulatory functions.[33] For bacteria, Brownian motion often serves as the primary source of intracellular movement, because these cells lack the highly developed cytoskeleton and motor proteins of eukaryotic organisms.[37]

1.4 Fluorescence Recovery After Photobleaching

Almost all intracellular diffusion studies rely on fluorescence spectroscopy. The most common method is fluorescence recovery after photobleaching (FRAP), in which a small region of a fluorescent sample is irreversibly bleached with a laser (Fig. 1.1).[38] The fluorescent recovery is then monitored over time as the unbleached fluorescent markers from the surrounding regions diffuse into the bleached area and the bleached markers diffuse into the surrounding regions. If the markers are freely diffusing, a recovery curve will be observed when the fluorescence intensity of the bleached region is plotted as a function of time. For ideal systems, the diffusion coefficient, D, is proportional to L^2/τ, where L is the size of the photobleached area and τ is the characteristic time for recovery.[39] This analysis assumes that there is an infinite supply of fluorescent markers to diffuse into the bleached region. Since bacteria cells are small (only about a tenth of the size of eukaryotic cells), this assumption is not valid.[37] Instead, data from FRAP studies in *E. coli* can be analyzed as described in Chapter 4. Briefly, the general solution to Fick's 2nd law and the appropriate boundary conditions is

a Fourier series, in which the amplitude is related to the diffusion coefficient.

For FRAP to be applicable, the protein of interest must be fluorescent and the fluorophore must be able to be irreversibly photobleached. Fluorescent recovery of individual fluorophores would contribute to τ, such that D would no longer be proportional to L^2/τ. Thus reversible photobleaching would complicate the data analysis. The fluorophore must also be stable enough that it is not bleached by the low-intensity observation beam.

1.5 Green Fluorescent Protein

Until recently, it was difficult to introduce a fluorescent marker that fulfilled these requirements into cells. This obstacle was overcome with the discovery of the 238-residue (27 kDa) green fluorescent protein (GFP) from the jellyfish, *Aequorea Victoria,* because GFP can be endogeneously expressed in cells.[40] GFP consists of an 11-stranded β-barrel that protects a central helix containing its chromophore. Autocatalytic cyclization reactions involving Ser[65], Gly[67], and Tyr[66] generate the fluorophore.[41] Furthermore, GFP is stable, maintaining its fluorescence even upon exposure to denaturing agents such as 1% sodium dodecyl sulfate.[42] GFP has become such an important tool in chemistry and biology that Osamu Shimomura, Martin Chalfie, and Roger Tsien were awarded the 2008 Nobel Prize in Chemistry for its discovery and its optimization. In one of the most important modifications, Tsien introduced a single point mutation, S65T, that increased the photostability and fluorescence of the protein while shifting the maximum excitation to a wavelength that is easily accessible with a typical Argon laser (488 nm).[43] The system was

further developed with the mutation F64L, which improves folding at 37ºC. This

popular form was named enhanced green fluorescent protein (EGFP) due to its

brighter fluorescence and better expression.[44] Like wildtype GFP, however, EGFP

tends to dimerize. Here, we use a non-dimerizable form of GFP developed from

Aequorea coerulescens called AcGFP1. AcGFP1 possesses an excitation

maximum at 475 nm and an emission maximum at 505 nm. While the primary

structure of AcGFP1 is 94% identical to EGFP, several of its surface hydrophobic

residues have been modified such that the protein is monomeric.

1.6 Promoters

A promoter is the region of DNA that regulates transcription of a particular

gene. Since a major goal of this project involves controlling protein expression, an

understanding of the general mechanism for several common promoters will be

useful. The pAcGFP1 vector contains the GFP gene under the control of the *lac*

promoter (Fig. 1.2). This sytem is controlled by both postive and negative

regulation.[45] In the absence of lactose, a repressor protein binds the lac operator.

As a result, RNA polymerase is blocked from binding to the DNA and transcription

is repressed. Lactose decreases the respressor's affinity for the lac operator.

Consequently, RNA polyerase is free to bind the promoter and turn on

transcription. Isopropyl β-D-1-thiogalactopyranoside (IPTG), an analog of lactose,

is usually used to induce expression in this system because, unlike lactose, it is not

metabolized by the bacteria.

A second factor that regulates transcription is the level of cyclic AMP

(cAMP) and cyclic AMP receptor protein (CAP) in the cell. CAP and cAMP form a complex that stimulates transcription by binding upstream of the lac promoter. The carbon source plays a major role in the level of cAMP. Glucose, an easily metabolized monosacchride, results in low cAMP levels. At low cAMP levels, CAP cannot effective bind the DNA, decreasing transcriptional efficiency. Thus, glucose repress transcription. In contrast, glycerol stimulates transcription by elevating cAMP levels.

The *T7* promoter provides an extra level of control to the *lac* system (Fig. 1.3).[46] *E. coli* RNA polymerase does not bind the *T7* promoter. Instead, transcription is initiated when T7 RNA polymerase, which is not native to *E. coli*, binds to the *T7* promoter. Specific strains, such as BL21(DE3) GOLD cells, are genetically engineered to contain a *T7* RNA polymerase gene under the control of a *lac* promoter. Thus the addition of IPTG turns on transcription of both T7 polymerase and the target gene.

Mutations in the *lac* and *T7* promoters can partially inhibit protein expression by decreasing their binding affinities for the RNA polymerase. The promoters contain two conserved sequences upstream of the transcription initiation site: TTGACA sequence (-35) and TATAAT (-10).[47] Specific modifications to these regions, or to the number of base pairs between these regions, cripple the promoters to varying degrees.[48,49]

The *araBAD* promoter allows control over the level of protein expression without the need for promoter mutations (Fig. 1.4).[50] Specifically, the level of the inducer, arabinose, that is added stipulates the amount of protein produced.

Arabinose binds to araC, releasing this repressor protein from the araBAD operon, turning on expression. CAP and cAMP are also involved in turning on transcription in this system. Glucose lowers cAMP levels, decreasing the binding affinity of CAP, and consequently inhibiting protein expression.

The T7 promoter can also be induced with arabinose in the BL21-AI strain. BL21-AI cells were designed to provide tight regulation for the expression of toxic proteins. For this strain, the T7 RNA polymerase gene is under the control of the *araBAD* promoter rather than the *lac* promoter. Consequently, protein expression is induced by arabinose rather than IPTG.

Armed with the above information, I have attempted to study how protein expression affects intracellular diffusion. Such information is important for understanding cellular metabolism, but also for understanding the results from in-cell NMR experiments. In this dissertation I describe the experiments used to gain further insight about protein expression and their results. Specifically, Chapter 2 describes strategies for controlling the level of protein expression while measuring the diffusion coefficient of that protein in *E. coli* cells. Chapter 3 focuses on the co-expression of two proteins. One protein (GFP) is expressed at a constant level and serves as the tracer molecule for diffusion measurements, while the expression level of the second is altered. This dual protein system is then used to conduct the FRAP experiments described in Chapter 4. The last chapter presents the development of a new technique, TIR-continuous photobleaching, which is complimentary to FRAP.

Figures 1.7

A.
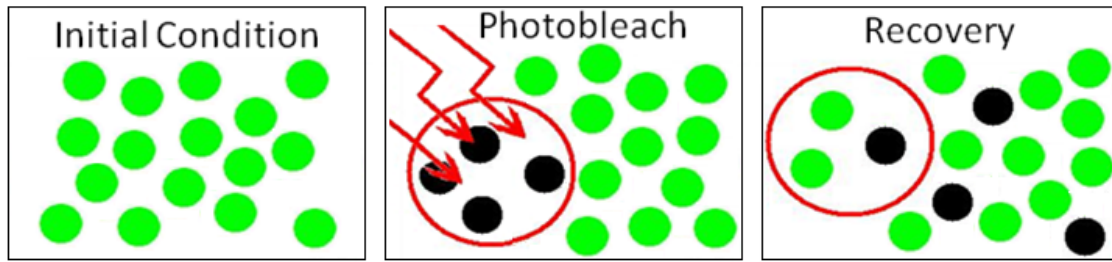
B.
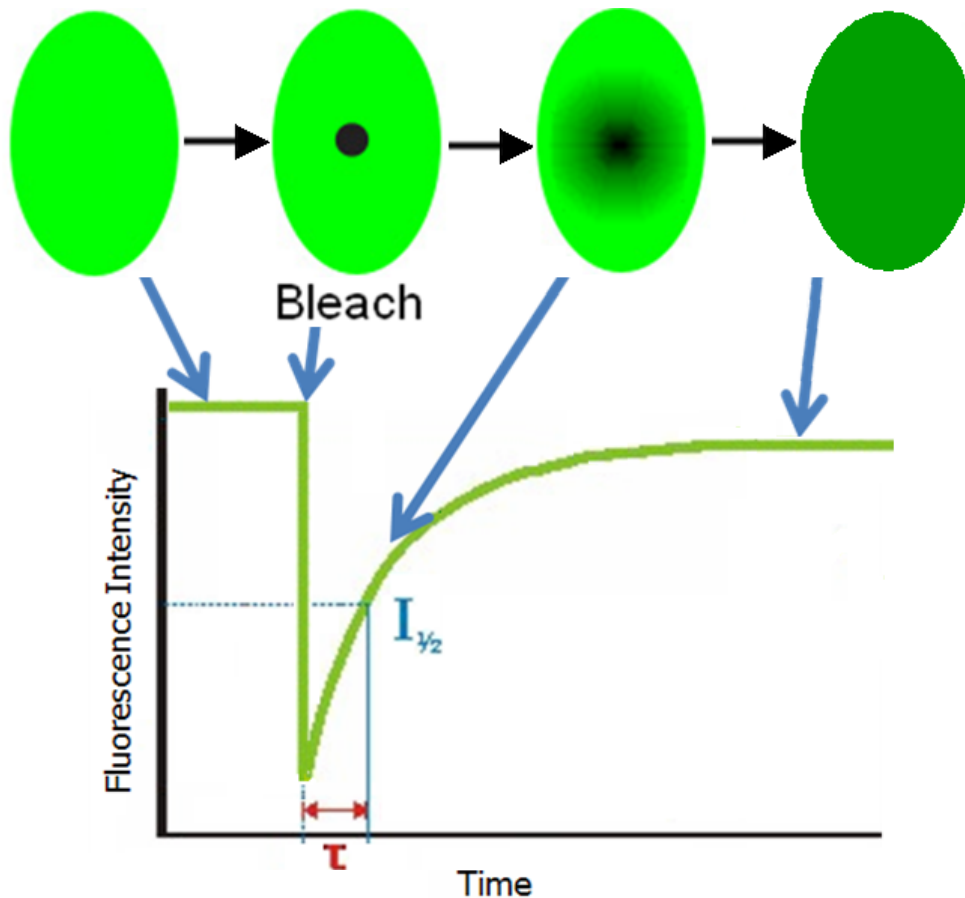
Figure 1.1 Schematic of fluorescence recovery after photobleaching (FRAP). (A) Fluorescent molecules (green) are bleached (black) with a laser (red). If the molecules are mobile, the bleached ones will diffuse out of the bleached region to establish a new equilibrium with the fluorescent molecules. (B) A typical recovery curve observed from the bleached region of a sample during a FRAP experiment. Immediately after the bleach, the fluorescence intensity drops. With time, the fluorescence gradually increases as fluorescent molecules diffuse into the bleached region. The characteristic recovery time, τ, is related to the diffusion coefficient of the molecules.
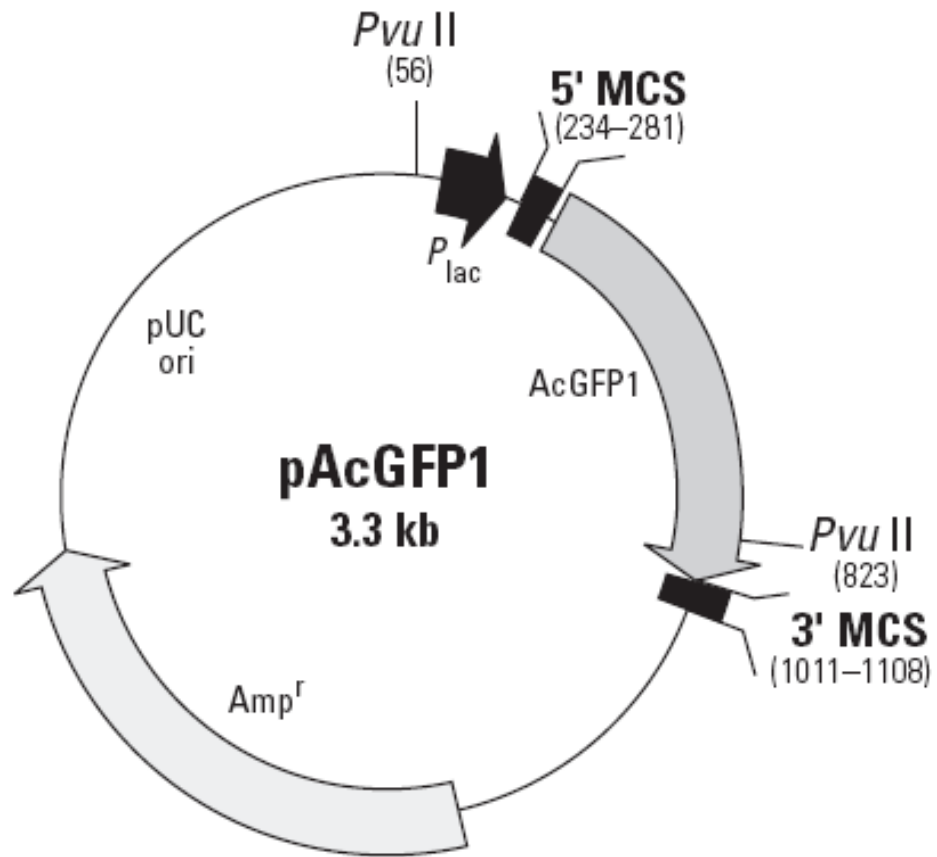
Figure 1.2 pACGFP1 vector map. (Clontech) This vector contains the GFP gene (AcGFP2) flanked by multiple cloning sites (MCS). The gene is regulated by the *lac* promoter (P_{lac}).
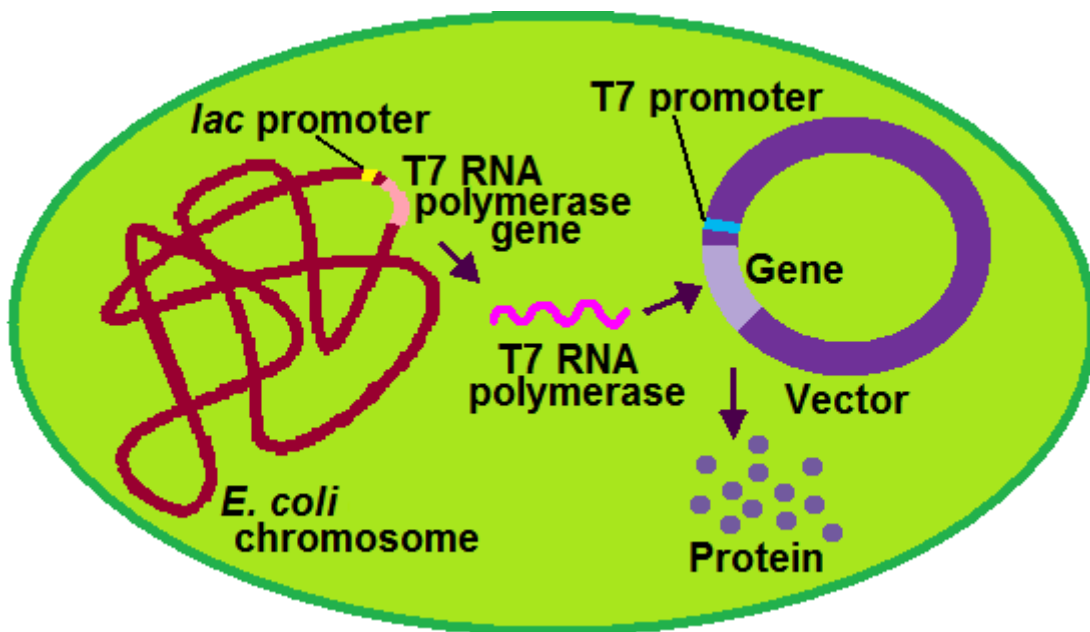
Figure 1.3 The *T7* Promoter. The DNA (maroon) of a BL21(DE3) GOLD *E. coli* cell (green) contains the *T7* RNA polymerase gene (pink) under the control of a *lac* promoter (yellow). T7 RNA polymerase binds to the T7 promoter (blue) of the recombinant vector (purple) to turn on transcription of the gene of interest.
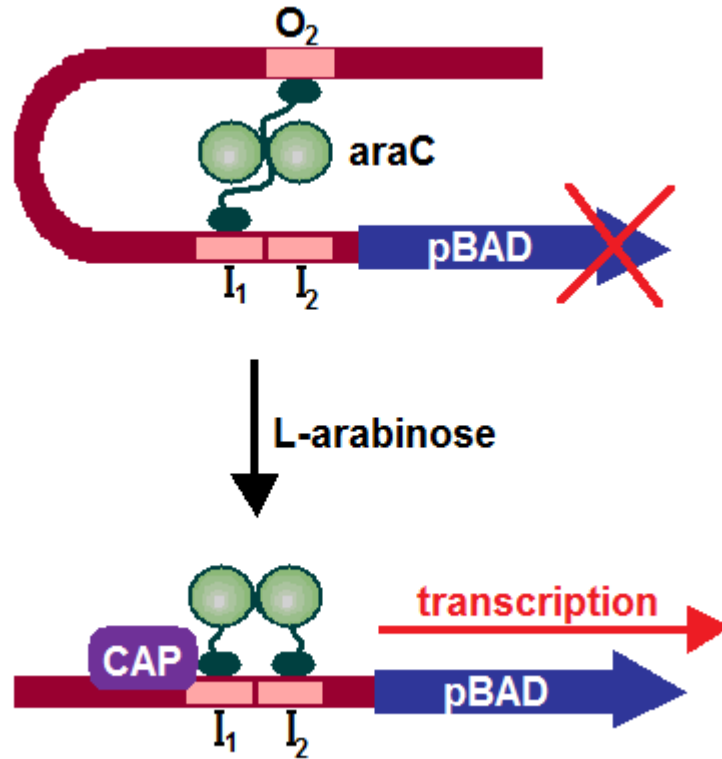
**Figure 1.4 The *araBAD* Promoter.** The araC protein binds the *araBAD* promoter to prevent transcription. Arabinose binds to araC, releasing this repressor protein from the araBAD operon, turning on expression. CAP and cAMP are also involved in turning on transcription in this system. Glucose lowers cAMP levels, decreasing the binding affinity of CAP, and consequently inhibiting protein expression.

Chapter 2: Attempts to Control the Expression of a Protein and Measure its Diffusion Coefficient

Summary

This chapter describes my initial attempts to regulate protein expression in *E. coli,* with the intention of then measuring the translational diffusion coefficient of the protein being controlled. Since the principal means for obtaining diffusion coefficients involves fluorescence, GFP was the primary protein. To provide more flexibility for choosing a protein of interest, however, fluorescent labeling with FlAsH dye and isotopic enrichment with ^{19}F were also explored. Here, I report the results, obstacles, and caveats of creating a system for controlled protein expression.

2.1 Introduction

To gain insight about the intracellular ramifications of protein expression level, the translational diffusion coefficient of a protein can be monitored as a function of protein expression levels in *E. coli* cells. Before this experiment can be performed, however, a system for controlling protein expression must be created. At the genetic level, this task may be accomplished by crippling the promoter with site-direction mutagenesis.[48,49] Also, the structural gene of interest can be coupled with various promoters that are known to provide better

control of expression, such as the *araBAD* promoter.[50] At the protein level,

control may be achieved by changing conditions, including expression media,

concentration of inducer, temperature, and incubation time.

      In addition to controlling protein expression, it is also necessary that the

protein be detectable. FRAP, for example, requires the protein of interest to be

fluorescent. Although this obstacle may be overcome with the use of GFP,

FlAsH, a fluorscein arsenical helix dye, may provide more flexibilty to monitor the

diffusion of a range of other proteins.[51-54]   This small fluorescent dye (480 Da)

can be used to selectively label a protein of interest within living cells by

genetically encoding a tetracysteine motif into the protein (Fig. 2-1). Specifically,

the protein is modified to include a surface exposed Cys-Cys-Gly-Pro-Cys-Cys.

The arsenic groups of the FlAsH bind the thiols of the cysteine residues with high

affinity. Non-specific binding can be minimized by adding of 1,2-ethanedithiol

(EDT) or 2,3 dimercaptopropanol (British Anti-Lewisite, BAL). Furthermore, the

unbound form of the dye, FlAsh-EDT$_2$, is essentially non-fluorescent, but

becomes 50,000 times more fluorescent when bound to the tetracysteine motif.[51]

While FlAsH is typically permeable to eukaryotic cells, the peptidoglycan layer of

E. coli, decreases the permeablity.[55] Consequently, bacterial cells must be

pretreated with small amounts of lysozyme to disrupt the outer cell membrane

before adding the dye. Cells exposed to this pretreatment appear to be viable

and morphologically intact.[56]

      If fluorecent probes prove impractical, NMR may provide an alternative

means to measure protein diffusion by using pulse field gradients.[57] To this end,

the sample is exposed to a linear magnetic field gradient, such that chemically equivalent spins experience different magnetic fields based on their position. Since the magnetization vectors point in different directions, the net magnetization is zero, and no NMR signal is detected. After a set amount of time, an equivalent linear gradient is applied in the opposite direction. If the nuclei remain in the same position during both pulses, then the two gradients cancel one another yielding a signal of unit intensity. If the molecules diffuse between pulses, complete refocusing cannot occur and the signal is attenuated by an amount dependent on the molecule's translational diffusion coefficient. Thus, the diffusion coefficient can be obtained from plotting the gradient strength versus peak intensity of the resulting NMR signal.

While NMR is not restricted to fluorescently-labeled systems, it does require that the protein of interest contain NMR-active nuclei. Most in-cell NMR studies depend on the detection of ^{15}N enriched proteins, but ^{19}F labeling provides several advantages.[31] First, the sensitivity of ^{19}F in NMR experiments is greater than that of ^{15}N and ^{13}C due to their respective gyromagnetic ratios (γ). Second, the ^{19}F spectra are simple since only one, or a few, labeled amino acids are used. [19]F labeling is ideal for biological studies because fluorine it is not native to the cell. Consequently, only the protein of interest will contain ^{19}F, minimizing background metabolite signals.

Since fluorine is not native to the cell, the protein of interest must be selectively enriched with [19]F. Enrichment can be accomplished by expressing the protein in media containing only 3-fluorotyrosine.[58] Alternatively, Mehl and co-

workers have developed a site-specific technique for *E. coli* in which a codon in the gene of interest is mutated to an amber stop codon, TAG.[59,60] The vector containing the gene of interest is then cotransformed with another plasmid, pDULE, which encodes the orthogonal RNA synthetase needed to incorporate a fluorinated unnatural amino acid, L-4-trifluoromethylphenylalanine (tfmF) at the inserted amber stop codon.

This chapter aims to address several strategies for controlling protein expression in *E. coli* while detecting the protein's diffusion coefficient. These strategies include enriching the protein with ^{19}F, labeling with FlAsH dye, cloning the gene of interest into various promoters, crippling the promoter, and changing expression conditions. The disadvantages and shortcommings of each tatic are discussed.

2.2 Materials and Methods

2.2.1 Creating the Vectors

All mutations were completed with a site-directed mutagenesis kit (Stratgene) and sequences were confirmed at the UNC-CH genome analysis facility. PCR was performed with pfu turbo polymerase and an Eppendorf Mastercycler®. Primers were obtained from UNC-CH nucleic acids core facility. Restriction enzymes were purchased from NEB. DNA fragments were analyzed by means of 1% (w/v) agrose (Sigma) gels electrophoresed at 150 V for 30 min. Ligations were performed with T4 DNA Ligase (NEB) at 16 ºC overnight.

Lac mutations

The *lac* promoter of the pAcGFP1 (Clontech) vector was mutated according to Table 2-1.

T7-GFP

The GFP gene from the pAcGFP1 vector was amplified by using PCR with primers 5' CACTTTATGCTTCCGCGGCTCGTATGTTGTGTG 3' and 5' TAAACA AATAGGGGTTCCGCGCACATTTCC 3'. The PCR product was digested with *NcoI* and *EcoRI* to yield three DNA fragments, which were separated by agrose-gel electrophoresis. The 742 bp band was excised and purified with a gel extraction kit (Quiagen). The pT7-7 vector containing the human α-synuclein gene was mutated to incorporate a *NcoI* site just before the start codon by using forward primer 5' CTTTAAGAAGGAGATATACCCATGGATGTATTCGTGAAAGGAC 3'. The resulting vector was digested with *NcoI* and *EcoRI*, such that the GFP gene could be cloned into the *Nco I* and *EcoRI* sticky ends of vector to create the T7-GFP vector. The T7-GFP vector was then mutated according to Table 2-2.

pBAD-GFP

The GFP gene from the pAcGFP1 vector was amplified by using PCR with primers 5' CACCCATGGTGAGCAAGGGCGC 3' and 5' TCATCACTTGTACAG CTCATCCATGC 3'. The PCR product was TOPO-cloned into pBAD-TOPO (Invitrogen) and the polyhistidine region was removed by digestion with *NcoI* and subsequent ligation.

Stop codons (TAG) were incorporated into the GFP gene of the *pBAD-GFP* vector according to Table 2-3 to create five separate mutants.

FlAsH mutations

The α-synuclein-100 gene in the pT7-7 vector[61] was modified to include mutations coding for Cys-Cys-Gly-Pro-Cys-Cys, which specifically binds FlAsH, as an extension of its N-terminus. This mutation was accomplished by using forward primer 5' TCAAAAAGGACCAGTGCTGCCCGGGCTGCTGCTAGG GCAAGAATGAA 3'. The α-synuclein-100 gene codes for a truncated version of the protein containing only the first 100 residues of wild type α-synuclein. It was created by inserting a stop codon after the first 300 nucleotide residues. To generate wild-type α-synuclein that could be tagged with FlAsH, the stop codon was removed from the α-synuclein-100 FlAsH-modified gene by using forward primer 5' GGGCTGCTGCTTGGGCAAGAATGAAGAAGG 3'.

For FlAsH experiments with CI2, the CI2 gene in the pET-28a(+) vector[62] was modified with forward primer 5' GGCACCATCGTGACCATGTGTTGTCCG GGCTGTTGTGAATATCGCATCGATCGC 3'. The FlgM gene[63] was modified with forward primer, 5' CAGACCATGAGCATTGACTGTTGTCCGGGCTGTTG TCGTACCTCACCTTTG 3'.

19

2.2.2 Protein Expression

pAcGFP1, lac mutations, T7-GFP, CI2, and FlAsH mutations

The vectors were individually transformed into BL21-Gold (DE3) competent *E. coli* cells and streaked onto Luria Broth plates containing 100-µg/mL ampicillin (LB_{AMP}) or, for the CI2 containing cells, 60-µg/mL kanamycin (LB_{KAN}). A 5-mL starter culture of liquid LB_{AMP} or LB_{KAN} was inoculated with a single colony and grown overnight at 37 ˚C with constant shaking at 225 rpm. One-mL of this starter culture was used to inoculate 25 mL of fresh LB_{AMP} or LB_{KAN} in a 250-mL flask. Once the optical density at 600 nm (OD_{600}) was between 0.5 to 0.7 or 1.0 (for the FlAsH mutants), the culture was induced with a final concentration of 1-mM isopropyl-β-D-thiogalactopyranoside (IPTG) and allowed to grow at 37 ˚C with constant shaking at 225 rpm.

pBAD-GFP and T7-GFP vectors

The pBAD-GFP vector was transformed into TOP10 or DH10B competent *E. coli* cells and streaked onto LB_{AMP} plates. The *T7*-GFP vector was transformed into BL21-AI competent *E. coli* cells and plated on LB_{AMP} plates. The cultures were then prepared as described above. Once the OD_{600} reached between 0.5 to 0.7, the cultures were divided into 5-mL aliquots, induced with varying amounts of arabinose (0.0001% to 2% w/v), and grown at 37 ˚C with constant shaking at 225 rpm for 3 h.

^{19}F mutants: Incorporation of an unnatural amino acid

Each pBAD-GFP TAG mutant was co-transformed with the pDule vector[60] into TOP10 competent *E. coli* cells and plated on Luria Broth plates containing 100-µg/mL ampicillin and 25-µg/mL tetracycline ($LB_{AMP/TET}$). A 50-mL starter culture of liquid $LB_{AMP/TET}$ was inoculated with a single colony and grown overnight at 37 °C with constant shaking at 225 rpm. This starter culture was used to inoculate 1-L of autoinduction media [25-mM phosphate buffer (pH 7), 50-mM NH_4Cl, 5-mM Na_2SO_4, 1-mM $MgSO_4$, 10% glycerol (w/v), 40% glucose (w/v), 20% arabinose (w/v), 5% aspartate (w/v), 4 mg/mL leucine, 0.2 mg/mL each the amino acids excluding tyrosine and cysteine, 4-µM $CaCl_2$, 2-µM $MnCl_2$, 2-µM $ZnSO_4$, 0.4-µM $CoCl_2$, 0.4-µM $CuCl_2$, 0.4-µM $NiCl_2$, 0.4-µM Na_2MoO_4, 0.4-µM H_3BO_3, 10-µM $FeCl_3$] as described by Hammill *et al.*[14] *L*-4-trifluoromethylphenylalanine (tfmF) was added to a final concentration of 1 mM, 30 min after inoculation, and cultures were grown at 37 °C with constant shaking at 225 rpm for 16-20 h. For a negative control, cultures were grown as described above, except that tfmF was never added.

^{19}F Tyr expression

3-Fluorotyrosine was incorporated into CI2 and GFP using the method described by Khan *et. al.*[58] Briefly, the pET-28a(+) vector containing the CI2 gene under the *T7* promoter and pAcGFP1 were individually transformed into BL21-Gold (DE3) cells and streaked onto plates containing 60- µg/mL LB_{KAN} or 100 µg/mL LB_{AMP}, respectively. A 5-mL starter culture of liquid LB_{KAN} or LB_{AMP}

was inoculated with a single colony and grown overnight at 37 °C with constant shaking at 225 rpm. This starter culture was used to inoculate 400 mL of 2x TY media containing either 30-µg/mL kanamycin or 100-µg/mL ampicillin and grown at 37 °C with constant shaking at 225 rpm. Once the OD_{600} was between 0.8 and 1.0, the culture was centrifuged (Sorvall RC-5B with SS-34 rotor) at 4,000 rpm for 15 min and the pellet was resuspended in 2 L of M9 media [70-mM sodium phosphate buffer (pH 7.3), 0.5-g/L NaCl, 1-g/L NH_4Cl, 2-mM $MgSO_4$, 0.4 % glucose, 0.1-mM $CaCl_2$]. This culture was then allowed to grow at 25 °C with constant shaking at 225 rpm. When the OD_{600} reached 0.6, 120 mg of *L*-tryptophan (Sigma), 120 mg of *L*-phenylalanine (Sigma), 140 mg of 3-^{19}F-tyrosine (Sigma), and 1 g of glyphosate (Sigma) were added. After 30 min, IPTG was also added to a final concentration of 1 mM, and the culture was incubated overnight at 25 °C with constant shaking at 225 rpm.

2.2.3 Protein Analysis

One-mL aliquots of culture were centrifuged at 10,000 g for 15 min and the pellets were resuspended in 30 uL of 20-mM sodium phosphate buffer (pH 7) with 10 uL of protein loading dye. After centrifuging at 17,000 g for 30 min to remove cellular debris, 15-uL samples were loaded onto 18% Tris-HCl Criterion Precast sodium dodecyl sulfate polyacrylamide gels (SDS-PAGs). The gels were electrophoresed at 200 V for 60 min. Fluorescence emission scans of the gels were collected with the BIO-RAD Versa Doc Imaging System (GE Healthcare). The gels were stained with Comassie Brillant Blue and destained with an

aqueous solution of 10% (v/v) methanol and 10% (v/v)acetic acid.

2.2.4 [19]F Protein Purification

Protein purification was performed by Heidi Scronce**.** The cultures were
harvested by centrifugation (Sorvall RC-3B rotor) at 1,600 g for 30 min at 4˚C.
The pellets were resuspended in lysis buffer [20-mM Tris (pH 8.0), 1-mM EDTA,
1-mM PMSF or 20-mM Tris (pH 7.4), 200-mM NaCl, 1-mM EDTA, 10-mM 2-
mercaptoethanol for MBP] and pulse sonicated (Branson Ultrasonics) at 4˚C
(18% amplitude) for 10 min. Cell debris were removed by centrifugation (Sorvall
RC-5B with a SS-34 rotor) at 27,000 g. The supernatant was stirred with
streptomycin sulfate (10 mg/mL) at 4˚C for 30 min and then centrifuged. The
new supernatant was stirred with ammonium sulfate (244 mg/mL) at 4˚C for 1 h
and then centrifuged. More ammonium sulfate (200 mg/mL) was added to the
supernatant and the GFP was extracted twice with ethanol (300 µL/mL). Butanol
(250 µL/mL) was added to the ethanol layer and GFP partitioned into the
aqueous layer. Equal parts chloroform and aqueous layer (containing GFP) were
mixed. The top layer, containing GFP, was dialyzed against water and then
concentrated in an Amicon Ultra MWCO 3,000 centrifugal filter unit (Millipore).
The protein was further purified by using size exclusion chromatography (16/60
Superdex[TM] 75) with water as an eluant.

2.2.5 FlAsH-labeling

One-mL aliquots of culture were centrifuged at 10,000 g for 15 min and
the pellet were resuspended in 700 µL of HEPES containing 25 µL of 0.1-mg/mL

lysozyme. One µL of FlAsH-EDT$_2$ reagent (Invitrogen) was diluted with 99 µL of

HEPES buffer (Invitrogen) and added to the resuspended pellet. After incubating

for 45 min at room temperature, the sample was centrifuged (Eppendorf model

5418) for 10 min at 4,000 g and the pellet was washed twice with 1.6 mL of 1X

BAL buffer (Invitrogen). After the second wash, the pellet was resuspended in

100 µL of minimal media [60-mM K$_2$HPO$_4$, 7.6-mM (NH$_4$)$_2$SO$_4$, 2-mM MgSO$_4$, 20-

µM FeSO$_4$, 1-mM EDTA (pH 6.8)] and injected into the sandwich slides prepared

as described in Chapter 3. After 30 min at room temperature, the sample

chamber was rinsed with minimal media and sealed with vacuum grease. As a

negative control, cells expressing the wild-type protein were also exposed to the

FlAsH dye as described above.

2.2.6 Microscopy

   Images of *E. coli* cells containing FlAsH-tagged α-synuclein were acquired

with a Zeiss 510 Meta scanning confocal inverted microscope equipped with a

30-mW argon laser. The sample was imaged with a 63x, 1.4 NA plan-

apochromat oil-immersion objective and a pinhole of 2.0 Airy-disc units. An

excitation wavelength of 488 nm was selected. A 505-nm long pass filter was

used for detection. Entire cells were bleached with a ~240 µW bleach beam

rastering across the surface for 52 ms. Both bleached and unbleached cells

were then monitored with an observation beam (~0.2 µW) for 6-10 s.

2.2.7 CD

Circular dichroism (CD) studies were performed on an Aviv 62DS spectropolarimeter with a rotary sample changer. Protein samples of 0.02 mM were loaded into 0.1 cm quartz cuvettes. Data were collected at 1-nm intervals from 180 to 375 nm with an averaging time of 3 s/point. Three scans were collected per sample.

2.2.8 NMR

NMR was performed by Charlie Li. Cultures were centrifuged at 4,000 x g (Sorvall, RC-3B) for 30 min at 4 ^{o}C and the pellets were resuspended in 2 mL of LB media. A 90:10 mixture of culture:D_2O was placed in a 5-mm NMR tube for data acquisition. After acquisition, the supernatants were collected by gentle centrifugation to assess leakage. The pellets were resuspended in lysate buffer (50-mM Tris, pH 8.0) to a final volume of 1 mL and the slurries were sonicated (Fisher Scientific, Sonic Dismembrator Model 500) on ice for 10 min with a duty cycle of 2 sec on, 5 sec off. The lysate was collected after centrifugation at 14,000 x g (Eppendorf, model 5418) for 10 min.

^{19}F spectra were acquired at 37 ^{o}C on a Varian Inova 600-MHz spectrometer equipped with a 5-mm ^{19}F(1H) z-gradient probe. The spectra were acquired with a 30-kHz sweep width and consisted of 2048 scans with a 30-kHz sweep width. The pulse width was 60°. The pulse spacing was 2 s.

2.3 Results

2.3.1 FlAsH

Proteins α-synuclein-100, α-synuclein, CI2, and FlgM were each tagged with FlAsH as shown from both inspection of fluorescence-emission scans of the SDS-PAGs and of confocal fluorescent images. Only a fraction of the cells, however, were fluorescent, indicating inefficient labeling. Cells expressing wild-type protein that were exposed to the FlAsH dye showed little to no fluorescence.

To determine whether FlAsH was an appropriate dye for FRAP, entire cells were exposed to the bleach beam and the fluorescence of the individual cells was observed with time (Fig. 2-2). After the bleach, the cell's fluorescence had greatly decreased, but recovered with time. Furthermore, the fluorescence intensity of cells exposed only to the observation beam decreased over the 6-s period of observation.

2.3.2 *Lac*-GFP

Several growth conditions were investigation for the *lac*-GFP system. First, the IPTG concentration was varied from 0 to 2 mM in both the BL21-Gold (DE3) and DH5α strains. The expression level was determined based on fluorescent and Coomassie-stained SDS-PAGs of the cell lysates. For both strains, the GFP expression level was independent of inducer concentration (Fig. 2-3A). Higher levels of GFP were expressed in the BL21-Gold (DE3) cells than in the DH5α cells. Lengthening the post-induction incubation increased the GFP expression (Fig. 2-3 B); however, the OD_{600} increased with incubation. This

result suggests that the increased GFP expression is more likely due to higher cell density than to more protein per cell. The incubation temperature (room temperature, 30 °C, and 37 °C), culture media (M9, Martex, and LB), and glucose or glycerol concentrations (0-3%) were also altered. Although different combinations of these factors appeared to give varying levels of GFP expression, none of the conditions yielded reproducible amounts of GFP relative to one another (Fig. 2-4). High concentrations of salt in the media appeared to hinder the growth of cells containing the pAcGFP1 vector. Thus, the salt concentration in the LB medium was limited to 5 g/L. GFP could not be detected for constructs harboring the *lac* promoter mutations (Fig. 2-4 B).

2.3.3 *T7-GFP*

The expression alterations described above were repeated with GFP under the T7 promoter. Using a combination of glucose and IPTG, the GFP expression level was able to be reproducibly controlled (Fig. 2-5 A). Likewise, the various mutations of the T7 promoter resulted in varying amounts of GFP expression (Fig. 2-5 B).

2.3.4 *T7-*CI2

To assess the level of control provided by the *T7* promoter, *E. coli* cells expressing CI2 were induced with varying amounts of IPTG and incubated for a range of times after induction. Unlike the *Lac*/GFP system, CI2 expression could be controlled with inducer concentration (Fig. 2-6A) and with time (Fig. 2-6B).

2.3.5 *araBAD-GFP*

The level of GFP expression under the *araBAD* promoter could be controlled by varying the arabinose concentration. As shown in Fig. 2-7, however, the maximum expression level of GFP under the *araBAD* promoter is significantly lower than that with the *lac* promoter. In attempts to increase expression, several growth conditions were altered, including incubation temperature (room temperature, 30˚C, and 37˚C), *E. coli* strain (TOP10, DH10B, and BL21-AI cells), and incubation time after induction (3h, 4h, 5h, and overnight). Of these changes, only the overnight incubation increased GFP expression according to fluorescent band intensities from SDS-PAGs (not shown). Based on optical density, though, this increase is more likely due to higher cell density than to more protein per cell.

2.2.6 ^{19}F

To incorporate ^{19}F into GFP, the gene was mutated to include an amber stop codon (TAG) sequence at residue 37, 47, 144, 222, or 240. The Phe222 mutant was expressed in the presence of the unnatural amino acid, tfmF. No GFP was observed without the addition of tfmF (Fig. 2-8 A). Similar results were obtained for the Tyr37 mutant (data not shown). For the position 240 mutant, GFP was expressed regardless of whether the unnatural amino acid was present (Fig. 2-8 B). No expression was observed for the Phe47 and Tyr144 mutants. To confirm that incorporation of the unnatural amino acid did not alter the structure of GFP, CD spectra were collected. The CD spectra of Tyr37 and

Phe222 variants closely resemble that of wild type GFP (Fig. 2-9). ^{19}F NMR spectra were collected for both the purified and in-cell samples of Tyr37 and Phe222 (Fig. 2-10). The in-cell protein resonances are broader than the peaks from the purified proteins. ^{19}F NMR was also performed on the 3-fluorotyrosine GFP and CI2. Both the in-cell and cell lysate spectra for CI2 showed two resonances, while NMR spectrum of the purified CI2 sample had only a single peak (Fig. 2-11). The -49.45 ppm resonance is from the 3-fluorotyrosine. No signal was observed for the GFP sample.

2.4 Discussion

The ultimate goal of this project is to measure the diffusion coefficient of a protein in *E. coli.* Since the primary method for accomplishing this task is FRAP, it was first necessary to find a suitable fluorescent probe. FlAsH seemed to be an ideal canidate because of its ability to permeate the cell mebrane and to specifically bind the protein of interest. Based on the location of the fluorescent bands on SDS-PAGs, the desired proteins could be tagged with FlASH. While these preliminary results showed promise, confocal imaging revealed that only a small fraction of cells exhibited the desired fluorecence. This poor labeling efficiency could reflect the inability of the dye to penetrate the additional cell wall not present in eukaroytic cells. Lyosyme has previously been shown to disrupt the cell peptoglycan layer enough to allow the FlAsH to pass through;[56] however, the concentration may need optimization. FlAsH experiments were abandoned before performing such optimization, because the dye was inappropriate for

29

FRAP studies. The recovery observed from bleaching whole cells indicates that

FlAsH bleaches reversibly (Fig. 2-2). Furthermore, FlAsH is reversibly bleached

even by a low-insensity observation beam. Due to these results and the lack of

options for appropriate intrcellular fluorecent probes, GFP was used for further

experiments.

Once GFP had been established as the probe of interest, it was necessary

to regulate its intracellular concentration. Although options for controlling protein

expression are available, altering inducer concetration provides a straight-

forward approach without the need to modify the vector. In theory, IPTG should

induce expression under the *lac* and *T7* promoters. However, SDS-PAG

analysis showed that GFP under the control of the *lac* promoter, was expressed

regardless of whether inducer was added (Fig. 2-3). This conclusion was

supported by the observation of the green color of colonies on the agar plates

(for cells containing pAcGFP1) that had not been exposed to IPTG. In contrast,

GFP expression under the *T7* promoter was lower without IPTG (Fig. 2-5 A).

Unfortunately though, even a low IPTG concentration (5 µM) induced the same

amount of GFP as higher concentrations (data not shown). This observation was

somewhat surprising since Elowitz *et al*. concluded that the concetration of IPTG

is directly related to the level of protein expression after measuring a decrease in

GFP diffusion when the IPTG concentration was doubled. We repeated the

expression experiment in DH5α cells, the same strain of *E. coli* used by Ellowitz,

but still failed to obtain a correlation between IPTG concentration to the amount

of GFP expressed.

Although changing the IPTG concentration failed to provide control, more protein was produced with longer incubation times for both the *T7* and *lac* (Fig. 2-3 B) systems. Since a constant volume of cell lysate was loaded per lane of the gel, this result could either be due to more protein per cell, more cells per volume of culture, or both. To distinguish between the two extremes, the OD_{600} of the cultures were measured. These optical densities increased with incubation time. Furthermore, when the band intensities on the gel were divided by the corresponding OD_{600}'s, the normalized intensities were within error of each other. These observations suggest that the increased protein is a result of higher cell densities, not more protein per cell.

Another option for controlling protein expression without genetic modifications is by growing the cells in different media. For in-cell NMR experiments, the cells are often transferred to a minimal media, such as M9 or Spectra 9, before inducing expression. As the name implies, minimal media contains only the essential ingredients necessary for cell survival. Consequently, protein expression is often lower in M9 or Spectra 9 than in LB. This result was observed in Fig. 2-4, (A, lanes 5, 8, 9 and B, lanes 1, 5, 9), but it was not reproducible (A, lanes 1-3).

The carbon source in the media has also been reported to affect protein expression by controlling the intracellular cAMP levels.[64,65] As expected, addition of 0.5 % glucose significantly decreased the intensity of the GFP band on the SDS-PAG by decreasing the cAMP concentration (Fig. 2-4 A, lanes 4 and 10), and 1% glycerol increased the band intensity (A, lane 6; b, lane 13).

Unfortunately, confocal imaging showed that the fluorescence varied considerably from cell to cell, suggesting that each individual cell contained a different amount of GFP.

Unlike the *T7* and *lac* promoters, the *araBAD* promoter is advertised to provide "tight modulated regulation". As expected, we observed more GFP in cells incubated with higher levels of arabinose (Fig. 2-7). Unfortunately, even at the highest arabinose concentrations, GFP expression (lane 9) was lower than that from the *lac* promoter (lane 1). Such low expression levels result in poor signal-to-noise for the FRAP experiments. In attempts to improve expression levels, several conditions were altered, including incubation time and temperature. Due to the slow folding rate of GFP, GFP expression can be optimized with lower incubation temperatures.[40] Neither incubation at room temperature, nor at 30 °C improved expression. Overnight incubation increased the fluorescence intensity of GFP bands on the SDS-PAG, but only because the extra time results in more cells. Since FRAP is conducted on single cells, this alteration does not improve the signal-to-noise ratio of the experiment.

Since altering the conditions for protein expression failed to provide the desired control, we tried to use mutagenesis to cripple the promoters to varying degrees. The *lac* mutations completely crippled the promoter such that GFP expression was undetectable (Fig. 2-4 B). In contrast, the T7 mutations were more successful (Fig. 2-5 B). The -5C to A mutation completely crippled the T7 promoter. This observation is consistent with a report that the -5C to A mutation decreases protein expression to 1% of the wild-type level.[48] The effects of the

32

-6A to T mutation, which is expected to have 13% activity, were less severe and a small level of GFP could be detected. Furthermore, this level could be altered by using BL21-AI cells and by varying the amount of arabinose inducer (Fig. 2-5 B lanes 2 and 3). Surprisingly, the +3G to A mutation, which according to the literature should have only 74% activity, consistently exhibited a higher expression level than the wild-type promoter (Fig. 2-5 B, lanes 5 and 6). Nevertheless, these mutations provided a range of GFP expression levels. In addition, these levels were greater than those obtained with the *araBAD* promoter, such that diffusion coefficients cold be obtained by FRAP for the highest expression systems. Unfortunately, decreasing the GFP expression level decreased the signal-to-noise for the FRAP experiments, such that diffusion coefficients could not be easily measured for lower concentrations. Likewise, combinations of glucose and IPTG with the *T7* promoter provided control of GFP expression, but again, the less than maximum expression levels made it impossible to obtain reliable diffusion coefficients.

Since diffusion measurements were not obtainable by FRAP at suboptimal GFP concentrations, ^{19}F NMR was investigated as an alternative. Fluorine can be incorporated into proteins by exposing the cells to media containing 3-fluorotyrosine. This method has advantages compared to incorporation of an unnatural amino acid, in that no genetic modifications are required. Short of obtaining a ^{19}F NMR spectrum, however, there is no quick control to ensure that ^{19}F has been incorporated. No ^{19}F signal was observed for GFP, which is consistent with the lack of signal obtained from ^{15}N-enriched GFP in in-cell NMR

studies attempted in our lab. For CI2, NMR spectra of both the purified and in-cell CI2 sample confirm that the protein had been successfully enriched with ^{19}F (Fig. 2-11). Unfortunately, one of the peaks in the in-cell spectrum is at the same chemical shift as the purified peak, suggesting that CI2 leaks out of the cells. This conclusion is confirmed by the spectrum of the cell lysate in which the relative intensity of that peak increases. The second peak in the in-cell and cell lysate spectra is the free amino acid. Thus, the protein is not able to be detected inside of the cells.

Often inability to detect proteins by in-cell NMR is a result of the protein's rotational diffusion being inhibited by the crowded cytoplasm or protein-protein interactions. While global movement of GFP may be restricted, a freely rotating probe attached to the protein may provide enough local movement to allow NMR detection. In the unnatural amino acid, L-4-trifluoromethylphenylalanine (tfmF), the fluorines are free to rotate. Furthermore, since there are three fluorines per tfmF residue, compared to the single fluorine on the modified tyrosine, this strategy should provide an increase in signal.

In a second attempt to enrich GFP with ^{19}F, a TAG mutation allowed incorporation of tfmF. Since tfmF is a phenylalanine analog, mutation sites were chosen by identifying the phenylalanines and tyrosines in the protein and then eliminating those involved in the chromophore or structure. In addition, the stop codon of the GFP (codon 240) was mutated to TAG. This mutation allowed the unnatural amino acid to be incorporated at the end of the protein so as not to interfere with the structure. Based on SDS-PAG, no GFP expression was

observed for the position 47 and 144 mutants. The poor expression may be due to a decrease in protein stability. For the position 144 mutant, tfmF replaces a tyrosine on the edge of the GFP β-barrel. Crystallographic data suggest that this residue stabilizes the protein through an edge-face interaction with a nearby benzyl ring.[41] For the position 222 mutant, a low level of expression was detected. The observed low level was not discouraging, because even the wildtype GFP was expressed at lower levels when the second vector (pDULE) and tfmF were present (Fig. 2-8 A). This reduced expression could be the cell's response to the stress of carrying two vectors and expressing both GFP and the amino acid tRNA synthetase (from pDULE).

As with the ^{19}F-tyrosine method, there is no quick way to prove that the unnatural amino acid was incorporated into the GFP. There is, however, a negative control that can strongly suggest if the experiment is proceeding as planned. Without the unnatural amino acid and the pDULE vector, the incorporated TAG sequence in the GFP gene should serve as a stop codon. Consequently, the mutant GFPs can only be expressed in their entirety with both tfmF and pDULE present. As shown in Fig. 2-8 A, no GFP band was observed for the position 222 mutant in the absence of tfmF. Unfortunately, this negative control was not applicable to the position 240 mutant. Since residue 240 is already the GFP stop condon, the protein will be expressed in the presence or absence of tfmF (Fig. 2-8 B). As a result, this mutant was abandoned once the CD spectra confirmed that incorporation of tfmF at residues 37 and 222 did not significantly alter GFP structure (Fig. 2-9).

As anticipated, the in-cell NMR spectrum for the tfmF studies exhibited broader peaks than the purified sample. This observation is a result of the protein's slower tumbling in the crowded environment of the cell. The sharp peak observed in both the supernant and in the in-cell spectra is the free tfmF. The lack of any other peaks in the supernatant suggests that GFP is not leaking out of the cells. Unfortunately, the signal from the in-cell spectra is too small to perform the desired pulsed-field gradient experiments.

2.5 Conclusion

The original goal of this thesis involved controlling the expression of a single protein inside *E. coli* cells and then measuring the diffusion coefficient of that same protein. Neither FlAsH nor ^{19}F NMR provided appropriate methods. The bleaching prosperities of FlAsH render it unsuitable for FRAP, and the NMR signal from ^{19}F enrichment with tfmF was too small to conduct the desired pulsed-field gradient experiments. The failure of these preliminary experiments limited further work to focus solely on GFP. Intracellular GFP concentrations could be controlled by both promoter mutations and by changing expression conditions. However, the general problem remained that that expressing GFP at lower levels decreases the signal for any type of diffusion study. Consequently, Chapter 3 focuses on the creation of a two-protein system. One protein will be expressed at a constant level and its diffusion coefficient will be monitored while the expression of a second protein is alerted.

2.6 Tables

Table 2-1. Mutations to the *lac* promoter in the pAcGFP1 vector

Mutation[a]	Forward Primer
A168G	5' CACTTTATGCTTCCGGCTCGT**G**TGTTGTGTGGAATTGTGAGCG 3'
T143G	5' CTCATTAGGCACCCCAGGC**G**TTACACTTTATGCTTCCGGC 3'
CG 161 insertion	5' CACTTTATGCTTCCG**CG**GCTCGTATGTTGTGTG 3'

a The number indicates the base in the pAcGFP1 vector where the mutation was created.

Table 2-2. Mutations to the *T7* promoter in the pT7-7 vector

Mutation	Forward Primer
+3G to A	5' GAAATTAATACGACTCACTATAGG**A**AGACCACAACGGTTTCCCTCTAG 3'
-6A to T	5' CGATTCGAACTTCTGATAGACTTCGAAATTAATACGACTC**T**CTATAGGGA 3'
-5C to A	5' CGAACTTCTGATAGACTTCGAAATTAATACGACTCA**A**TATAGGGAGACC 3'

Table 2-3. [19]F GFP mutations

Mutation Site	Forward Primer
Phe 47	5' CAAGCTGACCCTGAAG**TAG**ATCTGCACCACCGGCAAG 3'
Phe 222	5' CGATCACATGATCTAC**TAG**GGCTTCGTGACCGCCG 3'
Tyr 144	5' GCAATAAGATGGAG**TAG**AACTACAACGCCCAC 3'
Tyr 37	5' AGGGCGATGCCACC**TAG**GGCAAGCTGACCCTGAAGT 3'
Stop codon 240	5' GATGAGCTGTACAAG**TAG**TGAGCGGCCGCGACTC 3'

2.7 Figures

A.

FlAsH-EDT$_2$

B.
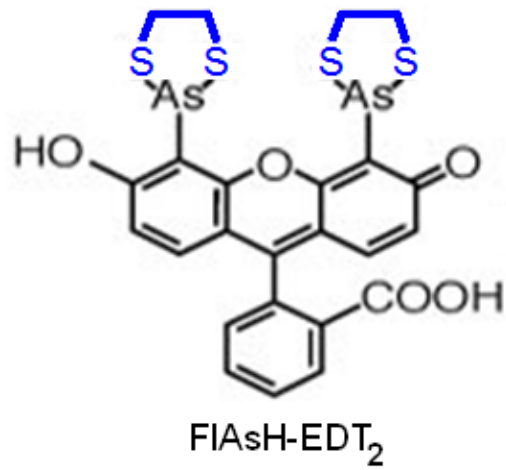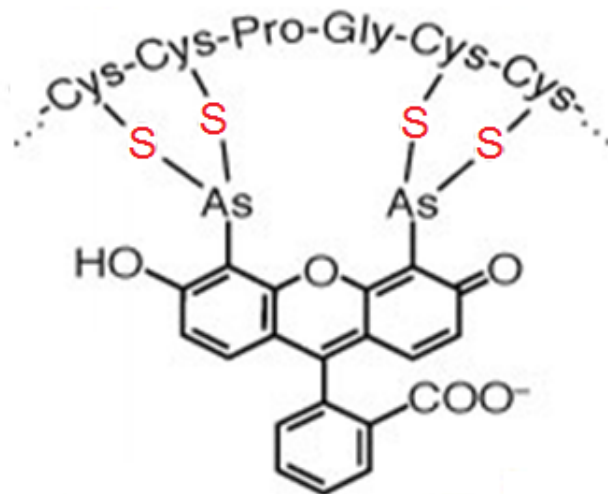
Figure 2-1 FlAsH dye. (A.) Structure of the FlAsH-EDT$_2$ dye (B.) Representation of the three-dimensional structure of a protein containing the CCXXCC mutation. Each As atom in the FlAsH dye forms covalent bonds with two cysteine residues of the protein.

A.

B.
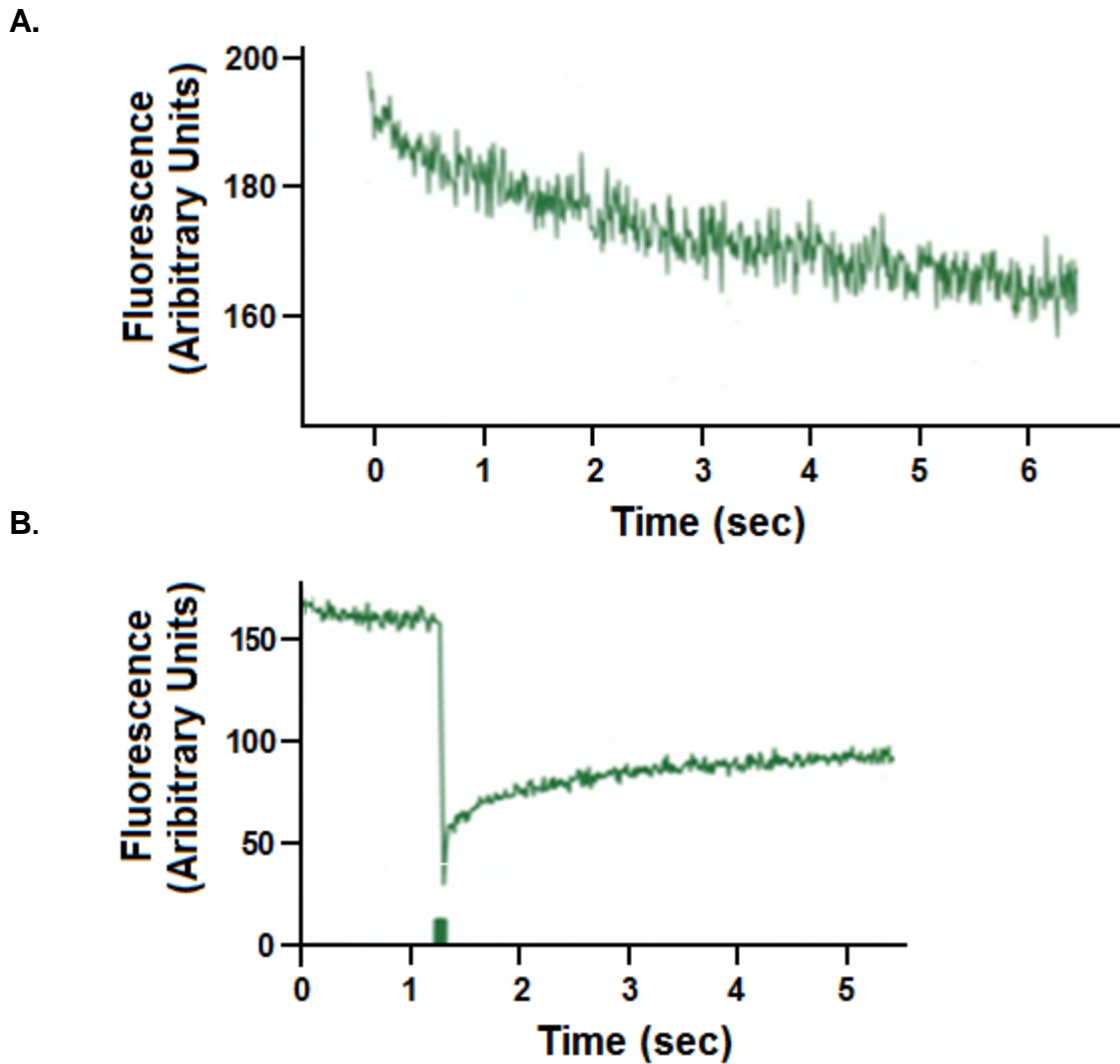
Figure 2-2 FlAsH is not suitable for FRAP experiments. (A.) *E. coli* cells containing FlAsH-tagged α-synuclein monitored by confocal microscopy using a low-power, 488-nm observation beam. (B.) Whole cells were subjected to a 52-ms bleach before monitoring their fluorescence with the observation beam.

A.

    1   2   3   4

B.

    1   2   3   4
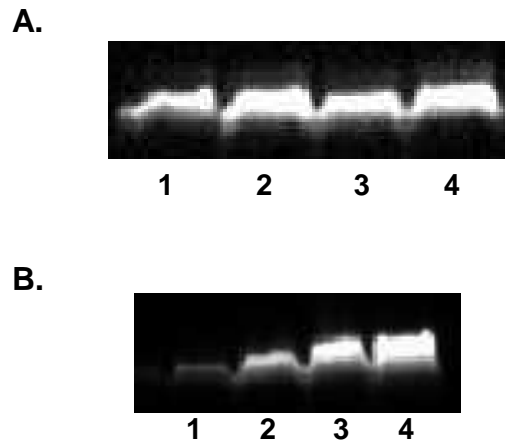
Figure 2-3 GFP expression can be controlled with incubation time, but not with inducer concentration. (A.) *E. coli* expressing GFP under the *lac* promoter were induced with 0-mM (lane 1), 0.1-mM (lane 2), 0.25-mM (lane 3) or 2-mM (lane 3) IPTG and incubated for 3 h. (B.) The *E. coli* were incubated for 30 min (lane 1), 1 h (lane 2), 2 h (lane 3) and 3 h (lane 4) after induction with 1-mM IPTG. Cell lystates were run on SDS-PAGs and the gels were fluorescently scanned.
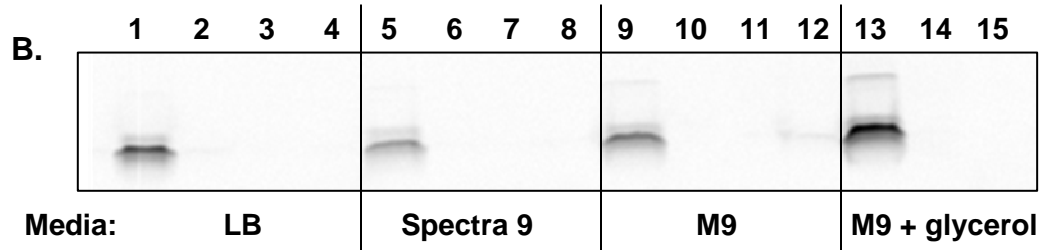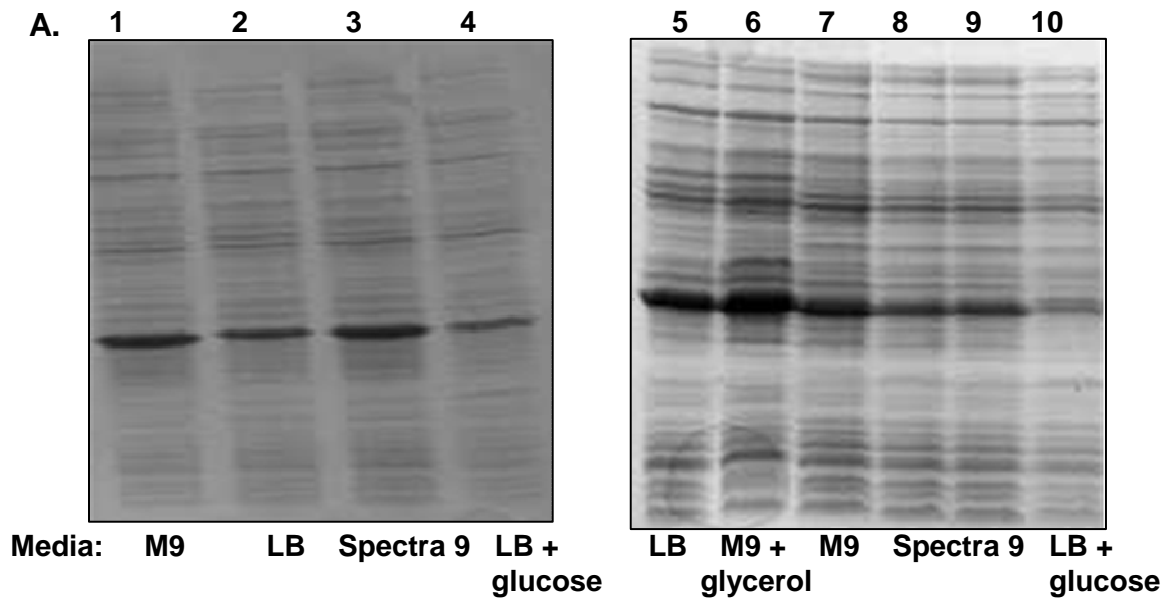
Figure 2-4 GFP expression controlled with media. (A.) *E. coli* expressing GFP were grown in M9 minimal media (lanes 1, 7), LB (lanes 2, 5), Spectra 9 (lanes 3, 8, 9), LB with 0.5% glucose (lanes 4, 10) or M9 with 1% glycerol (lane 6). Cell lystates were run on SDS-PAG and Coomassie stained. (B.) SDS-PAG of *E. coli* cell lysates with mutated *lac* promoters were fluorescently scanned. The *lac* mutations were A168G (lanes 2, 6, 10, 14), T143G (lanes 3, 7, 11, 15), or a CG insertion at position 161 (lanes 4, 8, 12, 16). For comparision, GFP expressed with the wt *lac* promoter are also shown (lanes 1, 5, 9, 13). Cells were grown in LB (lanes 1-4), Spectra 9 (lanes 5-8), M9 minimal media (lanes 9-12), or M9 with 1% glycerol (lanes 13-15).

A.

	1	2	3	4
1% Glucose	+	-	+	-
1-mM IPTG	-	-	+	+

B.

	1	2	3	4	5	6
Arabinose	2%	0.01%	0.2%	0.002%	0.2%	2%
Mutants	-5C to A	-6A to T		+G to A		wt
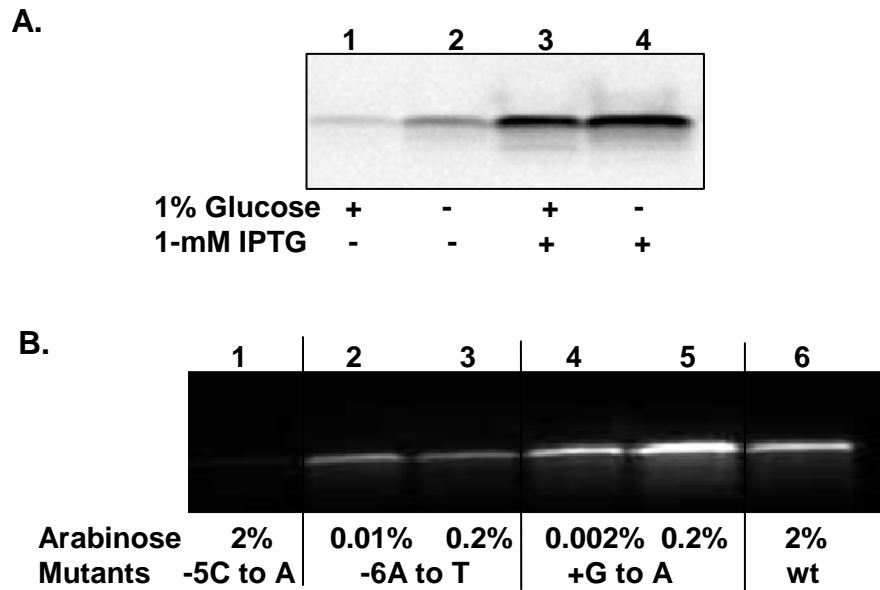
Figure 2-5 GFP expression under the T7 promoter. (A.) BL21-Gold(DE3) *E. coli* cells were grown in the presence (lanes 1, 3) or absence (lanes 2, 4) of 1% glucose. Some samples were induced with 1-mM IPTG (lanes 3, 4), while other samples were allowed to grow without inducer (lanes 1, 2). Cell lystates were run on SDS-PAG and the gel was fluorescently scanned. (B.) The fluorescent image of an SDS-PAG containing cell lysates from BL21-AI *E. coli* harboring mutated T7 promoters. The T7 mutations were -5C to A (lane 1), -6A to T (lanes 2, 3), +3G to A (lane 4, 5). For comparision, GFP expressed with the wt T7 promoter is also shown (lanes 6). The cells were induced with either 0.002% (lane 4), 0.01% (lane 2), 0.2% (lanes 3, 5), or 2% (lanes 1, 6) arabinose.
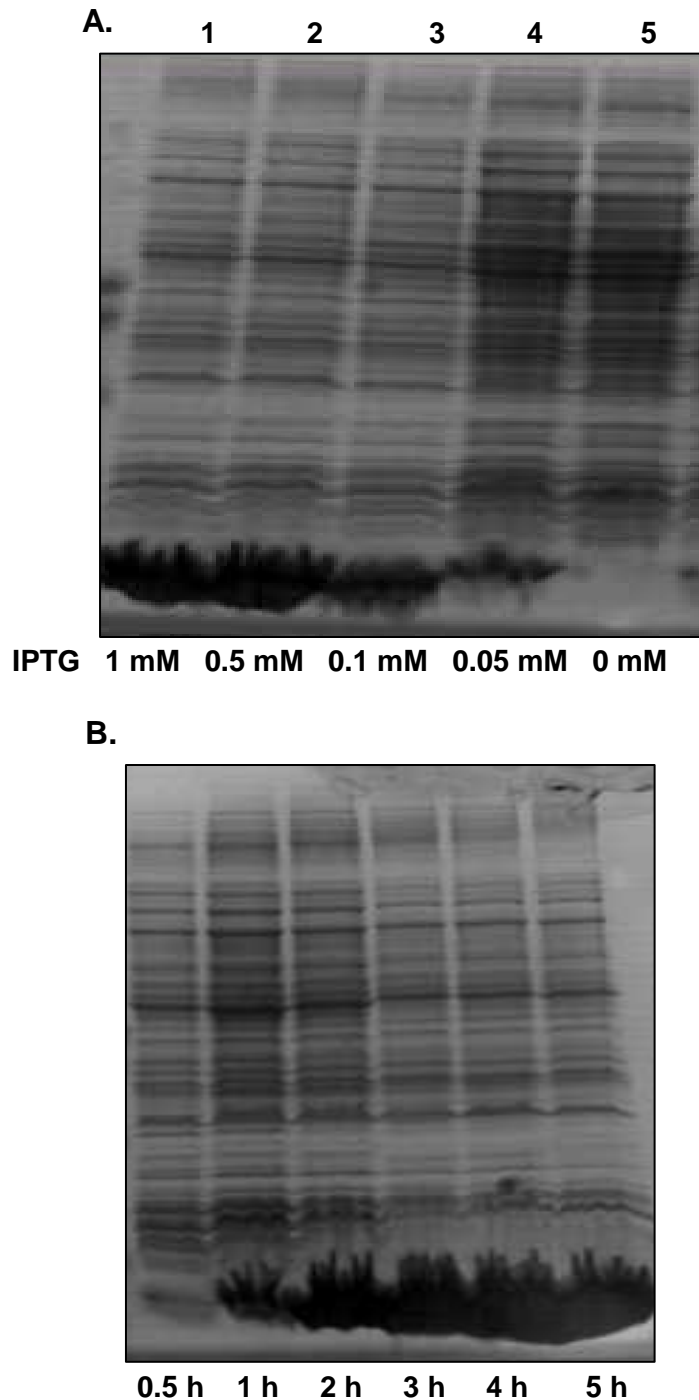
Figure 2-6. CI2 expression controlled with incubation time and inducer concentration. (A.) *E. coli* were induced with 0.01-mM (lane 1), 0.1-mM (lane 2), 0.5-mM (lane 3), 1-mM (lane 4), or 0-mM (lane 5) IPTG and incubated for 3 hours. (B.) Cells that were induced with 1-mM IPTG were incubated for 30 min (lane 1), 1 h (lane 2), 2 h (lane 3), 3 h (lane 4), 4 h (lane 5), or 5 h (lane 6). Cell lystates were run on SDS-PAG and Coomassie stained.
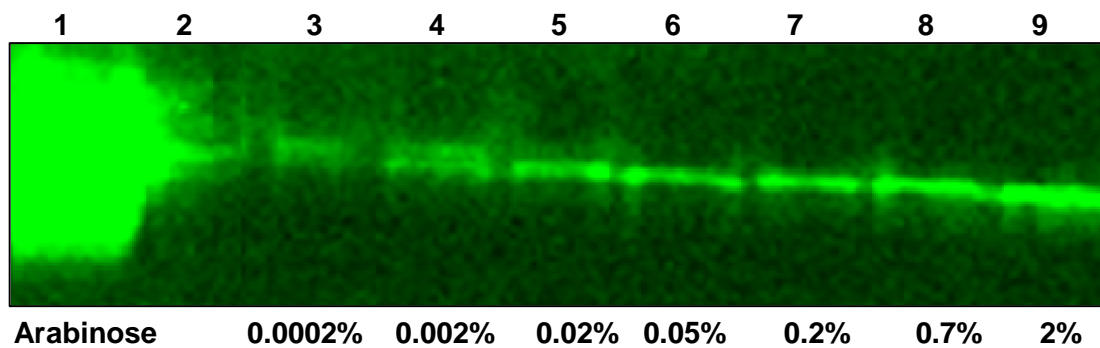
Figure 2-7 GFP expression under the *araBAD* promoter. TOP10 *E. coli* cells containing the pBAD/GFP vector were induced with varying amounts of arabinose as indicated. Cell lystates were run on SDS-PAG and the gel was fluorescently scanned. For comparision, cell lysate from Bl21-Gold (DE3) cells containing pAcGFP1 that have been induced with 1-mM IPTG is shown in lane 1. Lane 2 is intentionally empty.
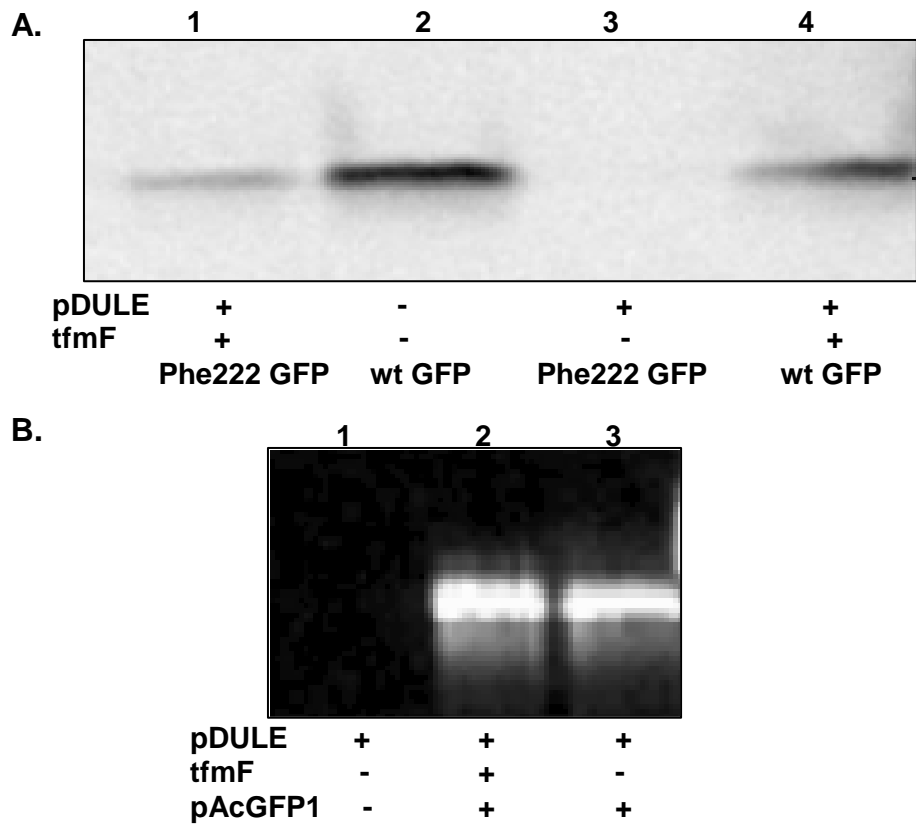
Figure 2-8 Incorporation of the unatural amino acid, L-4-trifluoromethylphenylalanine (tfmF), into GFP. After incubation in the presence and absence of tfmF, cell lystates were electrophoresed on a SDS-PAG and the gel was fluorescently scanned. (A.) The pAcGFP1 (wt) and the Phe222 mutatant were each co-transformed with the pDule vector into *E. coli* cells. (B.) The stop codon mutant (240) was also co-transformed with the pDule vector into *E. coli* and the cells were grown the presence (lane 2) and absence (lane 3) of tfmF. Lane 1 shows cells containing only the pDULE vector.
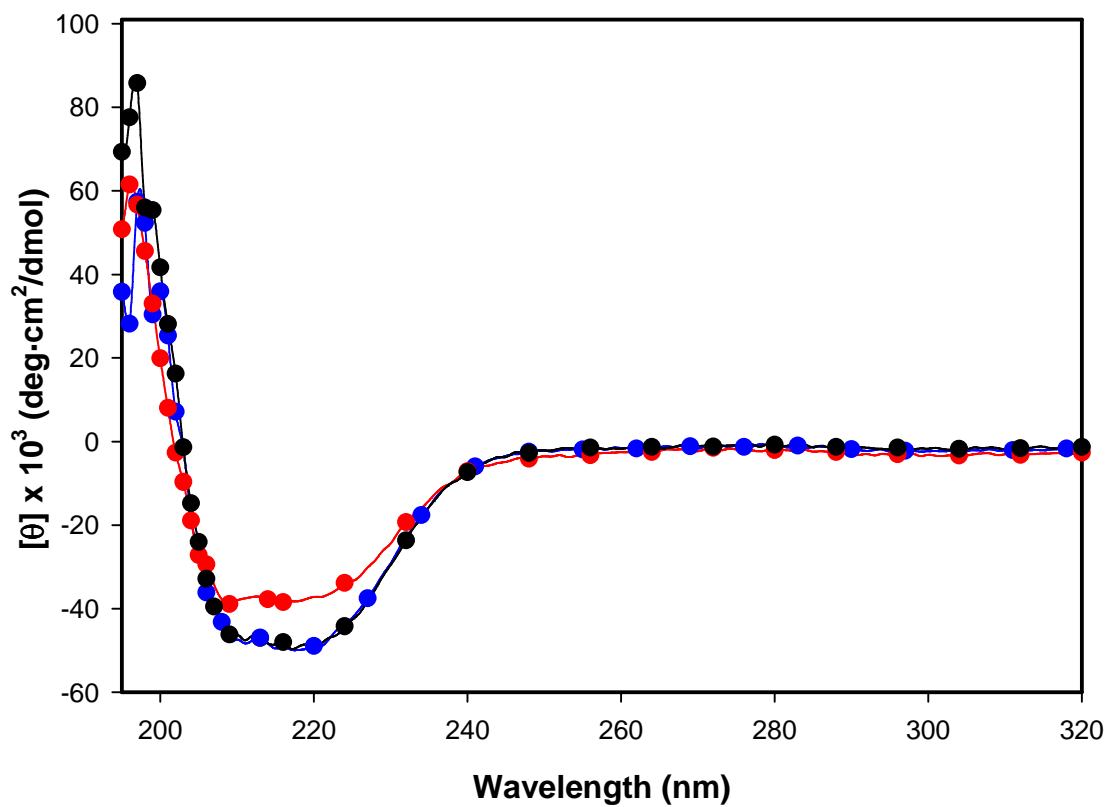
Figure 2-9 CD spectra of wildtype GFP and mutants containing tfmF.
Phe222 GFP (blue), Tyr37 GFP (red), and wt GFP were purified and CD spectra were collected.
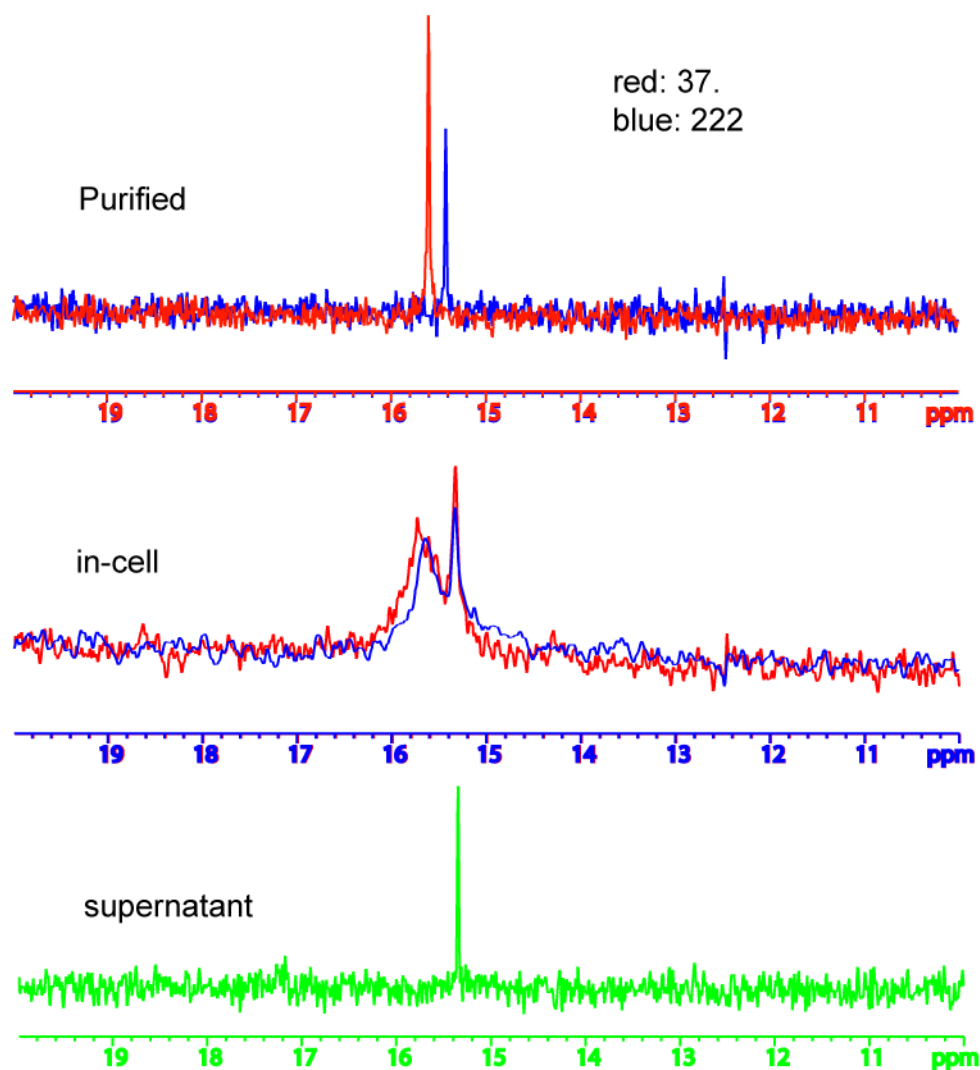
Figure 2-10 ^{19}F NMR of GFP containing tfmF. ^{19}F NMR was performed on *E. coli* cells expressing Phe222 GFP (blue) or Tyr37 (red) GFP mutants containing L-4-trifluoromethylphenylalanine at the mutation site.
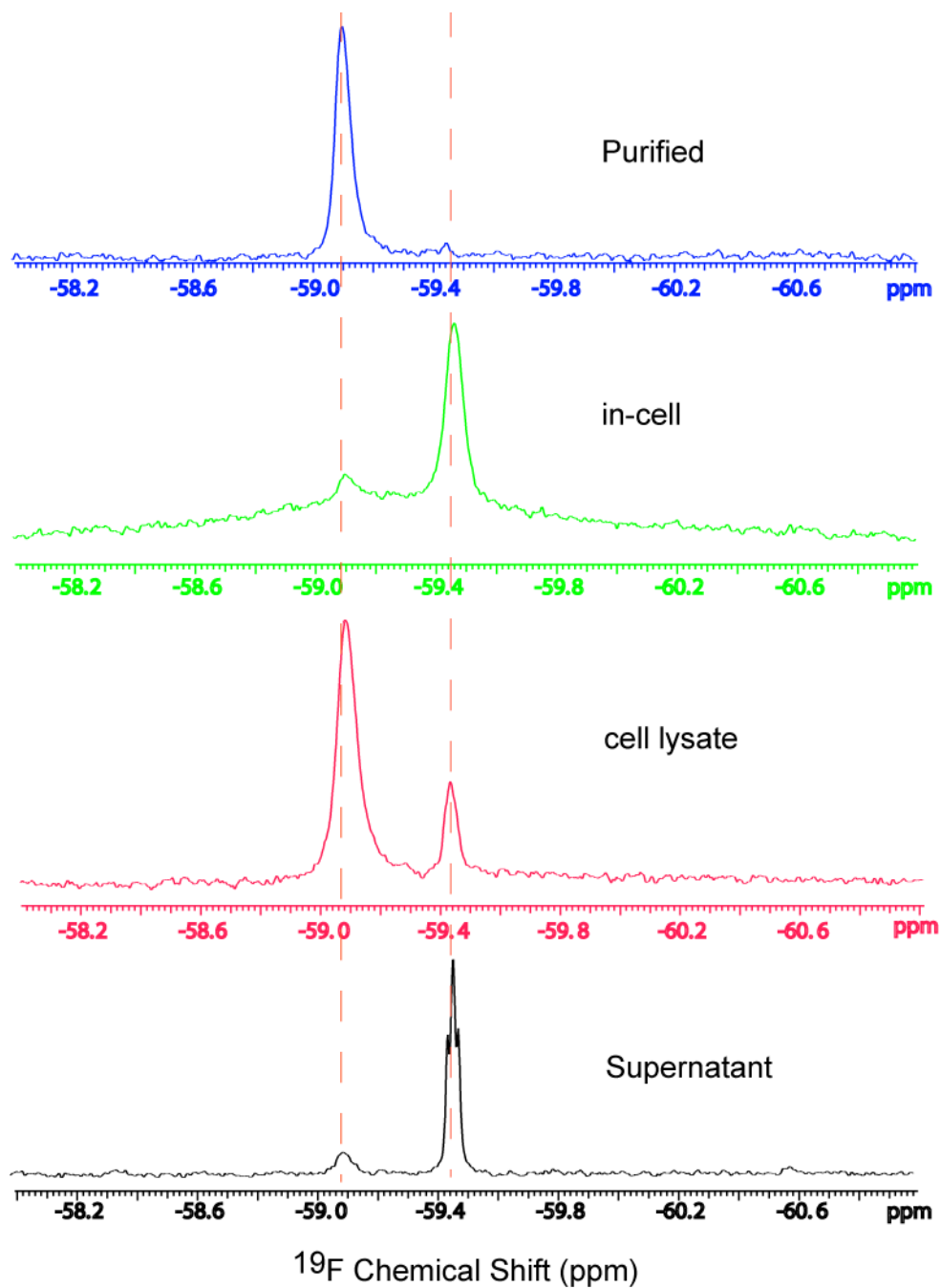
Figure 2-11 ^{19}F NMR of CI2 containing 3-fluorotyrosine. ^{19}F NMR was performed on purified CI2 and on *E. coli* cells expressing CI2. A spectrum of the supernantant was also collected. Finally, a spectrum of the cell lysate was acquired.

48

Chapter 3: Optimizing Expression with a Two-Protein System

Summary

To gain insight about the effects of protein expression on intracellular diffusion, two proteins are co-expressed. The test protein is expressed at varying levels. At the same time, green fluorescent protein is expressed at a constant level, and its diffusion is measured by using FRAP. Chymotrypsin inhibitor 2 was originally chosen as the test protein, but because it leaks from *E. coli* cells, it was quickly replaced by α-synuclein, maltose binding protein, tau-40, and calmodulin. After several attempts with multiple combinations of vectors and promoters, I created a single vector, pGFP/asyn*,* containing green fluorescent protein under the *lac* promoter and the test protein under the *araBAD* promoter. Here, I report the problems with the failed systems and the process of optimizing expression with the pGFP/asyn vector.

3-1 Introduction

As discussed in the last chapter, a two-protein system is necessary for studying the effects of protein expression on intracellular diffusion. When a single protein was used for both controlled expression and diffusion measurements, the signal-to-noise was too low at the low expression levels to

obtain reliable diffusion coefficients. Chymotrypsin inhibitor 2 (CI2) was originally chosen as the additional protein to co-express with GFP because it is stable and well-characterized in terms of its structure, folding, and stability.[66-69] It is a small (7.5 kDa), globular protease inhibitor. After experiments revealed that CI2 leaks from *E. coli* cells,[70] α-synuclein was selected as the new protein for co-expression. Like CI2, α-synuclein is small (14 kDa) and well-characterized.[71]

One way to co-express two proteins is by using the Duet vector designed from Novagen (Fig. 3-1). The Duet vector contains two *T7* promoters, each up stream of multiple cloning sites. This design allows insertion of the desired genes. Unfortunately, the Duet vector does not allow independent control of each gene. Since the promoters are identical, addition of IPTG turns on expression of both proteins.

If a single vector will not provide independent control of two genes, two vectors can be co-transformed. When co-transforming *E. coli* cells, one must consider their compatibility. Compatibility refers to the ability of two vectors to coexist stably within a single cell and is thought to depend on the vectors' origins of replication (ori). Vectors with the same ori compete for replication factors.[72,73] The vectors need to have an independent means of regulating replication if they are to co-exist.[74] However, Velappan *et al.* showed that vectors may be more compatible than the earlier studies suggest.[75] They co-transformed two vectors with the same ori but different antibiotic resistances. While the loss of these vectors from the cell was eventually observed, the process often took several days. Vectors with smaller differences in their rates of replication were able to

coexist for longer periods of time. Thus, factors such as plasmid size and toxicity play a major role in vector compatibility.

Antibiotic resistance is a common tool for ensuring that the vector of interest has been taken up by the *E. coli* cells. The recombinant vector contains a gene that allows the cells to survive in the presence of a specific antibiotic. Cells that do not contain the recombinant vector do not survive. Similarly, when co-transforming two vectors, the vectors should each contain a resistance to a unique antibiotic and the cells should be streaked on a plate containing both antibiotics. In this way, only cells containing both vectors survive.

Antibiotics are also useful in optimizing independent expression of the two proteins. Rifampicin binds bacterial RNA polymerase and blocks the initiation of transcription.[76] It does not, however, inhibit RNA elongation. Thus, any DNA already bound to polymerase when the rifampicin is added, will continue to be translated and then transcribed. Furthermore, since T7 polymerase is not native to *E. coli*, transcription with the T7 promoter is rifampicin resistant.[77] Thus, the use of rifampicin in a two-promoter system may provide an extra level of control for distinguishing the expression level of the two systems. Chloramphenicol also inhibits protein expression, but by another mechanism. It binds to the 50S ribosome, obstructing peptidyl transferase activity, and preventing peptide bond formation.[76] Unlike rifampicin, chloroamphenicol is effective against the T7 promoter. Another antibiotic of importance to these experiments is cephalexin. Cephalexin prevents bacterial cells from dividing by inhibiting the synthesis of their cell walls.[78] Specifically, cephalexin binds

penicillin-binding proteins thereby interfering with transpeptidation, the final step of peptiodoglycan synthesis. Exposing *E. coli* cells to cephalexin results in elongated cells that can reach 10 to 20 μm in length compared to the 2-μm length of an average *E. coli* cell.[79] This effect may be of use for the FRAP experiments.

3-2 Materials and Methods

3.2.1 Vector Creation

All mutations were completed with a site-directed mutagenesis kit (Stratgene) and sequences were confirmed at the UNC-CH genome analysis facility. PCR was performed with pfu turbo polymerase and an Eppendorf Mastercycler®. Primers were obtained from UNC-CH nucleic acids core facility. Restriction enzymes were purchased from NEB. DNA fragments were analyzed by means of 1% (w/v) agrose (Sigma) gels electrophoresed at 150 V for 30 min. Ligations were performed with T4 DNA Ligase (NEB) at 16 °C overnight.

pGFP/CI2 vector

The region of the pBAD-GFP vector (Chapter 2) containing the *araBAD* promoter and GFP gene was PCR amplified with primers 5'AAGCTTCCGTC AATTGTCTGATTCGTTACC 3' and 5' CTACTCCGTCAAGCTTCGTCAATTG 3'. The PCR product was digested with *HindIII*. A *HindIII* site was inserted into the pET-28a(+) vector containing the CI2 gene under the *T7* promoter. The forward primer was 5'GGTATCAGCTCACTCAAAGCTTGTAATACGGTTATCCACAG 3'.

The *HindIII*–digested *araBAD*-GFP PCR product was cloned into pET-28a(+) at its unique *HindIII* site to create the vector: pGFP/CI2.

DUET vector: pBad-asyn

A second *NdeI* site was added to the pRSFDuet-1 vector (Novagen) 2144 base pairs downstream of the first *NdeI* site. The forward primer was 5' CAGTGCAATTTATCTCTTCATATGTAGCACCTGAAGTCAG 3'. *NdeI* digestion of the pBAD-asyn vector gave an insert coding for the araC gene and the *araBAD* promoter followed by the α-synuclein gene. To separate this 2286 base-pair insert from the remaining 2100 base-pair fragment of the vector, an *ScaI* digest was used to cleave only the unwanted fragment and agrose gel electrophoresis was used to separate the fragments. The gel-extracted insert was cloned into pRSFDuet-1 using the *NdeI* sticky ends.

pBAD-asyn Vector

To remove the *EcoRI* restriction site in the human α-synuclein gene of the pT7-7 vector, the GGA codon at amino acid 111 was mutated to with forward primer, 5'GAGCCCCACAGGAAGGTATTCTGGAAGATATGCC 3'. The gene was amplified by using PCR with primers 5' CTTTAAGAAGGAGAATTCATATGGAT GTATTCATG 3' and 5'GAACATCTGTCAGCAGAATTCAAGAAACTGGGAGC 3'. An *EcoRI* site was also inserted into the pBad/His C vector (Invitrogen) at position 310 with forward primer 5'GGCTAACAGGAGGAATTCATGGGGGGTTCTCATC 3'. After removal of the polyhistidine region by *EcoRI* digestion, the α-synuclein PCR

product was ligated into this modified pBad vector. Finally, an *Spe I* site was incorporated into the vector 1548 residues downstream of the α-synuclein gene stop codon, with forward primer 5'CGTGAGTTTTCGTTCCACTAGTCGTCAG ACCCCGTAGAAAAG 3'.

pGFP/asyn Vector

An *Spe I site* was added to the pAcGFP1 vector (Clontech) at position 31 with the forward primer 5'CGCAAACCGCCTCTCCCCACTAGTTGGCCGATT CATTAATG 3'. *Spe I* digestion of the modified pAcGFP1 vector gave an insert coding for the *lac* promoter followed by the GFP gene, which was then cloned into at the *Spe I* site of the pBad-asyn vector. The forward primer, 5'CCGCGAC TCTAGAGAAGCTTGGCTGTTTTG 3' was used to remove an *EcoRI* site, (25 residues downstream of the α-synuclein gene stop codon) yielding the vector: pGFP/asyn.

Incorporating the MBP, calmodulin, and tau genes under the araBAD promoter

To remove the *EcoRI* restriction site in the calmodulin gene, the GAA codon at amino acid 12 was mutated to GAG by using the forward primer 5'CCGAGGAACAGATTGCAGAGTTCAAGGAAGCTTTCTCC 3'. This modified gene, along with the maltose binding protein (MBP) and human tau-40 genes were amplified by PCR with the primers shown in Table 3-1. The *EcoRI*-digested PCR products were cloned into the *EcoRI*-digested pGFP/*a*syn vector to create three new vectors.

3.2.2 Protein Expression

The pAcGFP1 and pET-28a(+) (containing CI2) vectors were co-transformed into BL21-Gold (DE3) cells and streaked onto Luria Broth plates containing 100 µg/mL ampicillin and 60 µg/mL kanamycin ($LB_{AMP/KAN}$). The modified DUET vector containing α-synuclein was co-transformed with pAcGFP1 into BL21-AI competent *E. coli* cells and streaked onto $LB_{AMP/KAN}$ plates. The other vectors were individually transformed into competent *E. coli* BL21-AI cells (Invitrogen) and streaked onto LB_{AMP} (or LB_{KAN} for pGFP/CI2) plates. A 5-mL starter culture of liquid $LB_{AMP/KAN,}$ $LB_{AMP,}$ or LB_{KAN} was inoculated with a single colony and grown overnight at 37 ˚C with constant shaking at 225 rpm. This starter culture was used to inoculate 25 mL of fresh $LB_{AMP/KAN,}$ $LB_{AMP,}$ or LB_{KAN} in a 250-mL flask at a 1:25 dilution.   Once the optical density at 600 nm (OD_{600}) reached between 0.5 to 0.7, the culture was divided into three, 5-mL aliquots, which were induced with varying concentrations of arabinose, ranging from 0.0002% to 20%, and grown at 37 ˚C with constant shaking at 225 rpm. After 3 h, the OD_{600} of each sample was measured, and chloramphenicol, to halt expression, was added to a final concentration of 50 µg/mL. Rifampicin (40 µg/mL) was added to selected samples at 2 h after induction.

3.2.3 Protein Analysis

One-mL aliquots of culture were centrifuged at 10,000 g for 15 min and the pellets were resuspended in 30 uL of 20-mM sodium phosphate buffer (pH 7) and 10 uL of protein loading dye [62.5-mM TRIS, pH 6.8, 50% (v/v) glycerol, 5%

(w/v) SDS, 10% (v/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue]. After centrifuging at 17,000 g for 30 min to remove cellular debris, 15-uL samples were loaded onto 10-to-20% Tris-HCl Criterion Precast SDS-PAGs. The gels were electrophoresed at 200 V for 60 min. Fluorescence emission scans of the gels were collected with the BIO-RAD Versa Doc Imaging System (GE Healthcare). The gels were stained with Comassie Brillant Blue and destained with an aqueous solution of 10% (v/v) methanol and 10% (v/v) acetic acid.

3-3 Results and Discussion

3.3.1 GFP and CI2 Co-expression

From previous attempts, we observed that GFP under the *lac* promoter is expressed at high levels regardless of whether IPTG is added to induce expression (Chapter 2). In contrast, IPTG increases CI2 expression under the *T7* promoter (Fig. 2-6 A). Based on these results, we co-transformed the pAcGFP1 and pET-28a(+) vectors into BL21-Gold (DE3) cells in an attempt to co-express GFP under the *lac* promoter and CI2 under the *T7* promoter. Addition of IPTG should elevate the amount of CI2, while the GFP concentration remains constant. Since the pAcGFP1 vector confers ampicillin resistance and the pET-28a(+) vector containing CI2 confers kanamycin resistance, addition of both antibiotics should ensure that only *E. coli* containing both vectors survive. Thus, this method provides a simple strategy for independent control of GFP and CI2 expression without the need to modify the original vectors.

Contrary to this rational, SDS-PAGs of the cell lysates revealed that CI2 and

56

GFP were never co-expressed (Fig. 3-2). CI2 bands were barely detectable in any of the samples. Fluorescent GFP bands were observed in samples that had been induced with arabinose (lanes 3 and 4), but no fluorescence was detected in the sample that had not been induced. The lysate from cells containing only the pAcGFP1 vector were also run on the gel to reveal the relative amount of GFP expressed in the presence and absence of the second vector (lane 5). The presence of the second vector drastically decreased the GFP expression level, even though CI2 was not expressed to any detectible level in these cells.

The lack of co-expression may be due to the poor compatibility of the two vectors. pAcGFP1 contains the pUC origin, which is a derivative of the BR322 vector contained in the pET-28a(+) vector. The origins of the two vectors are similar enough that they may be competing for essential replication factors. While recent studies suggest that the effects of vector incompatibility have been overestimated, these studies also show that the ability of two vectors to co-exist depends on their relative replication rates.[75,80] The approximate copy number for BR322 is 15-75, whereas pUC is a high copy plasmid, with 300-700 copies per cell.[80] The pUC-containing vector may be rapidly out-competing the other vector. This occurrence would explain why GFP levels are elevated by induction of arabinose, while a CI2 band is barely visible in any sample.

To avoid plasmid incompatibility, a single vector was created that contained the GFP gene under the *araBAD* promoter and the CI2 gene under the *T7* promoter. This vector was originally transformed into BL21-Gold (DE3) cells, but SDS-PAGs of the cell lysates suggested that only CI2 was being expressed.

Although there is evidence in the literature of expression with the *araBAD* promoter in BL21-Gold (DE3) cells,[81] TOP10 cells are typically used to achieve optimal expression with this promoter. When the GFP/CI2 vector was transformed into TOP10 cells, SDS-PAGs of the cell lysates showed that GFP was expressed, but CI2was not. Further investigation revealed that these results were consistent with the specific *E. coli* strains. BL21-Gold (DE3) cells are genetically modified to contain a *T7* RNA polymerase gene under the control of a *lac* promoter, while TOP10 cells do not. Likewise, TOP10 cells are genetically modified with a mutation in L-ribulose-phosphate 4-epimerase (*araD* mutation) so that the cells do not metabolize the arabinose inducer, and BL21-Gold (DE3) cells lack this mutation. Since protein has previously been expressed in BL21-Gold (DE3) with the *araBAD* promoter, we tried increasing the amount of arabinose (up to 20% w/v), anticipating that higher concentrations may allow the arabinose to induce expression before it is metabolized. Regardless of the amount of arabinose added, no GFP expression was observed. Thus, it is necessary to find a strain of *E. coli* that is compatible with both promoters.

Bl21-AI cells are designed to provide controlled expression of a toxic protein under the *T7* promoter, but their mechanism for providing this regulation also allows compatibility with the *araBAD* promoter. Since this system was created such that arabinose induces the *T7* promoter (Chapter 1), BL21-AI cells have been modified so that they cannot metabolize arabinose and so that they will synthesize *T7* RNA polymerase. As anticipated, SDS-PAGs of cell lysates from Bl21-AI cells containing the CI2/GFP vector showed expression of both

58

protein (Fig. 3-3). Furthermore, it was possible to control CI2 expression by increasing the amount of arabinose. Unfortunately, the level of GFP was also varied with this increase (Fig. 3-3B). Further attempts to optimize the system, however, were abandoned because concurrent experiments showed that CI2 was leaking out of the cells.[70]

3.3.2 GFP and α-synuclein co-expression

Due to the observed leakage of CI2, this protein was replaced by α-synuclein as the test protein for the remainder of the experiments. New vectors now needed to be created to co-express GFP and α-synuclein. When used as intended, the DUET vector is non-ideal because both genes are cloned under a *T7* promoter, and the addition of IPTG increases expression of both genes. The RSF-DUET vector is of use, though, because it has an origin of replication that is orthogonal to pUC, the ori of the pBAD-GFP vector. As a result, a DUET vector containing the α-synuclein gene can be co-transformed with the pBAD/GFP vector. After transforming these vectors into BL21-AI cells, SDS-PAGs of the cell lysates showed that the arabinose concentration determined the α-synuclein expression level, while the amount of GFP was relatively constant for the 0.2% and 0.002% arabinose concentrations (Fig 3-4 A, B). GFP was not detectable at lower arabinose concentrations. Unfortunately, these results were not consistently observed (Fig 3-5 C, D). Furthermore, the level of GFP was too low to perform FRAP experiments. Figure 3-5 shows the relative expression of GFP co-expressed with α-synuclein compared to GFP expression alone under both the *lac* and *araBAD* promoters.

Due to the poor reproducibility of the two-vector system, a single vector, pGFP/asyn, was created to co-express α-synuclein and GFP (Fig 3-6). To overcome the low GFP expression levels of the previous system, the α-synuclein gene was placed under the *araBAD* promoter and the GFP gene was placed under the *lac* promoter, which yields high levels of protein expression. Figure 3-7 shows a Coomassie-stained gel of the cell lysate from this system. The intracellular GFP concentration is constant and the α-synuclein expression varies with the arabinose concentration. The experiment was performed in triplicate to demonstrate the reproducibility of the system. Furthermore, adding rifampicin decreased the protein concentrations (Fig. 3-8). This additional level of control is useful for investigating intracellular effects of protein expression.

α-Synuclein is small and disordered. If the effects of protein expression depend on the protein's properties, α-synuclein is one of the least likely proteins to affect GFP diffusion. Rather, one might expect that overexpression of a large, globular protein would have greater influence on diffusion, since such a protein would exclude a greater volume of the cytoplasm to GFP. To determine if the size and shape of the protein affect whether its expression influences diffusion, GFP was co-expressed with one of four proteins. MBP (42 kDa) and bovine calmodulin (17 kDa) are globular, where as human tau-40 (45 kDa) and human α-synuclein (14 kDa) are disordered. Furthermore, calmodulin and α-synuclein are small, while MBP and tau-40 are much larger proteins. To co-express each of these test proteins with GFP, the α-synuclein/GFP vector was modified by replacing the α-synuclein gene with either the MBP, calmodulin, or tau-40 gene.

Figure 3-9 shows typical Coomassie-stained gels of the lysates from BL21-AI cells containing each of these vectors. For a given system, GFP is expressed at a constant level. The amount of co-expressed protein is controlled by the amount of arabinose used for induction. In the absence of arabinose, GFP is expressed, but the co-expressed protein is undetectable.

It was necessary to ensure that the intracellular protein concentrations observed on the SDS-PAGs did not change during the timescale of the FRAP experiments. Chloramphenicol was added to stop protein synthesis at 3 h after induction, and the aliquots of the culture were collected 3 h and 5 h after induction. SDS-PAGs from these cell lysates were identical, suggesting that the protein concentration of the cells exposed to chloramphenicol remains constant throughout the FRAP experiments. Finally, in anticipation of using cephalexin to elongate cells for the FRAP experiments, the effects of this antibiotic on protein expression needed to be explored. SDS-PAGs from the lysates of cells grown in the presence and absence of cephalexin are identical. This result confirms that cephalexin does not affect the protein concentrations.

3-4 Conclusion

With a single vector, we were able to co-express GFP with each of four proteins: MBP, calmodulin, tau-40, or α-synuclein. GFP was placed under the *lac* promoter, while the other gene was controlled by the *araBAD* promoter. In this way, GFP expression could be held constant while varying the expression level of the other protein.

3.6 Tables

Table 3-1. Primers for PCR amplification

Gene	Vector	Primer
Calmodulin	pT7-7	5' AAGGGGTTATGCTAGTTATTGCTCAG 3'
		5' GTTTAACTTTAAGAATTCGATATACCATGGC 3'
MBP	pMAL-c2x	5' CACGAGGAATTCACCAACAAG 3'
		5' GATGCCTGGCAGTTTATG 3'
Tau-40	pET-CaM	5' GCAGGGTTTGTAAGAATTCAAGCTTCTCGAGG 3'
		5' CACAACGGTTTCCCTCTAGAATTCATTTTGTTTAACTTTAAGAAG 3'

3.6 Figures
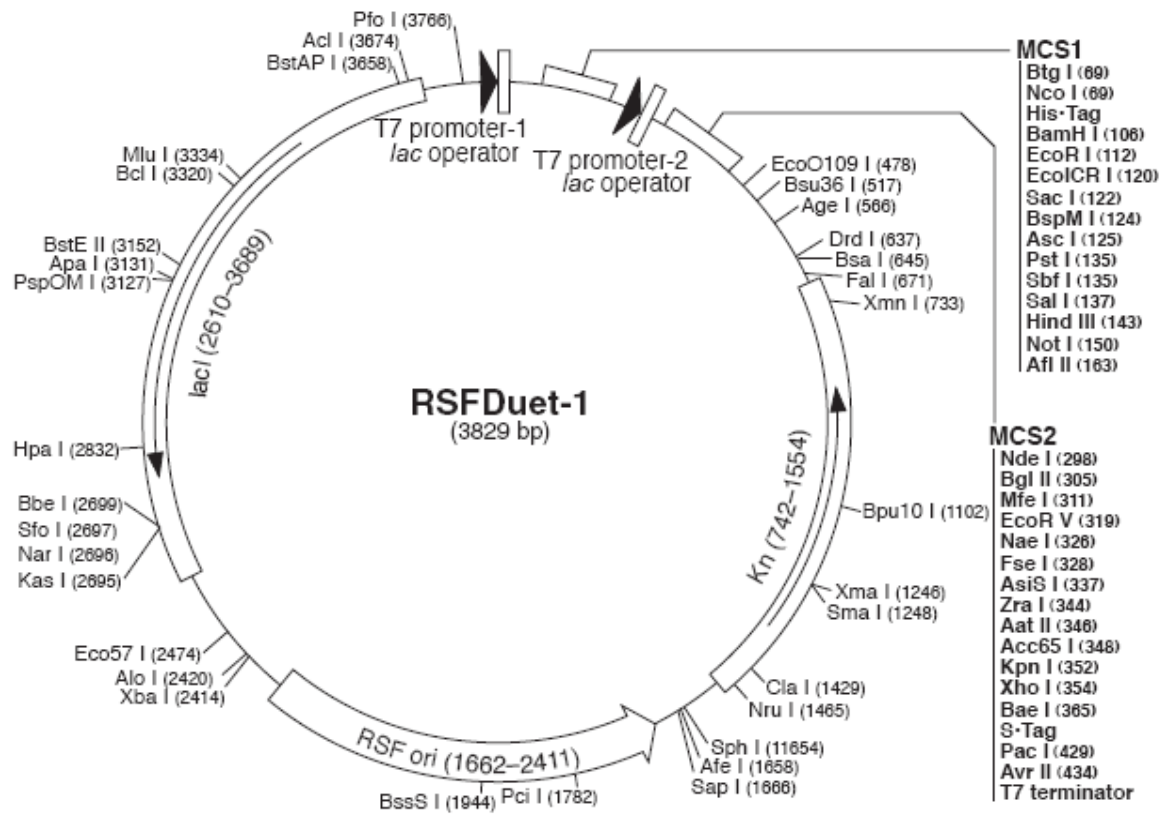
Figure 3-1 RSFDUET-1 vector. This vector allows the co-expression of two genes, each controlled by a *T7* promoter.
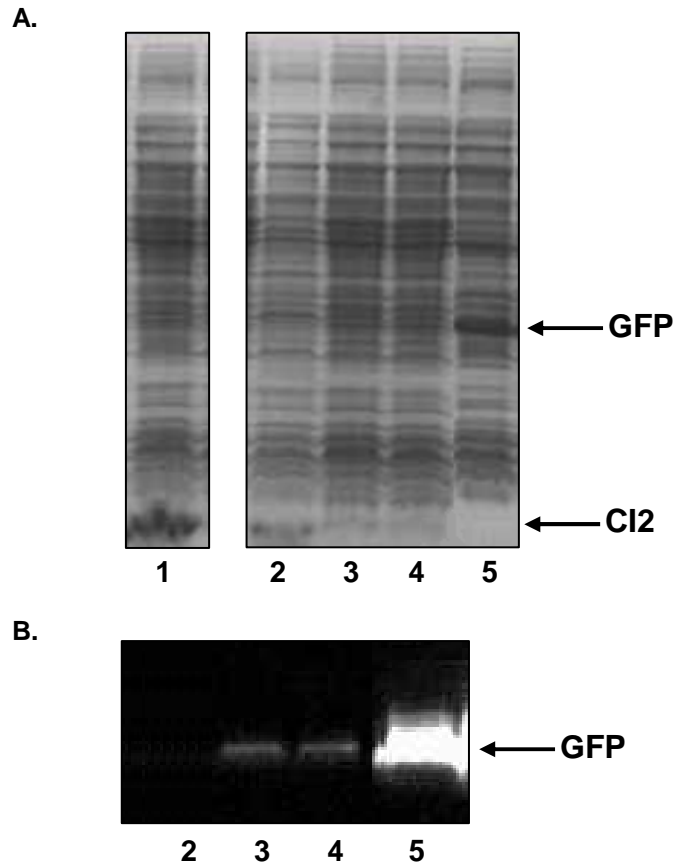
Figure 3-2 Co-expression of GFP and CI2 under the *lac* and *T7* promoters respectively. The pAcGFP1 and pET-28a(+) vectors were co-transformed into BL21-Gold (DE3) *E. coli* cells and induced with 0% (lane 2), 0.002% (lane 3), or 0.2% (lane 4) arabinose. (A) Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized with Coomassie staining. For comparison, lysates from cells expressing only CI2 (lane 1) or only GFP (lane 5) are also shown (B) The same gel was visualized by fluorescent scanning.

A.

GFP

CI2

Arabinose .0002% .002% .2%

B.

GFP

Arabinose .0002% .002% .2%
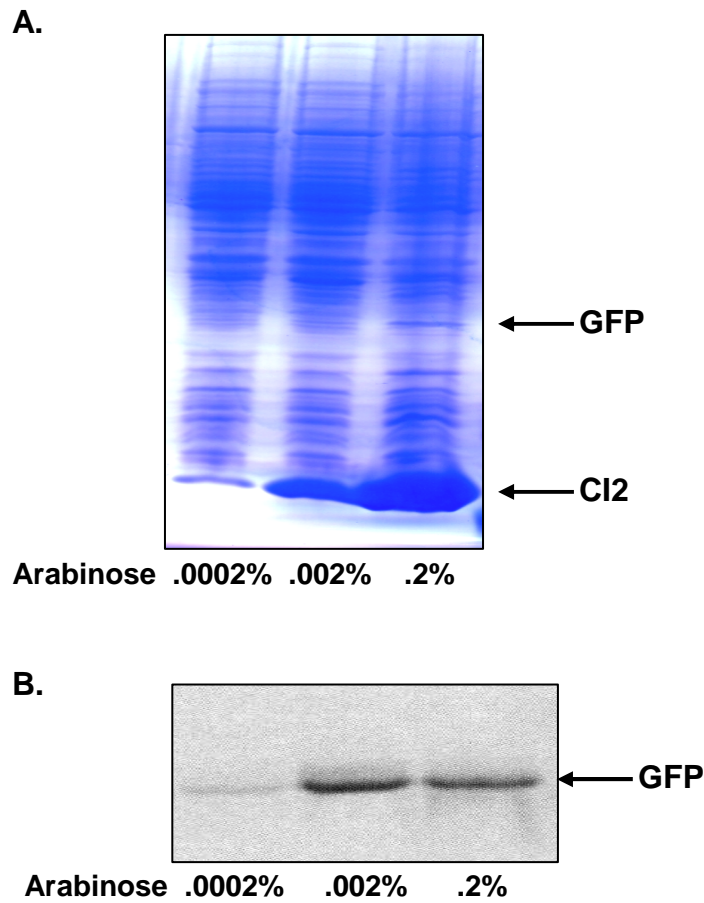
Figure 3-3 Co-expression of GFP and CI2 in *E. coli* using a single recombinant vector. The GFP gene was under the *pBAD* and the CI2 gene was controlled by the *T7* promoter. (A) Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized with Coomassie staining. (B) GFP can be visualized with a fluorescence scan of the same gel.
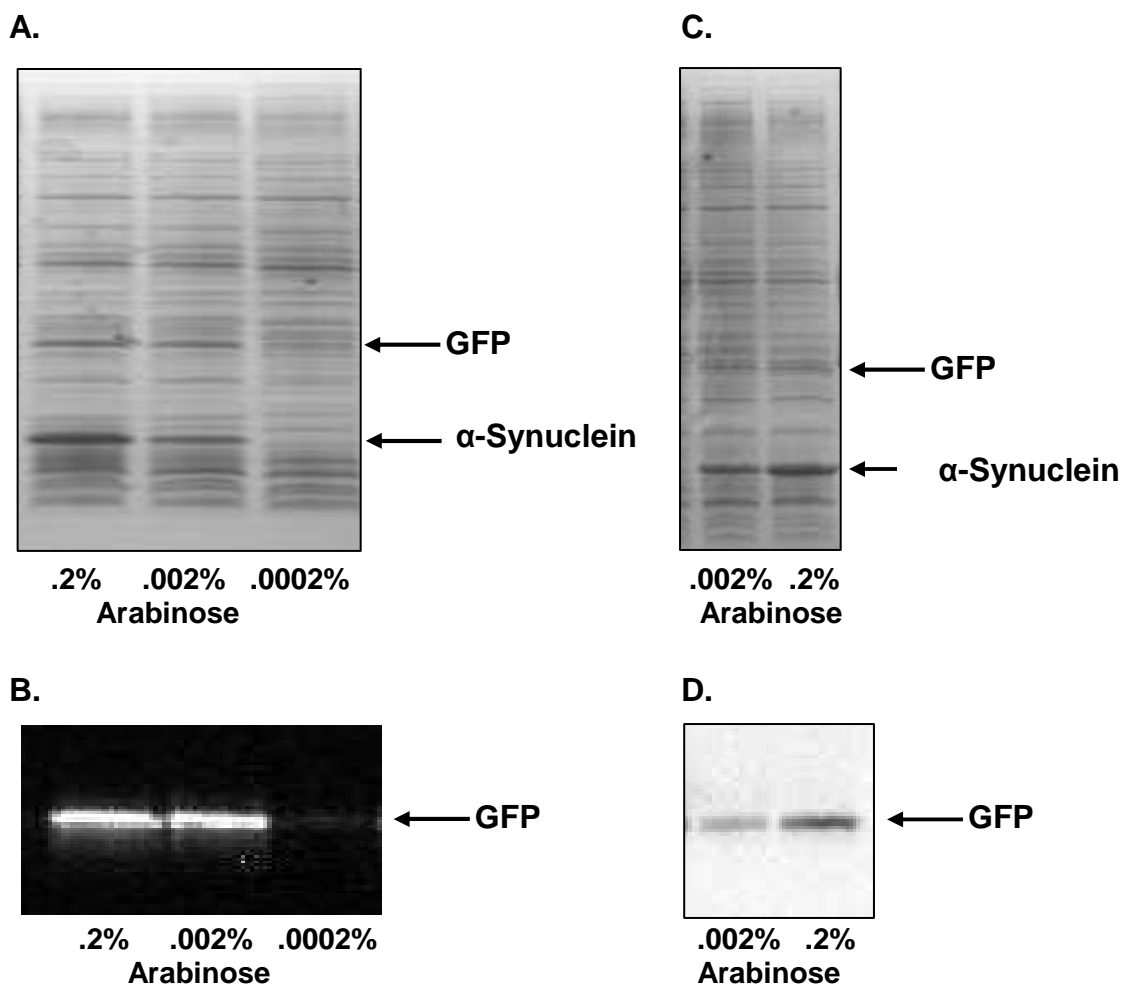
Figure 3-4 Co-expression of GFP and α-synuclein in *E. coli* under the *pBAD* and *T7* promoters respectively. (Panels A and C) Cell lysates were separated on 10%-to-20% gradient SDS-PAGs and visualized with Coomassie staining. (Panels B and D) GFP can be visualized with a fluorescence scan of the same gels.
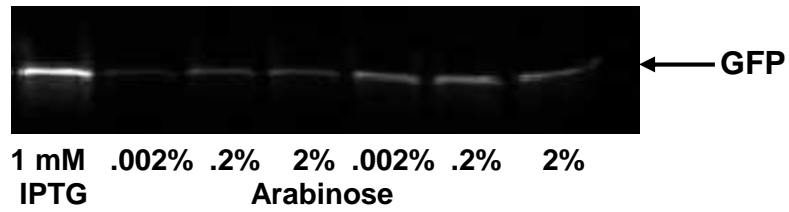
Figure 3-5 Co-expression of GFP and α-synuclein in *E. coli* yields lower GFP levels than when GFP is expressed alone. GFP was expressed under the *lac* promoter (lane 1) and under the *T7* promoter (lanes 2-7) in the presence (lanes 2-4) or absence (lanes 1, 5-7) of a second vector containing the α-synuclein gene. Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized by fluorescence.
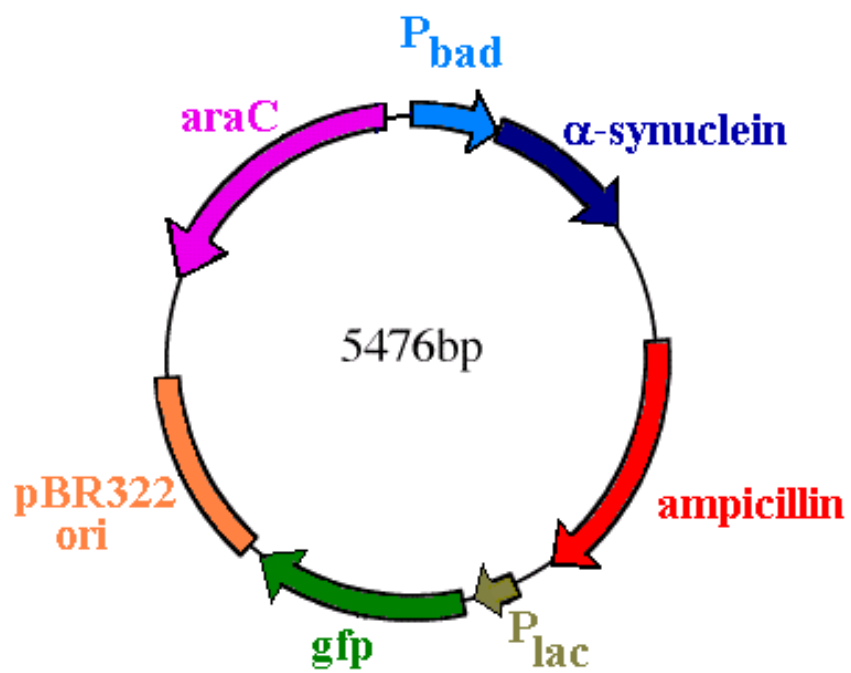
Figure 3-6 pGFP/asyn vector designed to co-express α-synuclein and GFP.
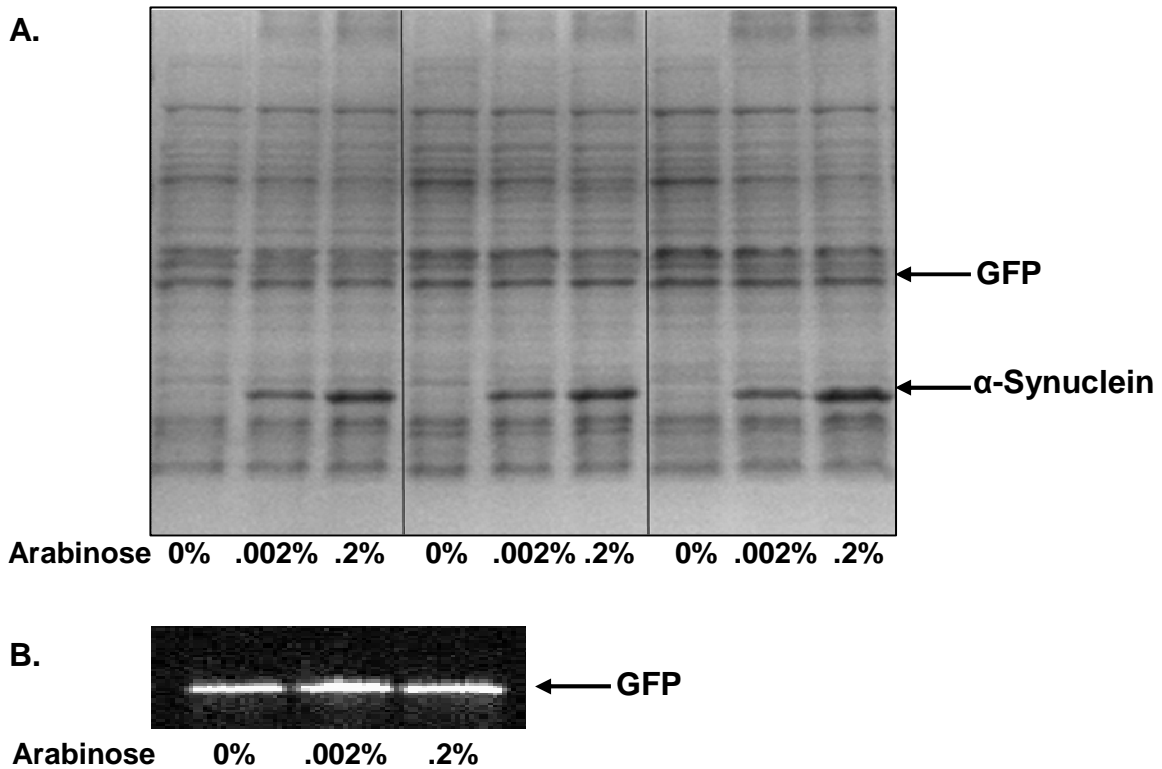
Figure 3-7 Co-expression of GFP and α-synuclein in *E. coli* under the *lac* and *pBAD* promoters, respectively. (A) Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized with Coomassie staining. The experiment was performed in triplicate to demonstrate reproducibility. (B) GFP can be visualized with a fluorescence scan of the same gel.

A.

	1	2	3	4	5	6	7	8
Arabinose	0%	.002%	.02%	0%	.002%	.02%	0%	.02%
Rifampicin	-	-	-	-	+	+	+	+

← GFP

← α-Synuclein

B.

	1	2	3	4	5	6	7	8
Rifampicin	-	-	-	-	+	+	+	+

← GFP
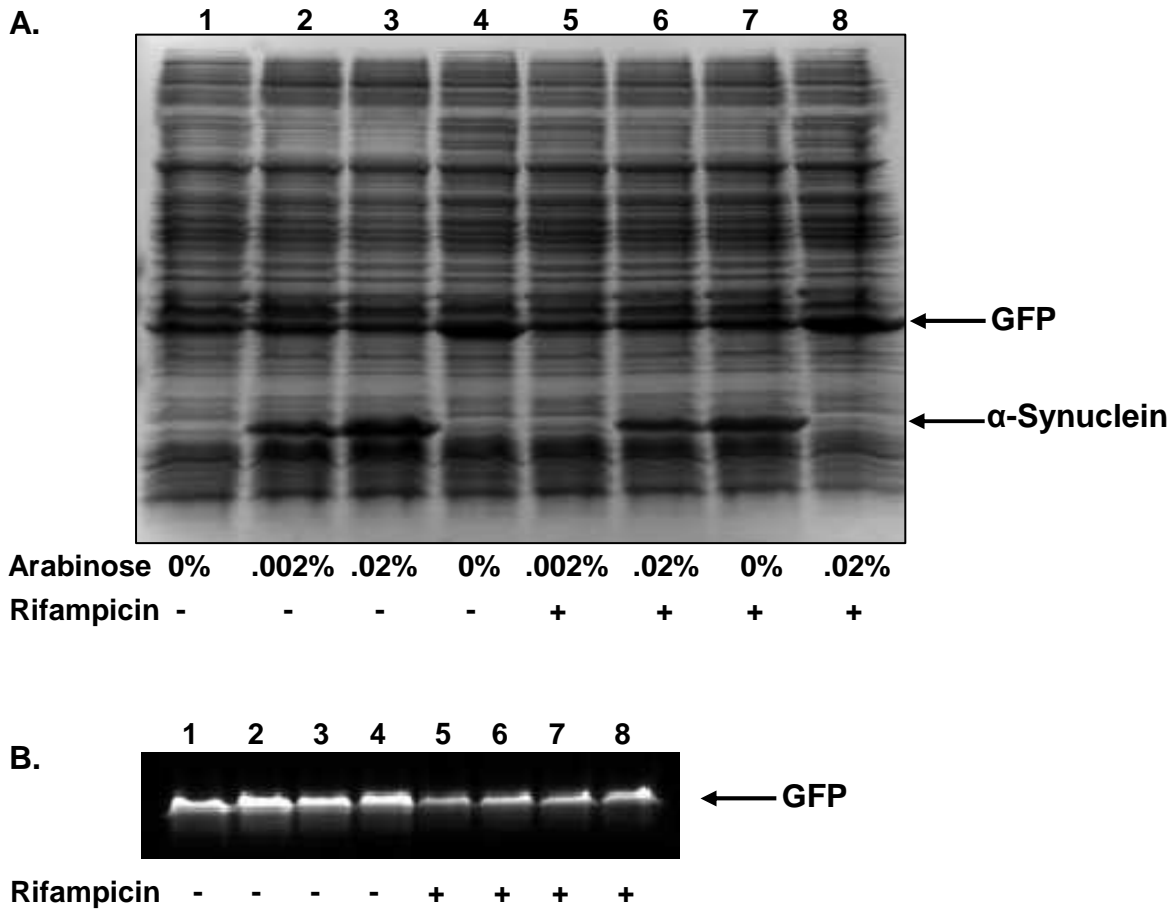
Figure 3-8 Rifampicin decreases protein expression. (A) GFP and α-synuclein were expressed in *E. coli* under the *lac* and *pBAD* promoters, respectively, both in the presence (lanes 1-3) and absence (lanes 5-7) of rifampicin. GFP under the *lac* promoter (pAcGFP1 vector) was also expressed in the presence (lane 4) and absence (lane 8) of rifampicin. Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized with Coomassie staining. (B) GFP can be visualized with a fluorescence scan of the same gel.
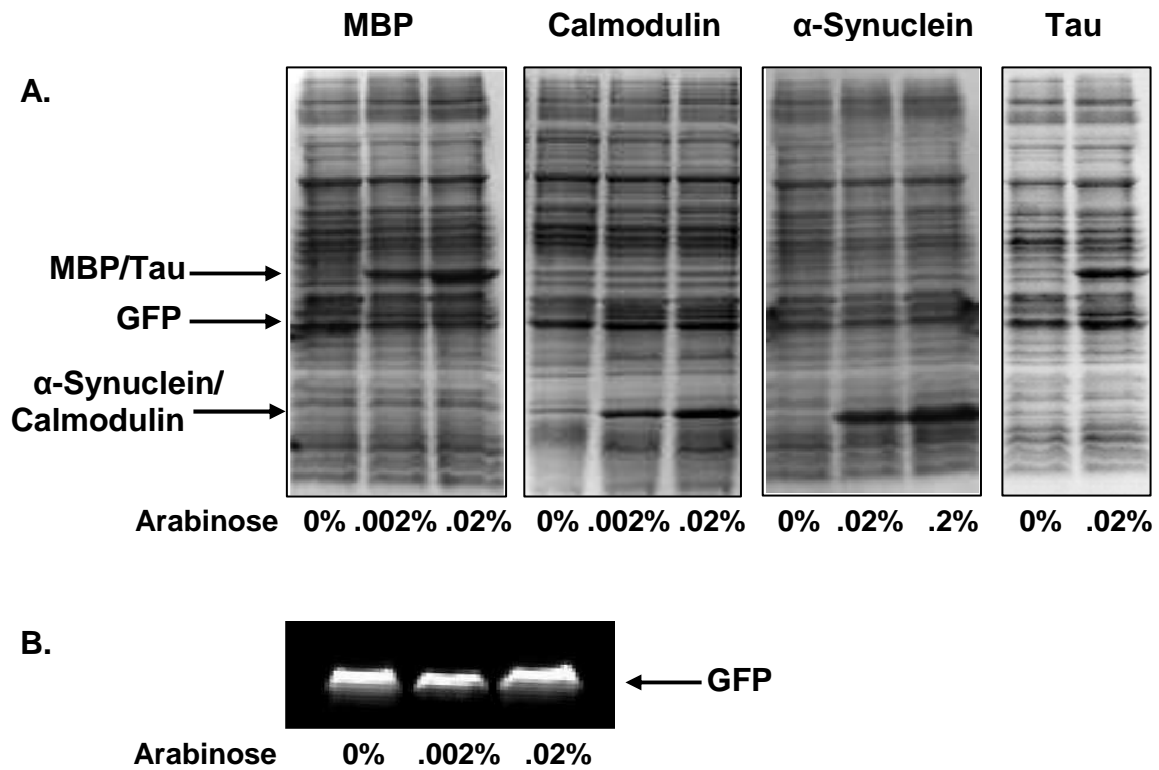
**Figure 3-9 Intracellular GFP concentration is independent of the amount of an individual co-expressed protein.** (A) Cell lysates were separated on a 10%-to-20% gradient SDS-PAG and visualized with Coomassie staining. The arabinose concentrations were adjusted for each system to maximize protein expression with the least amount of inducer. (B) GFP can be visualized by using fluorescence, as shown by these lysates from cells co-expressing GFP and MBP. Similar results were observed for the other three systems.

Chapter 4: Effects of Recombinant Protein Expression on Green Fluorescent Protein Diffusion in *Escherichia coli*

Summary

Fluorescence recovery after photobleaching was used to measure the diffusion coefficients of green fluorescent protein (GFP, 27 kDa) in *Escherichia coli* in the presence or absence of four co-expressed proteins: cytoplasmic maltose binding protein (42 kDa), tau-40 (45 kDa), α-synuclein (14 kDa), or calmodulin (17 kDa). Regardless of the type or amount of protein that was co-expressed, the GFP translational diffusion coefficient remains constant. We conclude that expression of these soluble proteins has little to no effect on the translational diffusion of GFP. These results have implications for the utility of in-cell nuclear magnetic resonance spectroscopy.

4-1 Introduction

While FRAP has been applied to a range of biological systems, using the technique in *E. coli* cells presents a challenge. Their small size complicates the data analysis since the assumption of an infinite supply of fluorescent markers is

no longer valid. Furthermore, it necessitates the need for a smaller bleached region. This requirement coupled with the faster recovery time observed in solution compared to conventional membrane experiments results in the need for a detector with short collection times, such as a video-rate camera.

Despite these complications, Elowitz and coworkers were able to measure the diffusion of GFP in *E. coli* by means of FRAP.[37] To account for the depletion of fluorescent markers at the edges of the cells, they included the boundary conditions $\frac{\partial C(0,t)}{\partial x} = \frac{\partial\, C(L,t)}{\partial x} = 0$, where L is the length of the cell and C(x, t) is the concentration at position x and time t. Applying these conditions to Fick's second law of diffusion yields the general solution:

$$C(x,t) = \sum_{n=0}^{\infty} A_n(t)\cos\left(\frac{n\pi x}{L}\right) \tag{4-1}$$

where D is the diffusion coefficient, n =1,2,3,4..., and $A_n(t)$ is the Fourier amplitude given by:

$$A_n(t) = A_n \exp\left(-Dt\left(\frac{-n\pi}{L}\right)^2\right) \tag{4-2}$$

Since the fluorescence intensity, I(x,t), is proportional to concentration, the Fourier amplitude can be determined by using Eq. 4-1 as:

$$A_n(t) = \frac{2}{L}\int_0^L \cos\left(\frac{n\pi x}{L}\right)I(x,t)dx \tag{4-3}$$

 Thus, to obtain diffusion coefficients, $A_n(t)$ is calculated by using the intensity values from the FRAP experiments in Eq. 4-3. The resulting Fourier amplitude versus time curve is fitted to the exponential $ae^{-bt} + c$, where $b = D(^{n\pi}/_L)^2$

73

according to Eq. 4-2. It is only necessary to compute $A_1(t)$ since the larger-numbered n-terms decay too fast to be measured by the detector.

This work paved the way for diffusion studies involving the effects of osmotic shock and crowding.[82,83] Of specific interest, van den Bogaart *et al.* reported no correlation between the GFP diffusion coefficient and the fluorescence intensity of *E. coli* cells (which was assumed to reflect protein expression levels).[84] This result is seemingly inconsistent with the observation by Elowitz *et al.* that increasing the inducer concentration significantly reduces the diffusion coefficient of GFP.[37] Since neither study directly quantified protein concentration, the discrepancy may involve a difference in expression levels.

Here, we use GFP (27 kDa) as a tracer molecule and measure its intracellular diffusion in the presence and absence of four different proteins expressed at varying levels in *E. coli*. The four proteins include two globular proteins, maltose binding protein (MBP, 42 kDa) and bovine calmodulin (17 kDa), and two disordered proteins, human tau-40 (45 kDa) and human α-synuclein (14 kDa). MBP is normally found in the periplasm, but we use a version that is expressed in the cytosol.[85] Calmodulin is a calcium binding protein found in eukaryotes that regulates numerous enzymes.[86] α-Synuclein is associated with Parkinson's disease.[87] When expressed in *E. coli*, this protein has been reported to be found exclusively in the periplasm.[88] We, however, find it in both the cytosol and the periplasm (*vide infra*). Tau-40 is a microtubule-associated protein commonly found in neurons. It is highly soluble, but, when misfolded, it can form aggregates that contribute to neurodegenerative disorders such as Alzheimer's

disease.[89] Together these four model proteins provide a range of size and structure for probing the effect of protein expression on intracellular dynamics.

4-2 Materials and Methods

4.2.1 Protein Expression

Four vectors derived from AcGFP1 (Clontech) and pBAD/HIS C (Invitrogen) were created to contain the gene for a nondimerizable GFP[90] under the *lac* promoter and the gene for either α-synuclein, tau-40, MBP, or calmodulin under the control of the *araBAD* promoter (Chapter 3). These vectors were individually transformed into competent *E. coli* BL21-AI cells (Invitrogen) and plated on LB_{AMP} plates. A 5-mL starter culture of liquid LB_{AMP} was inoculated with a single colony and grown overnight at 37˚C with constant shaking at 225 rpm. This starter culture was used to inoculate 25 mL of fresh LB_{AMP} in a 250-mL flask at a 1:25 dilution.   Once the OD_{600} reached 0.5 to 0.7, the culture was divided into three, 5-mL aliquots, which were induced with arabinose according to Fig. 3-9 and grown at 37˚C with constant shaking at 225 rpm. After 3 h, the OD_{600} of each sample was measured and chloramphenicol, to halt expression, was added to a final concentration of 50 µg/mL.

4.2.2 Protein Purification

pAcGFP1, pT7-7 containing tau-40, pMAL-c2x, pET-CAM1, and pT7-7 containing α-synuclein were individually transformed into *E. coli* BL21-Gold (DE3) competent cells (Stratagene) and plated on Luria Broth plates containing 100-

μg/mL ampicillin (LB$_{AMP}$). For calmodulin samples, 60-μg/mL kanamycin was used in place of ampicillin. Starter cultures of liquid LB$_{AMP}$ were inoculated with a single colony and grown overnight at 37°C with constant shaking at 225 rpm. These starter cultures were used to inoculate (1:25 dilution) 1-L of fresh LB$_{AMP}$ in 6-L flasks. The cultures were induced when the OD$_{600\ nm}$ reached 0.8, with a final concentration of 1-mM IPTG and allowed to grow at 37°C with constant shaking at 225 rpm for 4 h.

The cultures were harvested by centrifugation (Sorvall RC-3B rotor) at 1600 g for 30 min at 4°C. The pellets were resuspended in lysis buffer [20-mM Tris (pH 8.0), 1-mM EDTA, 1-mM PMSF or 20-mM Tris (pH 7.4), 200-mM NaCl, 1-mM EDTA, 10-mM 2-mercaptoethanol for MBP] and pulse sonicated (Branson Ultrasonics) at 4°C (18% amplitude) for 10 min. Cell debris were removed by centrifugation (Sorvall RC-5B with a SS-34 rotor) at 27000 g. The calmodulin sample was further purified by using a phenyl-sepharose column (GE Healthcare) as previously described.[91]

The MBP sample was added to 15 mL of amylase resin (New England Biolabs), agitated at 16°C for 2 h, and rinsed three times with MBP buffer. MBP was eluted with MBP buffer containing 10-mM maltose.

The GFP sample was further purified with anion exchange chromatography (HiLoad 16/10 Q Sepharose, AKTA FPLC UPC-900) by using a linear gradient from 0-M to 1-M NaCl in 20-mM Tris buffer (pH 7.5). The fractions containing GFP were concentrated in an Amicon Ultra MWCO 3,000 centrifugal filter unit (Millipore) and dialyzed against water. The protein was further purified

by using size exclusion chromatography (16/60 Superdex^{TM} 75) with water as an eluant.

The tau-40 and α-synuclein samples were boiled for 20 min and centrifuged at 27000 g for 30 min. The supernanant was stirred with streptomycin sulfate (10 mg/mL) at 4°C for 30 min and then centrifuged. The α-synuclein supernatant was further purified as described[92] except the freeze-thaw step was eliminated. The tau-40 supernatant was stirred with $(NH_4)_2SO_4$ (180 mg/mL) at 4°C for 30 min and centrifuged. The resulting supernatant was dialyzed overnight against water. The purity of each protein was confirmed by using SDS-PAGE (18%) with Coomassie Brilliant Blue staining.

4.2.3 Determining Protein Concentration

For each of the four cultures, 1-mL aliquots were collected 0 and 2 h after adding chloramphenicol. These aliquots were centrifuged for 10 min at 8000 g (Eppendorf model 5418). The pellets were resuspended in 20-mM potassium phosphate buffer (pH 7.5). Protein standards were obtained with the purification protocols described above. The concentration of each standard was determined by absorbance [GFP, $\varepsilon_{475\ nm}$ = 32,500 cm^{-1}M^{-1} as reported by the manufacturer, Clontech; MBP [93], $\varepsilon_{280\ nm}$ = 69,000 cm^{-1}M^{-1}; calmodulin,[94] $\varepsilon_{276\ nm}$ = 3030 cm^{-1}M^{-1}] and/or with the Lowry [95] method (Modified Lowry Protein Assay Kit, Pierce) using cytochrome c (for α-synuclein and calmodulin) or ovalbumin (for tau-40 and MBP) as a reference. The proteins in cell lysates and standards were resolved by electrophoresis on 10-to-20% gradient sodium dodecyl sulfate polyacrylamide

gels (SDS-PAGs, Criterion, Bio-Rad) for 60 min at 200 V. Gels were analyzed by both fluorescence and Coomassie staining with a VersaDoc MP imager (Bio-Rad). The exposure time was optimized to prevent pixel saturation, and Quantity-One software (BioRad) was used to quantify the band intensities. Calibration curves of the standards were used to determine the amount of the proteins of interest in the cell lysates. The cell densities of the cultures were determined from the OD_{600}.[96] These values were confirmed by dilution plating. The intracellular protein concentrations, C (mM), shown in Table 4-1 were ca.culationed by using the equation:

$$C = \left(\frac{n}{NV_{cell}} \right) \left(\frac{V_P}{V_{load}} \right)$$

where n (mmol) is the amount of protein contained in the band on the SDS-PAG (based on the standard curve made from pure proteins); V_p (µL) is the volume of resuspended pellet (in phosphate buffer with SDS-PAG loading dye); V_{load} (µL) is the volume of resuspended pellet loaded into the SDS-PAG; N is the number of cells in the 1-mL aliquot (determined by the OD_{600}); and V_{cell} is the volume of an *E. coli* cell (1×10^{-15} L). [79] Measurements were performed in triplicate.

4.2.4 Cell Fractionation

GFP was co-expressed with either MBP, tau-40, calmodulin, or α-synuclein as described in the Materials and Methods. For each of the four cultures, five 1-mL aliquots were centrifuged for 10 min at 8000 g. Cell pellets were then exposed to either osmotic shock, osmotic shock and lysozyme, chloroform, or TritonX detergent as described.[88,97-100] The pellets and supernants

were resolved by using electrophoresis on 10-to-20% gradient SDS-PAGs (Criterion, Bio-Rad) for 60 min at 200 V.

4.2.5 Sample Preparation

Cover glasses (22 x 22 x 0.17 mm, #1.5, Zefon) and glass microscope slides (3" x 1" x 1 mm, Fisher Scientific) were boiled in ICN detergent (MP Biomedicals) for 10 min, bath-sonicated for 30 min, rinsed thoroughly with deionized water, and dried overnight at 160°C. The dried slides were cleaned in an Ar-ion plasma cleaner (PDC-3XG, Harrick Scientific) for 15 min at 25°C immediately prior to use. The cover glass slides were pretreated with a 0.01% (w/v) poly-L-lysine solution (Sigma-Aldrich) for 15 min, rinsed with minimal media [7.6-mM $(NH_4)_2SO_4$, 60-mM K_2HPO_4, 2-mM $MgSO_4$, 20-μM $FeSO_4$, 1-mM EDTA (pH 6.8)], and attached to a microscope slide with double-sided tape (Part No. 021200-64988, 3M Corp) to form a sandwich. For osmotic stress and urea measurements, the minimal media rinse contained either 250-mM sorbitol (390 mOsm, measured with a Vapro osmometer) or 500-mM urea. The cultures, prepared as described above, were injected into the sandwiches. After incubating for 30 min at 25°C, the sample chamber was rinsed with minimal media and sealed with vacuum grease.

4.2.6 Microscopy

The intracellular GFP diffusion coefficient was measured by using FRAP.[37] Single cell images were recorded with a Zeiss 510 Meta scanning

confocal inverted microscope equipped with a 30-mW argon laser. The sample

was imaged with a 63x, 1.4 NA plan-apochromat oil-immersion objective and a

pinhole of 2.0 Airy-disc units. An excitation wavelength of 488 nm was selected.

A 505-nm long pass filter was used for detection. Cells were oriented in the x

direction, and the laser bleach spot was moved to one pole of the cell. For

photobleaching, the laser intensity was ~240 µW, with a ~0.2 µW observation

beam. After collecting 99 pre-bleach images, the region selected for bleaching

was scanned for 52 ms, and 500 postbleach images of 128 by 16 pixels (7.3 by

0.9 µm) were recorded at a rate of 13.1 ms/frame.

4.2.7 Data Analysis

Time series of fluorescence images were analyzed as described.[83] In

short, images were converted to one-dimensional intensity profiles, I(x,t), by

averaging columns of pixel intensities (perpendicular to the long cell axis) as a

function of distance from the cell edge (x). The one-dimensional profiles were

used to calculate Fourier amplitudes $A_n(t)$:

$$A_n(t) = \frac{2}{L}\int_0^L \cos(q_n x)I(x,t)dx \qquad\qquad (4\text{-}4)$$

where $q_n = n\pi/L$, t is time, L is the cell length in µm, and n=1, because the larger-

numbered n-terms decay too quickly to be measured. The Fourier amplitudes

were plotted as a function of time and fit to a single exponential decay, $ae^{-bt} + c$,

with a, b, and c as free parameters. Diffusion coefficients (D) were obtained from

the equation $b = Dq_1^2$. Analysis was performed by using Mathematica (Wolfram).

4-3 Results

4.3.1 Intracellular Protein Concentrations

Four vectors were created. The GFP gene was placed under the *lac* promoter.

The ara*BAD* promoter was used to control the expression of α-synuclein, MBP,

tau-40, or calmodulin. SDS-PAG electrophoresis was used to quantify the

concentration of a given protein in cells. The results are summarized in Table 1.

Comparing the induced and uninduced GFP concentrations shows that the value

is the same whether or not another protein is co-expressed. The level of GFP

expression varies, however, from system to system. The most extreme variation

is a 5-fold increase in GFP expression in the MBP system compared to the α-

synuclein system. With the exception of tau-40, the GFP concentration appears

to be inversely related to the concentration of the co-expressed protein.

4.3.2 Protein Location

Four fractionation methods were used to determine the locations of the co-

expressed proteins: osmotic shock, [88,99] exposure to detergent, [98] osmotic shock

plus lysozyme, [97] or chloroform treatment.[100] SDS-PAGs of the supernatants

(periplasmic proteins) and pellets (cytosolic proteins) confirm that MBP and tau-

40 are present exclusively in the cytoplasm (Fig. 4-1). The supernatant from the

chloroform fractionation contained calmodulin, yet results from the other three

methods suggest that calmodulin is exclusively localized to the cytoplasm. All

fractionation methods show the presence of α-synuclein in both the cytosol and

periplasm. The percent of cytosolic α-synuclein ranged from 40% (chloroform) to 70% (lysozyme), depending on the method.

4.3.3 Protein Expression

SDS-PAG electrophoresis of cell lysates provides bulk information about expression. To examine cell-to-cell variation, it is necessary to observe individual cells. The fluorescence intensity is constant from cell to cell in *E. coli* co-expressing GFP and α-synuclein (Fig. 4-2A). Similar images were obtained for the calmodulin and MBP systems (not shown). In contrast, the images of *E. coli* co-expressing tau-40 (Fig. 4-2B) show bright and dim cells, suggesting that the GFP concentration varies from cell to cell. Inexplicitly, this intensity variation for the tau-40 system is observed whether or not the tau-40 expressing cells are induced.

Since the levels of the **co-expressed** protein cannot be visualized in cells, GFP was expressed under the ara*BAD* promoter to assess the cell-to-cell expression-level variation for this promoter. The fluorescent intensity is constant for *E. coli* expressing GFP under *araBAD* at 0.02% arabinose (Fig. 4-2C). Histograms of the intensity versus cell number made from the data in Figure 4-2 indicate a slightly wider spread of fluorescence intensities for GFP under the *araBAD* promoter (Fig 4-2C) compared to GFP under the *lac* promoter (Fig 4-2A), but both appear to be normally distributed. This result suggests that, above the arabinose saturation concentration, *araBAD*-driven expression of GFP is similar from cell to cell. By inference, this result implies that the other proteins

driven by this promoter should also show constant expression from cell to cell. The cells in Fig. 4-2C are dimmer than the cells in Fig. 4-2A because the *araBAD* promoter is weaker than the l*ac* promoter for GFP. For all images, the distribution of GFP within a single cell is uniform within optical resolution, suggesting that GFP is neither aggregated nor localized to specific regions, under the conditions examined here.

4.3.4 Cephalexin-Treated Cells

To help with the challenge of performing FRAP experiments on small structures, the cells were treated with cephalexin to elongate them. According to Elowitz *et al.*, the treatment does not affect intracellular protein diffusion as long as cells with partial septa are not analyzed. Cephalexin-treated cells were 5-15 µm long. Untreated cells were 1-3 µm long. However cephalexin-treated cells were not used because many of them had partial septa.

4.3.5 Controls for FRAP Experiments

The photochemical properties of fluorophores can sometimes hinder experiments. For instance, GFP variants have been reported to photobleach reversibly and undergo photo-induced crosslinking.[101,102] To assess reversible photobleaching, confocal microscopy was used to bleach GFP throughout entire cells and to monitor the intracellular fluorescence over time. Post-bleach fluorescence recovery is not observed, indicating that GFP is irreversibly photobleached on the time scale of our experiments (Fig 4-3B). Likewise, when

experiments were conducted without a bleach pulse, no fluorescence decay is observed. To ensure that GFP does not undergo bleach-induced photochemistry, FRAP runs were repeated on the same cell several times. The same diffusion coefficient is obtained each time, showing that the system is not significantly altered by photochemistry.

4.3.6 Intracellular Diffusion of GFP

FRAP was used to measure the apparent diffusion coefficient of GFP in cells co-expressing an additional recombinant protein. Fig. 4-4A shows one-dimensional pixel intensity profiles of a typical cell as a function of distance from the cell edge at different post-bleach times. The observation that the fluorescence is constant with respect to the position at long times implies that either all or most of the GFP has long-range lateral mobility. Fig. 4-4B shows the amplitude of the first Fourier mode (n=1) for the same cell (blue), which is consistent with previous experiments.[83] The apparent diffusion coefficient was determined from the decay rate of this amplitude (Experimental Procedures). The average diffusion coefficients measured are summarized in Fig. 4-5. Regardless of the specific protein or the arabinose concentration used to induce expression, the intracellular diffusion coefficient of GFP, D_{GFP}, is the same within the uncertainty of the measurement. Likewise, the addition of 500-mM urea to the culture did not affect D_{GFP}, suggesting that GFP diffuses as a monomer.

4.3.7 Osmotic Shock

To confirm that our implementation of the FRAP experiment is sensitive enough to detect differences in D_{GFP}, cells were osmotically shocked with sorbitol. A 226-mOsm increase in the osmolarity decreased the decay rate of the first Fourier amplitude three fold (Fig. 4-4B). This decrease is consistent with previous studies.[83]

4-4 Discussion

We expanded on the simplistic suggestion of Dedmon *et al.* by examining the effects of protein expression on intracellular diffusion.[15] Dedmon et al. proposed that if intracellular protein concentration is constant, then overexpression of a single protein should not affect protein diffusion. However, overexpression could alter other factors such as cytoplasmic composition and protein interactions, both of which could affect protein diffusion. Thus, a more direct measurement of the effects of overexpression on diffusion was necessary.

For a given system, GFP was expressed at a constant level and its diffusion coefficient, D_{GFP}, was measured in the presence and absence of a second recombinant protein. Four proteins were chosen to cover a range of characteristics. Both large (MBP, 42 kDa; tau-40, 45 kDa) and small proteins (α-synuclein, 14 kDa; calmodulin, 17 kDa) were selected because evidence suggests that the size of the crowding agent affects diffusion.[103] Globular (MBP and calmodulin) and disordered (α-synuclein and tau-40) proteins were chosen to determine if shape is a factor. When induced with arabinose, the expressed

protein represented as much as 15% of the mass of the total intracellular protein.[104] Despite these variations in protein size, shape, and concentration, D_{GFP} remained within error of the commonly referenced GFP diffusion coefficient in *E. coli*, 7.7 ± 2.5 $\mu m^2 sec^{-1}$.[37] D_{GFP} was also consistent with the value recently measured by total internal reflection continuous photobleaching.[105] Our observations suggest that GFP diffusion is independent of the type or amount of protein co-expressed, but several points need to be addressed.

GFP aggregates could complicate the interpretation. Almost all intracellular protein diffusion studies in *E. coli* have involved GFP or GFP-fusion proteins. GFP may interact with other proteins inside the cell, which would decrease D_{GFP}. To test these potential protein-protein interactions, 500-mM urea was added to the cells. *E. coli* take up urea and remain viable at these concentrations.[106,107] The addition of urea did not alter the observed diffusion coefficient, implying that GFP is not interacting with itself or intracellular components.

FRAP could be too insensitive to changes in diffusion. Osmotic shock by sorbitol increases the intracellular concentration of macromolecules, which significantly decreases protein mobility.[108] To show that our instrumentation is sufficiently sensitive to detect differences in D_{GFP}, the GFP diffusion coefficient was measured in cells treated with 250-mM sorbitol. Dilution plating confirmed that sorbitol does not affect cell viability. We obtained a D_{GFP} of 2.4 ± 1.6 $\mu m^2 sec^{-1}$ in 397-mOsm buffer, which is consistent with other studies that report

diffusion coefficients of 1.8 $\mu m^2 sec^{-1}$ in 370-mOsm buffer [84] and 0.94 ± 0.55

$\mu m^2 sec^{-1}$ in 392-mOsm buffer.[83]

The issue of whether protein expression is uniform across the population

of cells must be addressed. Expression from the *araBAD* promoter is regulated

by arabinose, such that intermediate levels of protein expression can be

achieved by using subsaturating arabinose concentrations. The intracellular

concentrations of the co-expressed proteins (Table 4-1) were determined from

gels like the one shown in Fig. 3-9 and represent average intracellular

concentrations. It is known, however, that there is cell-to-cell variability of

expression levels at subsaturating inducer concentrations. Siegele and Hu [109]

showed that intermediate levels of bulk expression result from a mixture of fully

induced and uniduced cells. Thus, increasing the arabinose concentration

increases the fraction of induced cells, rather than the amount of protein

produced in an individual cell. Their results show that 0.02% arabinose is above

the saturation level such that most of the cells are fully induced. One explanation

for the few cells that were not induced is that they lost the vector contationg the

araBAD promoter. Our images of *E. coli* expressing GFP under the *araBAD*

promoter induced with 0.02% arabinose verify this observation (Fig. 4-2C).

Similar images of cells expressing GFP under the *lac* promoter show that GFP

expression is uniform from cell to cell, with the exception of tau-40 (Figs. 4-2A

and B).

Due to the behavior of the *araBAD* promoter at subsaturation, the diffusion

coefficients in Fig. 4-5 and the intracellular concentrations in Table 4-1 are

provided for only fully induced (≥0.02% arabinose) and uninduced cultures.

Nevertheless, intermediate levels of all co-expressed protein yield GFP diffusion

coefficients within error of those given in Fig 4-5 (not shown).

Interpreting the effect of the protein co-expression on D_{GFP} depends on

the assumption that the protein of interest is in the same compartment as GFP.

It is possible that the absence of an effect on GFP diffusion by the expression of

a second protein arises because the protein was not present in the same

compartment as the GFP. An *E. coli* cell comprises two main compartments: the

periplasm and the cytoplasm. The periplasm is the outer compartment bounded

by the plasma membrane and cell wall. It encloses 20-40% of the cell's total

volume.[110] The cytoplasm comprises the volume surrounded by the plasma

membrane. The cytosol is the cytoplasm without other subcellular structures,

such as ribosomes, and the fibrous proteins that determine cell shape, motility,

and material transport. GFP is found exclusively in the cytosol of *E. coli.* [111]

To determine if the co-expressed proteins are present in the cytoplasm or

periplasm, the cells were subjected to osmotic shock,[88,99] detergent, [98] osmotic

shock plus lysozyme,[97] or chloroform.[100] SDS-PAG electrophoresis of the

supermants (periplasmic proteins) and pellets (cytoplasmic proteins) confirmed

that the co-expressed proteins are soluble and present in the cytoplasm (Fig. 4-

1). The chloroform fraction of calmodulin was inconsistent with the results from

the other methods. However, even for MBP, α-synuclein, and tau-40, the

chloroform supernatant contains numerous proteins that are absent from the

supernatants from the other fractionation methods.   This observation suggests

that chloroform can lyse entire cells and not just the periplasm. Based on the osmotic shock, lysozyme, and detergent experiments, we conclude that calmodulin is a cytosolic protein. Our observation that 40% to 70% of the α-synuclein is cytosolic is inconsistent with a study that reports α-synuclein as exclusively periplasmic in *E. coli.*[88,100] Another study, which shows 60-85% of α-synuclein in the periplasm, suggests that the amount depends on the strain used for expression. The BL21-AI strain used here was not part of either previous study [100]. This difference in *E. coli* strain may explain the inconsistency in the amount of periplasmic α-synuclein. In summary, all the proteins except α-synuclein are found exclusively in the cytosol.

The tau-40 expression system was unlike the others. As shown in Table 1, tau-40 expression is three-fold lower than the other proteins when the same amount of inducer is added. Furthermore, GFP expression with this system exhibits extensive cell-to-cell variation as shown in Fig. 4-2B. This observation suggests that arabinose saturation has not been reached at 0.02%. The results with 0.2% arabinose are the same as those from a 0.02% sample, and 2% arabinose resulted in even lower expression. Nevertheless, the GFP diffusion coefficient for both induced and uninduced cultures were consistent with the other systems.

These control experiments strengthen our general conclusion that expression of these proteins has little or no effect on the diffusion of GFP. This diffusional homeostasis softens one criticism of in-cell NMR; namely, that the protein over expression required for observing in-cell spectra leads to an

unacceptably non-physiological environment. Therefore, the need to overexpress a protein for in-cell NMR does not necessarily invalidate the physiological relevancy of the results from in-cell NMR. The conclusion also raises new questions. It will be important to know what maintains diffusional homeostasis in the face of a single protein representing >10% of a cell's protein. In addition, the effects of inclusion bodies (which occur when overexpression leads to insolubility) need to be understood.

Our conclusion also highlights an enigma. NMR spectra from several soluble globular proteins cannot be observed in intact *E. coli* cells.[31] Yet, inclusion bodies are not observed, the soluble protein is easily purified, and simply lysing the cells causes the appearance of high-quality spectra. One explanation is that the high viscosity in cells slows the protein's rotation, making its resonances too broad to observe.[70]

The data presented here show that translational diffusion in cells is slowed about ten-fold compared to dilute solution. The enigma arises in trying to reconcile this ten-fold decrease with *in vitro* data. Adding enough of the macromolecular crowder, 40-kDa polyvinyl pyrrolidone, to slow translation of a small globular protein ten-fold, slows rotational diffusion by only two-fold.[31] A two-fold decrease in rotational diffusion is not enough to obliterate the spectrum of a small globular protein. It remains unclear how the cellular interior retards rotational diffusion to such a degree that protein NMR spectra cannot be observed, while having such a small effect on translational diffusion.

4-5 Conclusion

The data presented here indicate that protein co-expression has little to no effect on the intracellular translational diffusion of GFP. These observations suggest that in-cell NMR can provide biologically relevant data, despite the need to over express the protein being studied. However, the results also highlight new questions that must be addressed about cellular protein dynamics.

4.6 Tables

Table 1. Intracellular Concentrations[a]

	Co-expressed Protein	GFP	
	Induced (mM)	Uninduced (mM)	Induced (mM)
MBP	0.6 ±0.1[b]	0.9 ±0.5	1.0 ±0.2
Calmodulin	0.8 ±0.3	0.6 ±0.3	0.6 ±0.3
Tau-40	0.21 ±0.06	0.5 ±0.1	0.5 ±0.2
α-Synuclein[c]	1.27 ±0.09	0.16 ±0.06	0.19 ±0.07

[a] quantified by integrating pixel intensities of bands from Coomassie-stained SDS-PAGs

[b] standard deviation, n=3

[c] some of this protein is periplasmic (see text)

4.7 Figures
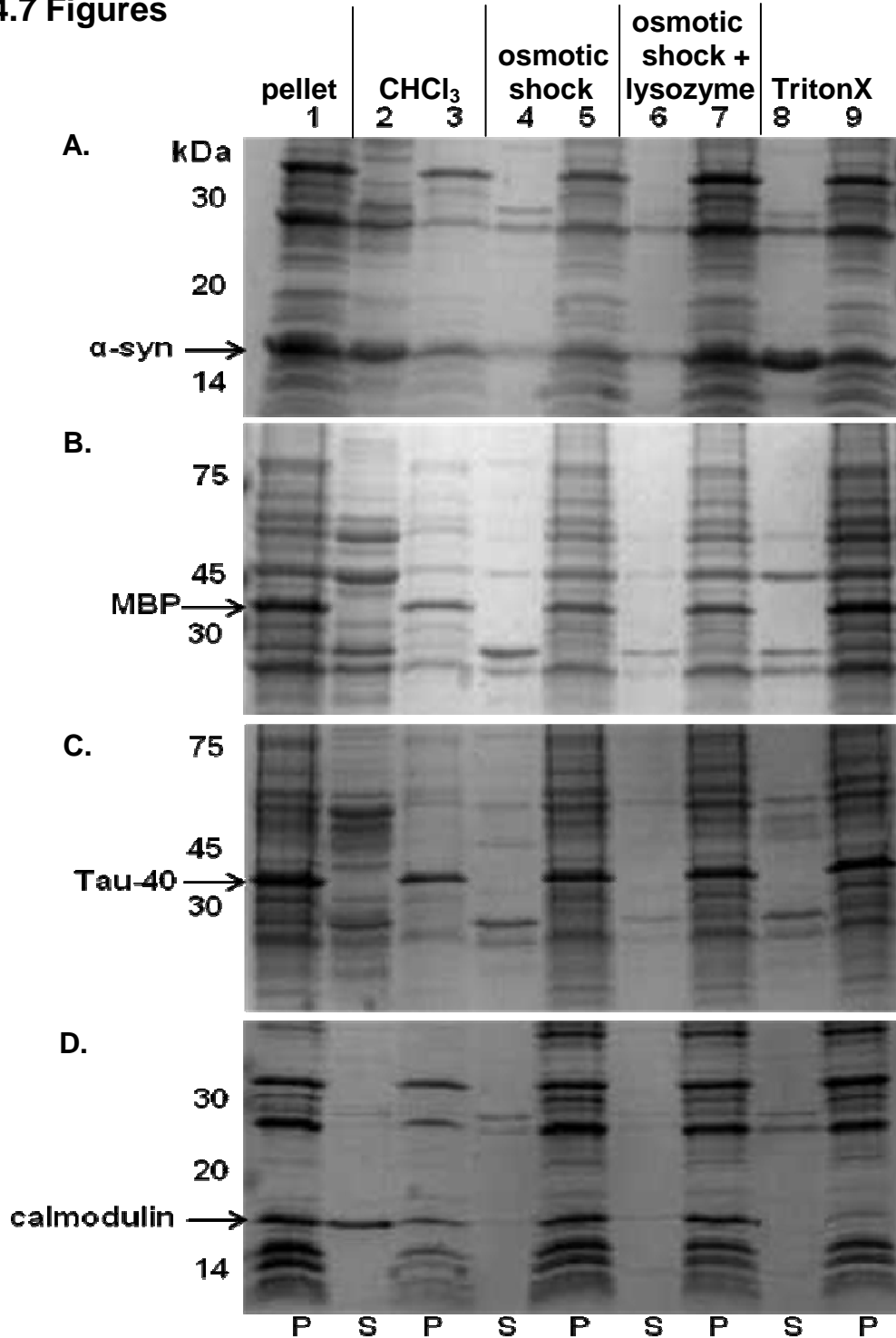
Figure 4-1 Protein Location. Aliquots of *E. coli* co-expressing GFP and either (A) α-synuclein (B) MBP (C) tau-40 or (D) calmodulin were centrifuged. The pellets were then exposed to either chloroform (lanes 2, 3), osmotic shock (lanes 4, 5), osmotic shock and lysozyme (lanes 6, 7), or TritonX detergent (lanes 8, 9). The pellets (P) and supernants (S) were resolved by electrophoresis on 10-to-20% gradient SDS-PAGs. Lane one is the untreated cell lysate.
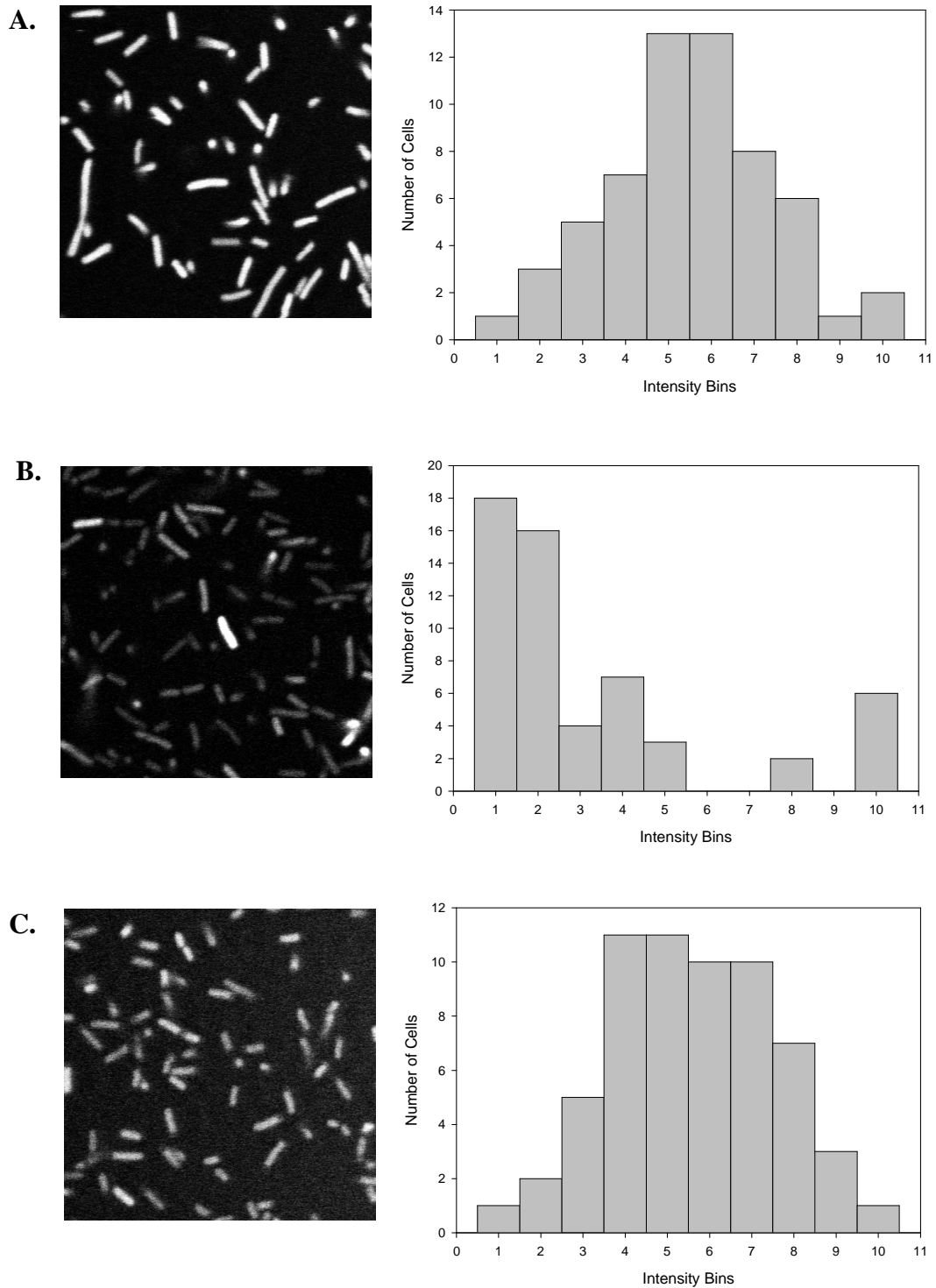
Figure 4-2 Fluorescent images and corresponding histograms of *E. coli* expressing GFP (35µm x 35µm). In panels A and B, GFP, under the *lac* promoter, is co-expressed with (A) α-synuclein or (B) tau-40, both under the *araBAD* promoter. In panel C, GFP is expressed under the *araBAD* promoter, instead of the *lac* promoter with no co-expressed protein.

A.
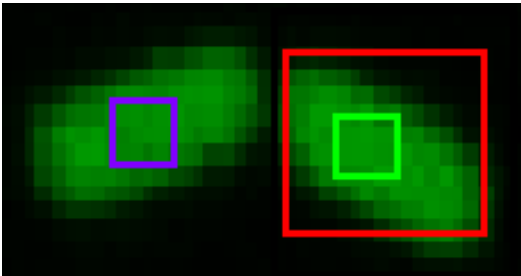
B.
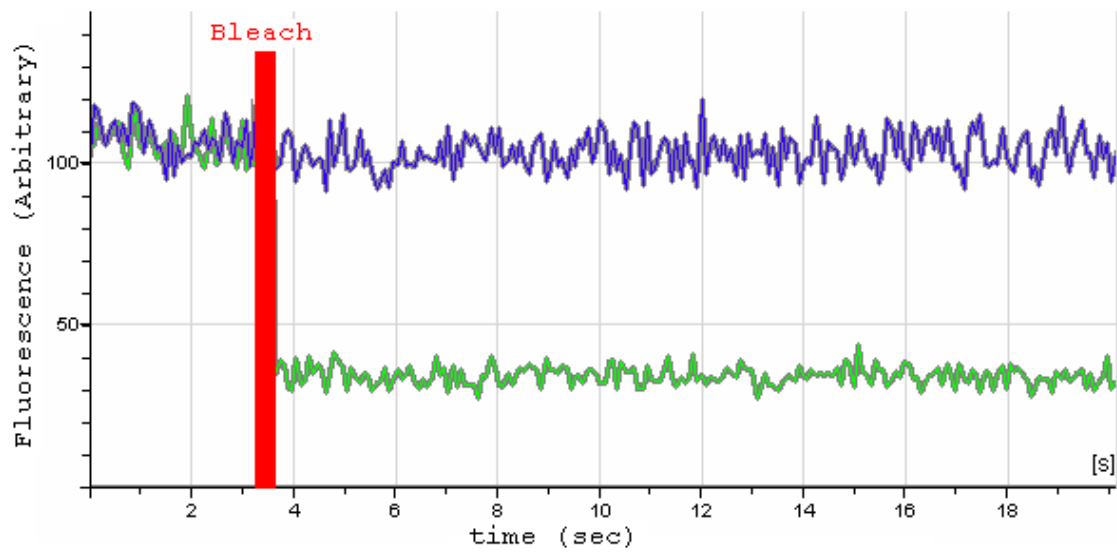
Figure 4-3 Controls for FRAP experiments. (A) Fluorescent images of *E. coli* cells expressing GFP. The red box indicates the region that was bleached. (B) Fluorescence intensity of the green and blue boxed regions in A.

A.
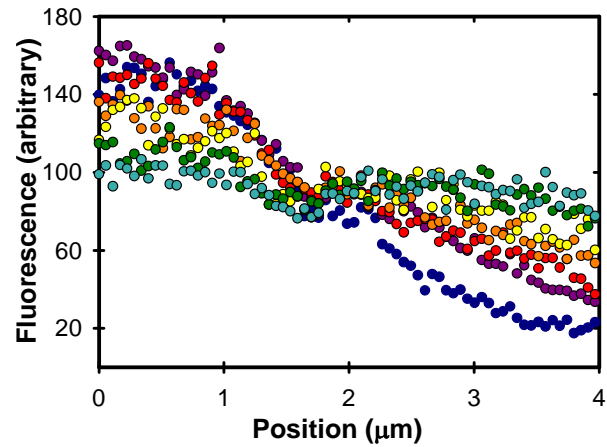
B.
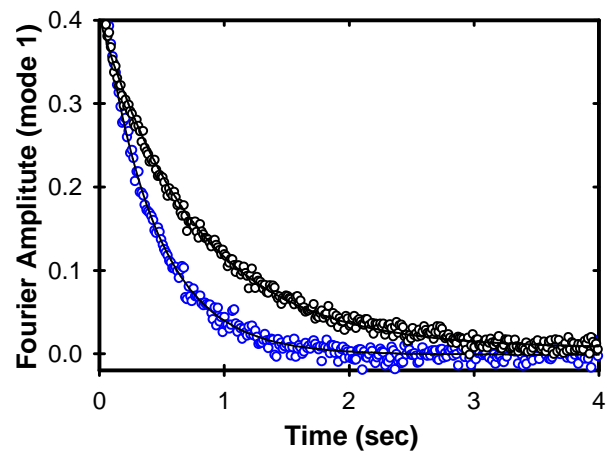
Figure 4-4 Photobleaching data. (A) Fluorescence intensity profiles at 0.03 (dark blue), 0.10 (purple), 0.17 (red), 0.30 (orange), 0.50 (yellow), 0.83 (green), 1.51 (cyan) s after photobleaching are shown for a BL21-AI cell co-expressing tau-40 and GFP. (B) Temporal decay of the first Fourier amplitude for the same cell (blue), and for a cell exposed to a 226-mOsm increase in osmolality (black). The solid lines are fits to the exponential function, aexp(-bt) + c, with a, b, c as free parameters.
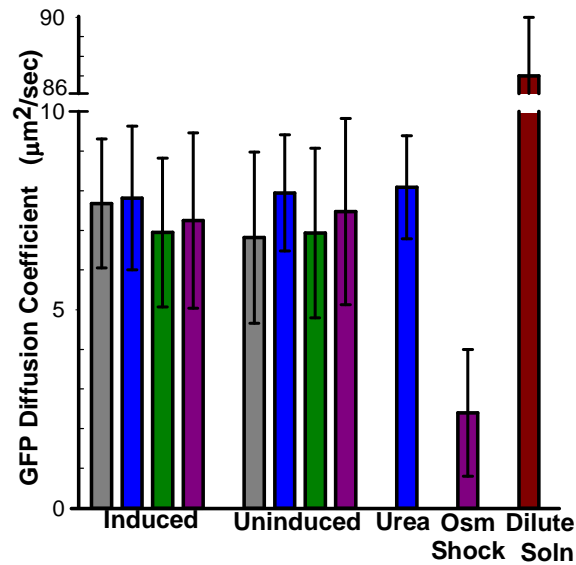
Figure 4-5 Co-expressing another recombinant protein does not affect intracellular GFP diffusion. GFP was co-expressed with either calmodulin (grey), MBP (blue), tau-40 (green), or α-synuclein (purple). The arabinose concentrations used for induction were 0.02%, 0.02%, 0.02%, and 0.2% respectively.   GFP diffusion in the absence of inducer (arabinose) from these same co-expression systems is also shown (n > 34). GFP was expressed constitutively from an uninduced *lac* promoter. The induced MBP sample was exposed to 500-mM urea (n = 21). Osmotic shock was introduced by adding sorbitol to a final concentration of 250 mM (n = 19). For comparison, the GFP diffusion coefficient in dilute solution is also shown.[112]

Chapter 5: Quantifying GFP Diffusion in *Escherichia coli* by Using Continuous Photobleaching with Evanescent Illumination

Summary

Fluorescence recovery after photobleaching and fluorescence correlation spectroscopy are the primary means for studying translational diffusion in biological systems. Both techniques, however, present numerous obstacles for measuring translational mobility in structures only slightly larger than optical resolution. We report a new method using through-prism total internal reflection fluorescence microscopy with continuous photobleaching to overcome these obstacles. Small structures, such as prokaryotic cells or isolated eukaryotic organelles, containing fluorescent molecules are adhered to a surface. This surface is continuously illuminated by an evanescent wave created by total internal reflection. The characteristic length describing the decay of the evanescent intensity with distance from the surface is smaller than the structures. The fluorescence decay rate resulting from continuous evanescent illumination is monitored as a function of the excitation intensity. The data at higher excitation intensities provide apparent translational diffusion coefficients for the fluorescent

molecules within the structures because the decay results from two competing processes (the intrinsic photobleaching propensity and diffusion in the small structures). We present the theoretical basis for the technique and demonstrate its applicability by measuring the diffusion coefficient, 6.3 ± 1.1 μm^2/sec, of green fluorescent protein in *Escherichia coli* cells.

5.1 Introduction

FRAP suffers from the fact that the smallest focused laser spot used for illumination is not much smaller than an *Escherichia coli* cell or an organelle.[113] Thus, not only is optical alignment difficult, but also the quantity of unbleached molecules that can contribute to fluorescence recovery is limited. As described in previous chapters, these effects reduce the signal-to-noise ratio and complicate data analysis. In addition, the recovery time associated with a small, focused spot and solute diffusion is too rapid for many conventional, simpler instruments.

Fluorescence correlation spectroscopy (FCS) is the other primary method for measuring translational diffusion in biological systems.[114-116] FCS suffers in the context of small, contained structures, since the size of the illuminated region is on the same order of magnitude as the structures. Consequently, there is only a small population of non-illuminated molecules. Under such conditions, the fluorescence fluctuations can be recorded for only a short time before the reservoir of unbleached molecules is depleted by photobleaching. This limitation significantly reduces the signal-to-noise ratio of the fluorescence fluctuation

autocorrelation function. Nonetheless, FCS has been successfully employed to measure protein diffusion coefficients in *E. coli* cells.[117] The method is limited, however, to low concentrations of fluorescent molecules.

Continuous fluorescence microphotolysis or continuous photobleaching (CP) is an alternate method for characterizing lateral diffusion in biological systems. In CP, a small region of a fluorescently labeled sample is continuously illuminated such that two competing processes arise—photobleaching of fluorophores in the illuminated region and fluorescence recovery due to diffusion of unbleached fluorophores from surrounding regions into the illuminated region. By monitoring the rate and shape of the fluorescence decay, rate constants for the two processes can be determined by fitting to appropriate theoretical forms.[118,119] Recently, CP has been used, along with FCS, to measure diffusion and compartmentalization in giant unilamellar vesicles and in large living cells.[120] The results of CP, FCS, spatial imaging, and confocal microscopy have been used together to analyze diffusion of intracellular molecules and binding to specific sites in cells.[121] CP has also been combined with 4Pi microscopy to obtain higher spatial resolution.[122] Pulsed FRAP, a modification that combines CP and FRAP, has been used with confocal microscopy to measure the diffusion coefficients of fluorescent proteins in *E. coli.*[123]

Total internal reflection (TIR) has previously been combined with FRAP to measure solute diffusion in eukaryotic cells.[124] The small penetration depth of the evanescent wave results in photobleaching only near the interface. Since this depth is much smaller than the diameter of the laser beam, a one-

dimensional geometry can be used to simplify the mathematical diffusion model. The small penetration depth, however, also leads to fast recovery, thus requiring special equipment (e.g., acousto-optic modulators).

Here, we combine total internal reflection and continuous photobleaching to demonstrate a new method, TIR-CP, for characterizing the translational diffusion of fluorescent molecules contained in structures only slightly larger than classical optical resolution. These structures are deposited on a surface at which a laser beam is internally reflected, such that the resulting evanescent intensity illuminates only those fluorescent molecules close to the surface. The sample is continuously photobleached as a function of the excitation intensity. As in CP, two competing processes contribute to the rate and shape of fluorescence decay. At low intensities, the decay is determined by the photobleaching rate within the evanescent wave. At higher intensities, the decay is also affected by diffusion of the fluorescent molecules. Thus, the diffusion coefficient of the fluorescent molecules within the small structures can be determined by acquiring data as a function of the excitation intensity. This chapter describes the theoretical basis for this new method and demonstrates its applicability by measuring the diffusion coefficient of GFP in *E. coli* cells.

5.2 Theory

5.2.1 Conceptual Basis

The notion behind the new method is illustrated in Fig. 5-1. Small structures of average length L are deposited on a surface at which a laser beam

101

is internally reflected. The evanescent intensity decays with distance z from the

surface and with a spatial profile in the x-y plane as

$$I(x, y, z) = I_0 \exp(-\frac{2x^2}{w_x^2}) \exp(-\frac{2y^2}{w_y^2}) \exp(-\frac{z}{d})$$ (5-1)

where I_0 is the intensity at the interface (z = 0) and at the spot center (x = y =

0).[125] The characteristic distance for the evanescent wave decay, d, depends on

the excitation wavelength, the incidence angle, and the refractive indices of the

two materials at which internal reflection occurs. Parameters w_x and w_y are $1/e^2$

values for the elliptical Gaussian shape of the internally reflected beam in the

sample plane and depend on the initial beam radius as well as the optical

parameters used to generate internal reflection.

The observed sample volume is defined by the depth of the evanescent

intensity and by a pinhole placed at a back image plane of an optical microscope

through which the fluorescence is collected. The pinhole, which is positioned to

correspond to the center of the illuminated region (x = y = 0), restricts

fluorescence observation to a small volume so that the collected fluorescence is

low enough to be measurable even at relatively high excitation intensities. For

the same reason, fluorescence is collected though a low numerical aperture

objective. The low numerical aperture does not compromise z-axis resolution,

which is very thin as defined by the evanescent excitation intensity. The

observed area is small enough so that only a few of the small structures are

present in this area.

The sample's fluorescence is monitored as a function of time, t, with t = 0

corresponding to the onset of illumination. The time-dependent fluorescence

decays to zero as molecules within the evanescent wave are photobleached and the reservoir of unbleached molecules within the small structure is depleted. Two competing processes, diffusion and photobleaching, contribute to the rate and shape of fluorescence decay. At low intensities, the decay is determined by the photobleaching rate within the evanescent wave. At high enough intensities, the photobleaching rate is fast enough that the diffusion of unbleached molecules into the evanescent wave becomes the rate limiting step. By acquiring data as a function of the excitation intensity, the diffusion coefficient of the fluorescent molecules within the small structure is determined.

5.2.3 Unbleached Molecule Concentration as a Function of Space and Time

The evanescently illuminated area is much larger than the observed area, and the pinhole is placed at the center of this illumination. Since the sample radius (~ 5 μm) is much less than w_x and w_y (22 and 65 μm respectively), the intensity at z = 0 does not vary much over the observed area in the x-y plane. The intensity at z = 0 also does not vary significantly as a function of x and y for a given cell because an *E. coli* cell is smaller than the observed area. Furthermore, the characteristic distance of the evanescent wave decay (d ~ 0.1 μm) is much less than the length of a cell in the z-direction (L ~ 2 μm). For these reasons, the mathematical problem is approximately one-dimensional in space with the key coordinate being z. The evanescent intensity (Eq. 5-1) is then approximated as

$$I(z) \approx I(0,0,z) = I_0 \exp(-\frac{z}{d}) \tag{5-2}$$

and the concentration of unbleached molecules is assumed not to depend on coordinates x and y. As a result, an approximate differential equation for the concentration of unbleached molecules is spatially one-dimensional and depends only on the distance from the interface, z, and the time, t. This equation is

$$\frac{\partial}{\partial t} U(z,t) = D \frac{\partial^2}{\partial z^2} U(z,t)$$ (5-3)

where U(z,t) is the concentration of unbleached molecules as a function of space and time, and D is the diffusion coefficient. Unlike similar theories,[118,119] Eq. 5-3 lacks a term describing photobleaching, because photobleaching occurs only at or near the illuminated surface. Instead, this process is described by a boundary condition (Eq. 5-6, see below).

The initial condition is

$$U(z,0) = C$$ (5-4)

where C is the total concentration of fluorescent molecules. One boundary condition is

$$[\frac{\partial}{\partial z} U(z,t)]_{z=L} = 0$$ (5-5)

where L is the length of the cell measured from the surface into the solution. This "reflection" condition expresses the notion that the flux across the boundary of the structure far from the interface is zero. The other boundary condition is

$$D[\frac{\partial}{\partial z} U(z,t)]_{z=0} = d\kappa I_0 U(0,t)$$ (5-6)

where κ is a proportionality constant describing the photobleaching propensity with units of intensity^{-1} time^{-1}.[126]

The general solution to Eq. 5-3 is

$$U(z,t) = [A\cos(\beta z) + B\sin(\beta z)]\exp(-D\beta^2 t) \qquad (5\text{-}7)$$

where A, B and β are constants. Eqs. 5.5-5.7 [i.e., the two boundary conditions and the general solution] imply that

$$A\sin(\beta L) = B\cos(\beta L)$$
$$BD\beta = d\kappa I_0 A \qquad (5\text{-}8)$$

These two equations yield a discrete, countably infinite number of β and x values defined by

$$\beta_n \tan(\beta_n L) = \frac{d\kappa I_0}{D} \qquad\qquad x_n \tan(x_n) = c$$

$$x_n \equiv \beta_n L \qquad\qquad c \equiv \frac{d\kappa I_0 L}{D} \qquad (5\text{-}9)$$

where x_n (for n = 1,2,3,...) and c are dimensionless quantities.

The values of x_1, x_2 and x_3 (determined numerically) are shown in Fig. 5-2a as a function of c, which is proportional to the excitation intensity, I_0 (Eq. 5-9). The values of x_n have the following properties. First, $(n-1)\pi \leq x_n \leq (2n-1)\pi/2$. Second, x_n increases with c. At extremely low intensities, $c \approx 0$, $\tan(x_1) \approx 0$, and $x_1^2 \approx 0$. At higher intensities, $\tan(x_1) \approx x_1$ and $x_1^2 \approx c$. As shown in Fig. 5-3, this approximation is accurate within 10% up to $c \leq 0.3$. At even higher intensities,

$$\tan(x_1) \approx x_1 + \frac{x_1^3}{3}$$

$$c \approx x_1 \tan(x_1) \approx x_1^2 + \frac{x_1^4}{3} \qquad (5\text{-}10)$$

$$x_1^2 \approx \frac{-3 + (9 + 12c)^{1/2}}{2}$$

As shown in Fig. 5-3, this approximation is accurate within 10% up to $c \leq 1.5$. At

105

even higher intensities,

$$\tan(x_1) \approx x_1 + \frac{x_1^3}{3} + \frac{2x_1^5}{15}$$

$$c \approx x_1 \tan(x_1) \approx x_1^2 + \frac{x_1^4}{3} + \frac{2x_1^6}{15}$$

(5-11)

$$x_1^2 \approx \frac{1}{6} \left\{ \begin{array}{l} -5 + \dfrac{(13)(5^{2/3})}{[-110 - 162c + 9(285 + 440c + 324c^2)^{1/2}]^{1/3}} \\ -(5^{1/3})[-110 - 162c + 9(285 + 440c + 324c^2)^{1/2}]^{1/3} \end{array} \right\}$$

As shown in Fig. 3, this approximation is accurate within 10% up to c ≤ 2.9. For

extremely large intensities, c → ∞ and x_1 → π/2.

Eqs. 5-8 and 5-9 also imply that

$$A_n = \frac{x_n}{c} B_n$$

(5-12)

Thus (Eqs. 5-7 and 5-12),

$$U(z,t) = \sum_{n=1}^{\infty} B_n [\frac{x_n}{c} \cos(\frac{x_n z}{L}) + \sin(\frac{x_n z}{L})] \exp[-D(\frac{x_n}{L})^2 t]$$

(5-13)

At time zero (Eqs. 5-4 and 5-13),

$$C = \sum_{n=1}^{\infty} B_n [\frac{x_n}{c} \cos(\frac{x_n}{L} z) + \sin(\frac{x_n}{L} z)]$$

(5-14)

Multiplying both sides of Eq. 5-14 by the factor,

$$\frac{x_m}{c} \cos(\frac{x_m}{L} z) + \sin(\frac{x_m}{L} z)$$

(5-15)

and integrating z from zero to L (with Eq. 5-9) implies that

$$U(z,t) = \sum_{n=1}^{\infty} \frac{2c^2 C}{x_n(c + c^2 + x_n^2)} [\frac{x_n}{c} \cos(\frac{x_n z}{L}) + \sin(\frac{x_n z}{L})] \exp[-D(\frac{x_n}{L})^2 t]$$

(5-16)

It can be shown in a straightforward manner (with Eq. 5-9) that each term in Eq. 15-6 satisfies Eqs. 5-3, 5-5 and 5-6. The consistency of Eq. 5-16 with Eq. 5-4 can be demonstrated numerically.

5.2.4 Fluorescence Decay During Continuous Photobleaching

The observed fluorescence during continuous photobleaching at $(x,y) = (0,0)$ is given by (Eqs. 5-2, 5-9 and 5-16)

$$F(t) = QI_0 \int_0^L dz \exp(-\frac{z}{d})U(z,t) = \sum_{n=1}^{\infty} F_n \exp[-D(\frac{x_n}{L})^2 t]$$

$$F_n = 2QI_0 CcdL \frac{(cd+L)x_n^2 - \exp(-\frac{L}{d})[L(c^2+x_n^2)\cos(x_n)]}{x_n^2(c+c^2+x_n^2)(L^2+d^2x_n^2)}$$

(5-17)

where Q is a proportionality constant. The initial fluorescence value is (Eqs. 5-4 and 5-17)

$$F(0) = QCI_0 d[1 - \exp(-\frac{L}{d})]$$

(5-18)

Thus,

$$f(t) \equiv \frac{F(t)}{F(0)} = \sum_{n=1}^{\infty} f_n \exp[-D(\frac{x_n}{L})^2 t]$$

$$f_n = \frac{2cL\left\{(cd+L)x_n^2 - \exp(-\frac{L}{d})[L(c^2+x_n^2)\cos(x_n)]\right\}}{[1-\exp(-\frac{L}{d})]x_n^2(c+c^2+x_n^2)(L^2+d^2x_n^2)}$$

(5-19)

For our experimental conditions, $d \approx 0.1$ µm and $L \approx 2.2$ µm (see below). Fig. 5-2 B shows the values of f_1, f_2 and f_3 as function of c, for these values of d and L. As shown, f_1 drops below 0.9 only for $c > 0.3$. Thus, for $c < 0.3$, the fluorescence decay can be accurately approximated by a single exponential with

rate $R_1 = D(x_1/L)^2$. For our experimental conditions, $c \leq 2$ (see below). As shown in Fig. 5-2 B, for $0.3 \leq c \leq 2$, f_1 ranges from 0.9 to 0.6. Therefore, one might expect that multi-exponential analysis would be required for the higher intensities. For these intensities, however, the second rate, $R_2 = D(x_2/L)^2$, ranges from 5-18 sec^{-1} (see below), faster than the time resolution of our software (50 msec), and contributions to f(t) from terms with n > 1 are negligible. Therefore, we evaluate all data as a single exponential with rate, R, where

$$f(t) \approx \exp[-Rt]$$
$$R \approx D(\frac{x_1}{L})^2$$

(5-20)

5.2.5 Limits as a Function of Intensity Parameter c

When c is small (≈ 0), $x_1^2 \approx 0$ (Eq. 5-9), and Eq. 5-20 predicts the expected result that $R \approx 0$ and $f(t) \approx 1$ is constant with time. In other words, for very low intensities, photobleaching does not occur. When c is small but nonzero, $\tan(x_1) \approx x_1$, $x_1^2 \approx c$ and Eqs. 5-9 and 5-20 predict that

$$f(t) \approx \exp(-\frac{\kappa I_0 d}{L} t)$$

(5-21)

The ratio $x_1^2/c \geq 0.9$ for $c \leq 0.3$ (Fig. 5-3B). In this limit, diffusion within the small structure is fast enough so that the fluorescence decay rate R does not depend on D but only on d, L, κ, and the intensity, as expected. In addition, R is linear with the intensity. For extremely high intensities ($c \to \infty$), $x_n \approx (2n-1)\pi/2$, $\cos(x_n) \approx 0$, $cd \gg L$, and $c^2 \gg c$. Also, $d \ll L$. By using these approximations and the method of partial fractions, one can thus show that (Eq. 5-19)

$$f(t) \equiv \frac{F(t)}{F(0)} = \sum_{n=1}^{\infty} f_n \exp\left\{-D[\frac{(2n-1)\pi}{2L}]^2 t\right\}$$

$$f_n \approx \frac{2dL}{L^2 + d^2 (\frac{(2n-1)\pi}{2})^2} \tag{5-22}$$

As expected, for high intensities, the fluorescence decay depends only on d, D and L and not on the intensity.

5.2.6 Measurements with Immobilized GFP

To determine the approximate value of κ, the decay of fluorescence with time was measured with purified GFP immobilized on the surface. In this case,

$$\frac{d}{dt}F(t) = -\kappa I_0 F(t)$$
$$\frac{F(t)}{F(0)} = \exp(-\kappa I_0 t) \tag{5-23}$$

where F(t) is proportional to the density of unbleached GFP on the surface.

5.3 Materials and Methods

5.3.1 GFP Expression

The pAcGFP1 vector (BD BioSciences Clontech, Palo Alto, CA), which contains the gene for a nondimerizable GFP[127] with 94% identity to EGFP, was transformed into *E. coli* BL21-Gold (DE3) competent cells (Stratagene, La Jolla, CA) and plated on Luria Broth plates containing 100 µg/mL ampicillin (LB_{AMP}). A starter culture of liquid LB_{AMP} was inoculated with a single colony and grown overnight at 37°C with constant shaking at 225 rpm. This starter culture was used to inoculate (1:25 dilution) 25 mL of fresh LB_{AMP} in a 250-mL flask. Once the optical density at 600 nm was between 0.5 and 0.7, the culture was induced

with a final concentration of 1-mM isopropyl-β-D-thiogalactopyranoside (IPTG) and allowed to grow at 37˚C with constant shaking at 225 rpm.

5.3.2 GFP Purification

One liter of culture prepared as described above was harvested 4 h after induction by centrifugation (Sorvall RC-3B, Sorvall Instruments, Newtown, CT) at 1,600 g for 30 min at 4˚C. The pellet was resuspended in 30 mL of lysis buffer [20-mM Tris (pH 8.0), 200-mM NaCl, 1-mM EDTA, 1-mM PMSF, 1-mM DNase, 1-mM RNase] and pulse sonicated (Branson Ultrasonics, China) at 4˚C (18% amplitude) for two rounds of 5 min each. Cell debris were removed by centrifugation (Sorvall RC-5B with a SS-34 rotor) at 27,000 g. After dialyzing overnight at 4˚C against 20-mM Tris buffer (pH 8), the sample was purified by using anion exchange chromatography (HiLoad 16/10 Q Sepharose, AKTA FPLC UPC-900, GE Healthcare, Piscataway, NJ) in 20-mM Tris buffer using a linear gradient from 0-M to 1-M NaCl. The fractions containing GFP were concentrated in an Amicon Ultra MWCO 3,000 centrifugal filter unit (Millipore, Billerica, MA) and dialyzed into water. The protein was further purified by using size exclusion chromatography (16/60 superdex 75, GE Healthcare) with water. The purity was confirmed by using SDS-PAGE (18%) with Coomassie Brilliant Blue staining.

5.3.3 Sample Preparation

Fused silica slides (1" x 1" x 1 mm, Quartz Scientific, Fairport Harbor, OH) and glass microscope slides (3" x 1" x 1 mm, Fisher Scientific, Fair Lawn, NJ)

were boiled in ICN detergent (MP Biomedicals, Solon, OH) for 10 min, bath-sonicated for 30 min, rinsed thoroughly with deionized water, and dried overnight at 160°C. The dried slides were cleaned in an argon-ion plasma cleaner (PDC-3XG, Harrick Scientific, Ossining, NY) for 15 min at 25°C immediately prior to use. The fused silica slides were pretreated with a 0.01% (w/v) poly-L-lysine solution (Sigma-Aldrich, St. Louis, MO) for 15 min, rinsed with minimal media [7.6-mM $(NH_4)_2SO_4$, 60-mM K_2HPO_4, 2-mM $MgSO_4$, 20-µM $FeSO_4$, 1-mM EDTA (pH 6.8)], and attached to a microscope slide with double-sided tape (Part No. 021200-64988, 3M Corp, St. Paul, MN) to form a sandwich. For osmotic stress measurements, the minimal media rinse contained 250-mM sorbitol (0.390 osmolal). The samples, which consisted of either bacterial cultures collected 3 h after induction or 2-µM purified GFP containing 5-mg/mL bovine serum albumin (Sigma-Aldrich, St. Louis, MO), were injected into the sandwiches. After incubation for 30 min at 25°C, the sample chamber was rinsed with minimal media and sealed with vacuum grease.

5.3.4 Total Internal Reflection Fluorescence Microscopy

Through-prism total internal reflection (TIR) bleaching experiments were carried out on an instrument consisting of an inverted microscope (Zeiss Axiovert 35, Thornwood, NY), an argon ion-laser (Innova 90-3; Coherent, Palo Alto, CA), and a single-photon counting photomultiplier (RCA C31034A, Lancaster, PA). The instrument was controlled with an in-house LabVIEW program and DAQ board (PCI-MIO-16XE-50, Texas Instruments, Austin, TX). Experiments were

conducted at 25°C by using an excitation wavelength of 488 nm and a laser power ≤ 500 mW. To achieve optimal excitation intensity without overloading the detector, a 100-µm diameter pinhole was inserted in a back intermediate image plane of the microscope and aligned to correspond with the center of the TIR illumination. Polarization paper (25% transmission, Edmund Optics, NT54-795) was also inserted after the dichroic mirror and barrier filter, but before the detector, to attenuate the signal. The bleaching intensity was varied by inserting neutral density filters in the beam path prior to excitation. The fluorescence decay was collected with a 10x, 0.25 numerical aperture objective for up to 90 sec.

5.3.5 Size of Evanescent Illumination

TIR was generated on a fused silica/microscope slide sandwich containing a 1-µM solution of Alexa Fluor 488 carboxylic acid, tetrafluorophenyl ester (Molecular Probes/Invitrogen, Carlsbad, CA) as described above. Five images of the TIR spot were collected with an AT200 CCD camera and software (Photometrics, Tucson, AZ). To determine w_x and w_y (Eq. 5-1), slices of the images with y = 0 or x = 0 were converted to pixel intensities W(x) or W(y) with the Photometrics software and plotted as a function of distance. The pixel dimension (0.87 µm) was determined by imaging a graticule. The data were fit in Sigma Plot (Systat Software Inc., San Jose, CA) to

$$W(x) = W_0 \exp(-\frac{2x^2}{w_x^2}) + \gamma \tag{5-24}$$

or

$$W(y) = W_0 \exp(-\frac{2y^2}{w_y^2}) + \gamma \qquad (5\text{-}25)$$

with W_0, w_x(or w_y), and γ (the background) as free parameters.

5.3.6 Intensity Values

For a given incident laser power, P, at z = 0 (Eq. 5-1),

$$P = I_0 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp(-\frac{2x^2}{w_x^2}) \exp(-\frac{2y^2}{w_y^2}) = \frac{\pi}{2} w_x w_y I_0 \qquad (5\text{-}26)$$

The value of I_0 was determined as

$$I_0 = \frac{2P}{\pi w_x w_y} \qquad (5\text{-}27)$$

by using Eq. 5-27 with the known values of P, w_x and w_y.

5.3.7 Cell Length

Samples were prepared as described above except that 5-µg/mL FM1-43 membrane stain (Molecular Probes/Invitrogen, Carlsbad, CA) was added to the samples and No. 1.5 glass coverslips (Fischer Scientific, Fair Lawn, NJ) were used in place of the fused silica slides. The cells were imaged on a Zeiss 510 scanning confocal inverted microscope (Carl Zeiss, Thornwood, NY) with a 63x, 1.4 NA oil-immersion objective, a 65-µm pinhole, and an excitation wavelength of 488 nm. Sixteen 36.6-µm x 36.6-µm images were collected in the x-y plane by moving away from the coverslip in 0.2-µm increments. From these images, the maximum lengths of 110 cells in the z-direction were determined and averaged, yielding L = 2.2 ± 0.5 µm. Repeating this process for 107 sorbitol-treated cells gave L = 1.8 ± 0.5 µm.

5.3.8 Data Analysis

Fluorescence decays were fit to the following function by using Sigma Plot:

$$F(t) = F(0)\exp[-Rt] + \varphi \qquad\qquad (5\text{-}28)$$

with F(0), R and φ (the background) as the free parameters. At least three (and often more) decay curves were collected for each I_0 and the R values averaged. Three complete data sets were acquired. The average values of R as a function of I_0, for the three data sets, were fit to Eq. 5-20 with x_1^2 given by Eq. 11 and c given by Eq. 5-9. In these fits, L was fixed to 2.2 µm (see above) and the free parameters were D and b = κd. The process was repeated for the sorbitol-treated cells using L =1.8 µm (see above).

5.4 Results

5.4.1 Size of Evanescent Illumination in the x-y Plane and I_0 Values.

To determine the values of w_x and w_y (Eq. 5-1), five images of fused silica slides coated with Alexa Fluor 488 were acquired using an imaging detector (see Methods). As shown in Fig. 5-4A, the lateral intensity profile of the evanescent illumination appeared to be of an elliptical Gaussian shape.[125] Images like those shown in Fig. 5-4A report the value of I(x,y,0) (Eqs. 5-1, 5-24, 5-25 and 5-26). The pixel intensities as a function of distance along the x-axis (y = 0) and along the y-axis (x = 0) were plotted for each image. These data were fit to Eqs. 5-24 and 5-25 and gave Gaussian-shaped curves as shown by the quality of the fits (Fig. 4B). The averages of the best-fit values were w_x = 22.4 ± 0.5 µm and w_y =

65.0 ± 0.4 μm. Excitation intensities I_0 were then determined by using Eq. 5-27.

5.4.2 Cell Length

The cell length L was measured by using confocal microscopy and *E. coli* treated with the membrane stain FM1-43. Sixteen 36.6-μm by 36.6-μm images were collected in the x-y plane by moving away from the coverslip to which the cells were adhered in 0.2-μm increments. Fig. 5-5 shows two slices in the x-y plane of cells at different distances from the interface. The majority of the cells were ellipsoidal and had their major axes parallel to the interface. The circles in Fig. 5-5 A are cells attached perpendicular to the interface and thus extend further into the solution. This result is apparent in Fig. 5-5 B (the slice farther from the interface) where the ellipsoidal shaped cells have begun to fade, while the fluorescence of the circular cells remains strong. These images also confirmed that the cells were immobilized by the poly-L-lysine and that the GFP was contained within the cells. By averaging the maximum length in the z-direction of 110 cells in different spatial orientations, L was determined to be 2.2 $\pm$ 0.5 μm, which is consistent with the literature value.[128,129] Repeating this process for 107 sorbitol-treated cells gave L = 1.8 $\pm$ 0.5 μm. With > 99% confidence (student's t-test), the sorbitol-treated cells are shorter than non-treated cellst, since osmotic stress shrinks the cells.[130]

5.4.3 Photobleaching Propensity

The propensity for photobleaching, described by the parameter κ (Eqs. 5-

6, 5-9 and 5-23), was measured by monitoring the evanescently excited fluorescence decay of pure GFP immobilized on fused silica slides, as a function of time and excitation intensity. Fig. 5-6A shows three representative decay curves and the corresponding best fits to Eq. 5-28. At very low excitation intensities, bleaching was almost negligible. As the laser power was increased, so did the initial fluorescence intensity and the decay rate. As shown in Fig. 5-6 B, the decay rate was linearly proportional to the excitation intensity. The slope of this line yields a κ (for immobilized GFP) of 0.21 ± 0.01 $\mu m^2 \mu W^{-1} sec^{-1}$.

5.4.4 Diffusion Coefficient of GFP in *E. coli*

Three complete data sets for the intensity-dependent, evanescently-excited fluorescence decays of GFP in *E. coli* were acquired. Fig. 5-7 shows examples of typical decay curves collected from continuous photobleaching of the cells close to the adherent surface. For each decay curve, the decay rate constant, R, was determined by fitting the data to Eq. 5-28 with the intensity, I_0, determined as described above. As expected, both F(0) and R increased with I_0. For each data set, the average values of R as a function of I_0 were then fit to Eq. 5-20 with L = 2.2 μm, x_1^2 given by Eq. 5-11, and c given by Eq. 5-9. The free parameters were D and b = κd. A representative plot showing the experimental values and their best fits to this theoretical form is shown in Fig. 5-8. The best fit values of the free parameters for the three data sets were averaged to give D = 6.3 ± 1.1 $\mu m^2 sec^{-1}$ and b = 0.026 ± 0.001 $\mu m^3 \mu W^{-1} sec^{-1}$. In all cases, the parameter c was calculated by using Eq. 5-9 and the best-fit values of b and D

along with the known values of L and I_0. The maximum value of c was 2.0, validating the use of single-exponential fits and Eq. 5-11 (see above).

The diffusion coefficient obtained by TIR-CP, 6.3 ± 1.1 $\mu m^2 sec^{-1}$, agrees well with that measured by confocal-FRAP on the same system in our laboratory (data not shown). The measured coefficient also agrees well with literature values for GFP diffusion in *E. coli* cells. The most commonly referenced value,[131] 7.7 ± 2.5 $\mu m^2 s^{-1}$, was obtained by using confocal FRAP. Similar studies found D = 6.1 ± 2.4 $\mu m^2 s^{-1}$ [132] and 3.2 $\mu m^2 s^{-1}$.[123] Other groups have measured diffusion coefficients of GFP fusion proteins in *E. coli* that are also consistent with our value, such as TorA-GFP (30 kDa), 9.0 ± 2.1 $\mu m^2 s^{-1}$;[133] cMBP-GFP (72 kDa), 2.5 ± 0.6 $\mu m^2 s^{-1}$;[131] and CheY-GFP (40 kDa), 4.6 ± 0.8 $\mu m^2 s^{-1}$.[117]

The b parameter from the fits contains information about the propensity for GFP to photobleach, since b = κd. For our experimental conditions, d ≈ 0.1 μm [134] and b = 0.026 ± 0.001 $\mu m^3 \mu W^{-1} sec^{-1}$. Taken together, these values yield κ ≈ 0.26 $\mu m^2 \mu W^{-1} sec^{-1}$, which is comparable to the value (0.21 ± 0.01 $\mu m^2 \mu W^{-1} sec^{-1}$) for the immobilized GFP (Fig. 5-6).

Several GFP variants have been reported to reversibly photobleach.[102,124] To assess the degree of reversible photobleaching that might be occurring, confocal microscopy was used to bleach GFP throughout entire *E coli* cells and to monitor the intracellular fluorescence over time. Post-bleach fluorescence recovery was not observed, indicating that GFP is irreversibly photobleached over the time scale of the experiment (data not shown).

5.4.5 Effects of Osmotic Shock

To further confirm the validity of the new technique, we measured GFP diffusion in sorbitol-treated cells. Sorbitol, and other forms of osmotic shock, increase the intracellular concentration of macromolecules, which significantly decreases protein mobility.[108] Data from confocal FRAP experiments show that GFP diffusion in *E. coli* cells decreases to 0.94 ± 0.55 $\mu m^2 sec^{-1}$ in 392 milliosmolal solution.[132] A similar study reported a diffusion coefficient of 1.8 $\mu m^2 sec^{-1}$ in a 370 milliosmolal buffer.[123] For TIR-CP experiments, three complete data sets were collected for GFP in *E. coli* exposed to a 390 milliosmolal sorbitol buffer and analyzed as described above. The diffusion coefficient decreased from 6.3 ± 1.1 $\mu m^2 sec^{-1}$, to 3.05 ± 1.0 $\mu m^2 sec^{-1}$ in a sorbitol buffer.

5.5 Discussion

As shown here, total internal reflection illumination with continuous photobleaching (TIR-CP) can be used to monitor the translational mobility of fluorescent molecules within small structures that are only slightly larger than optical resolution.   The structures are deposited on a surface such that the evanescent wave generated by internal reflection continuously photobleaches only those fluorescent molecules very near the surface (Fig. 5-1). The resulting fluorescence decay curves depend on two competing processes: photobleaching and diffusion. At low excitation intensities, the propensity for photobleaching determines the rate and shape of the fluorescence decay curves. At higher excitation intensities, the diffusion rate of the fluorescent molecules across the

length of the small structure also affects the fluorescence decay curves. By examination of the fluorescence decay as a function of the excitation intensity, the diffusion coefficient of the fluorescent molecules within the small structures can be determined. This new method was demonstrated by measuring the diffusion coefficient of GFP in *E. coli*. The measured diffusion coefficient agreed with those measured by different methods.[117,123,131-133] In addition, data were acquired for GFP in *E. coli* subjected to an osmotically stressed environment, and analysis of these data reported lower GFP diffusion coefficients. This result is also consistent with data obtained by different methods.[123,132]

As described in the Introduction, other techniques have also been developed for examining molecular mobility in small biological structures, particularly *E. coli*. The most well developed of these methods is confocal-FRAP. Two advantages of TIR-CP as compared to confocal-FRAP are that the evanescent illumination confines photobleaching to a smaller fraction of the *E. coli* volume and the simplicity of the required instrumentation. TIR-CP also avoids the complication of aligning a very small focused laser beam within the small structure of interest (e.g., *E. coli*). However, for both methods, either the shape of the *E. coli* must be approximated to obtain analytical theoretical forms for data analysis or the exact geometry of a given *E. coli* must be measured and used in numerical simulations.

Confocal-FRAP monitors diffusion on a cell-by-cell basis, and, although data analysis is somewhat tedious, this method can provide diffusion coefficient histograms and correlations of intracellular molecular mobility with other cellular

characteristics. In TIR-CP as described here, a few cells rather than one are in the observed volume. However, the possibility exists of generalizing TIR-CP by using a fast EMCCD camera and subsequent imaging. Because the evanescent intensity varies as a function of position (Fig 5-4), the entire range of intensity-dependent decay curves could be acquired from a single time-dependent image sequence. Given a dilute enough density of adherent cells, this method might also provide histograms of apparent diffusion coefficients, as well as correlations of mobility with other cellular properties.

A third method, TIR-FRAP, has also been used to measure the cytoplasmic mobility of fluorescent molecules close to the inner leaflet of membranes of surface adherent, large eukaryotic cells.[124] TIR-FRAP has potential for being generalized to small cells such as *E. coli*. TIR has also been combined with FCS to monitor translational mobility close to surface-adsorbed model membranes.[135] TIR-FCS could, in principle, measure membrane-local diffusion coefficients in small structures. Such potential TIR-FCS measurements, however, would suffer because the pool of unbleached molecules is limited in small structures. In addition, it has been demonstrated that the apparent diffusion coefficients of proteins close to membrane surfaces are significantly reduced due to hydrodynamic effects.[136] Such potential measurements, therefore, would not directly report the overall apparent diffusion coefficient of fluorescent molecules throughout the structure.

We have used TIR-CP to examine protein diffusion in *E. coli*. This new method may also be applicable to other small biological structures such as

phospholipid vesicles, isolated synaptic vesicles, and isolated organelles (e.g.,

mitochondria). Furthermore, the technique shows promise for use over a wide

range of protein concentrations. While FCS is strictly limited to very low

concentrations, confocal-FRAP becomes challenging at low concentrations due

to poor signal-to-noise. In contrast, TIR-CP is applicable both to low and high

concentrations, because of the range of excitation intensities available. Analysis

at these concentrations may be important as it has been reported that the level of

protein expression in *E. coli* may affect the protein diffusion coefficient.[131] In

addition, controlling the level of one protein while monitoring the diffusion

coefficient of another protein might reveal not only concentration-dependent

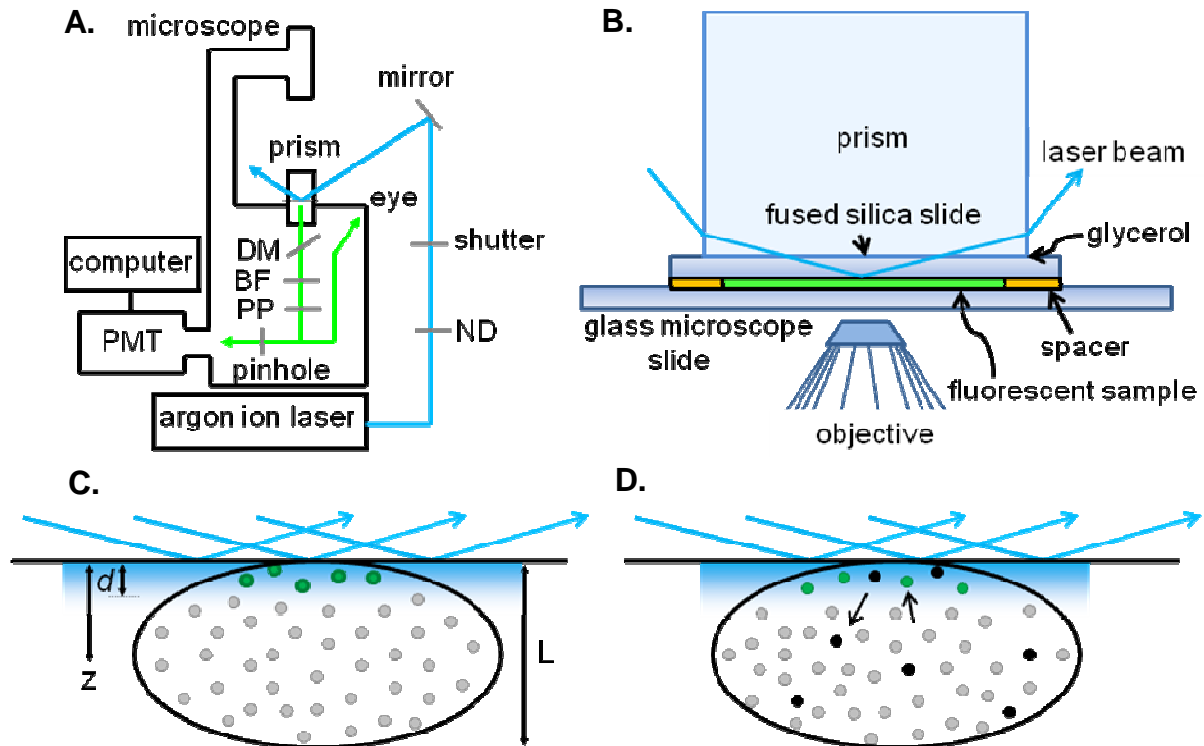diffusion, but also provide insight about protein-protein interactions.

5.5 Figures

Figure 5-1. Through-prism total internal reflection with continuous photobleaching (TIR-CP). (A) Schematic of the instrumentation: PMT, photomultiplier tube; BF, barrier filter; DM, dichroic mirror; PP, polarization paper (see Methods); ND, neutral density filter. (B) A fused silica prism is optically coupled with glycerol to a sandwich made from fused silica and microscope slides. (C) Small structures that extend a distance L into the solution are attached to the lower surface of the fused silica slide. A laser beam is internally reflected at the interface of the fused silica slide and the internal solution of the sandwich to create an evanescent field whose intensity decays exponentially with distance, z, from the interface. The characteristic distance of this decay, d, is much smaller than L. Molecules contained within the small structure do not fluoresce (grey circles) until they diffuse into the surface-associated evanescent field. (D) Eventually, the fluorescent molecules (green circles) are permanently bleached (black circles) by exposure to the evanescent field. At low excitation intensities, the decay of the evanescently excited fluorescence with time is dominated by the propensity for photobleaching. At high excitation intensities, the decay of fluorescence with time is dominated by diffusion through the length of the small structure.
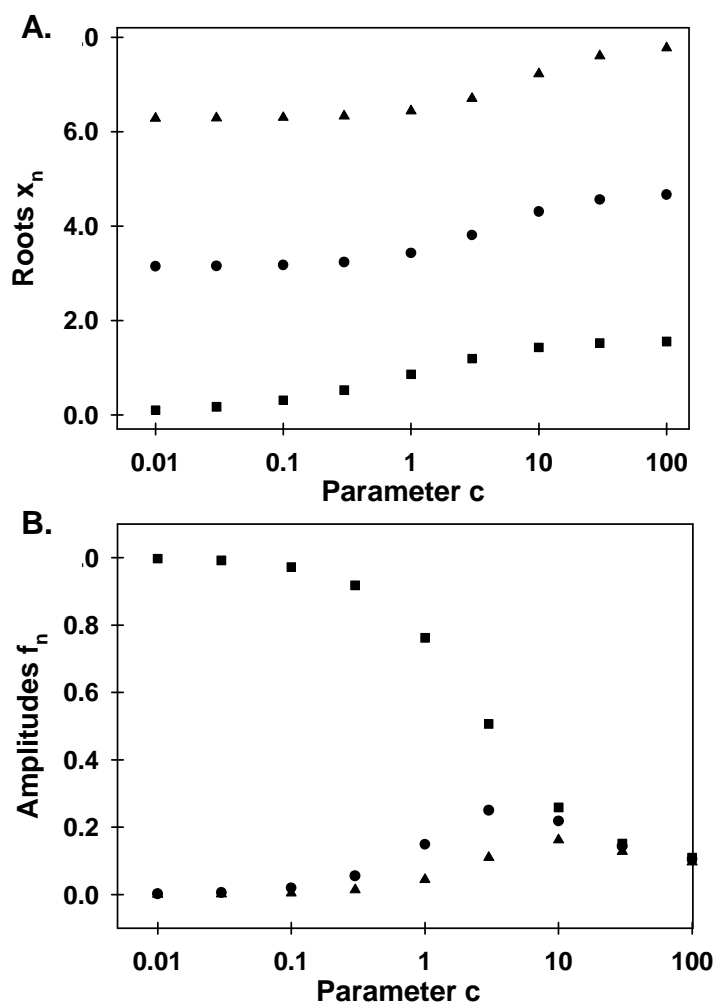
Figure 5-2. Parameters x_n and f_n. Fluorescence decay during continuous photobleaching is described by an infinite sum of exponentials with rates $D(x_n/L)^2$, where D is the diffusion coefficient and L is the cell length (Eqs. 9 and 19). The parameters x_n are a discrete, countably-infinite set of values that are determined by the value of the parameter c. Parameter c depends on experimental conditions (D, L, the depth of the evanescent wave d, and the bleaching propensity κ). Most importantly, c is proportional to the excitation intensity I_0. The parameters f_n are amplitudes associated with the different exponentially decaying terms and are defined, in general, by the values of c, L, d and x_n. (A) The values of x_1 (■), x_2 (●) and x_3 (▲) were calculated numerically as a function of c by using Eq. 9. (B) The values of f_1 (■), f_2 (●) and f_3 (▲) were calculated by using Eq. 19 with L = 2.2 μm and d = 0.1 μm. At low c values, the first amplitude, f_1 is much larger than the others and the fluorescence decay can be approximated as a single exponential with rate $D(x_1/L)^2$. In addition, for many experimental conditions, the terms associated with rates having n > 1 decay too rapidly to affect the observed fluorescence decay.

123

A.

B.
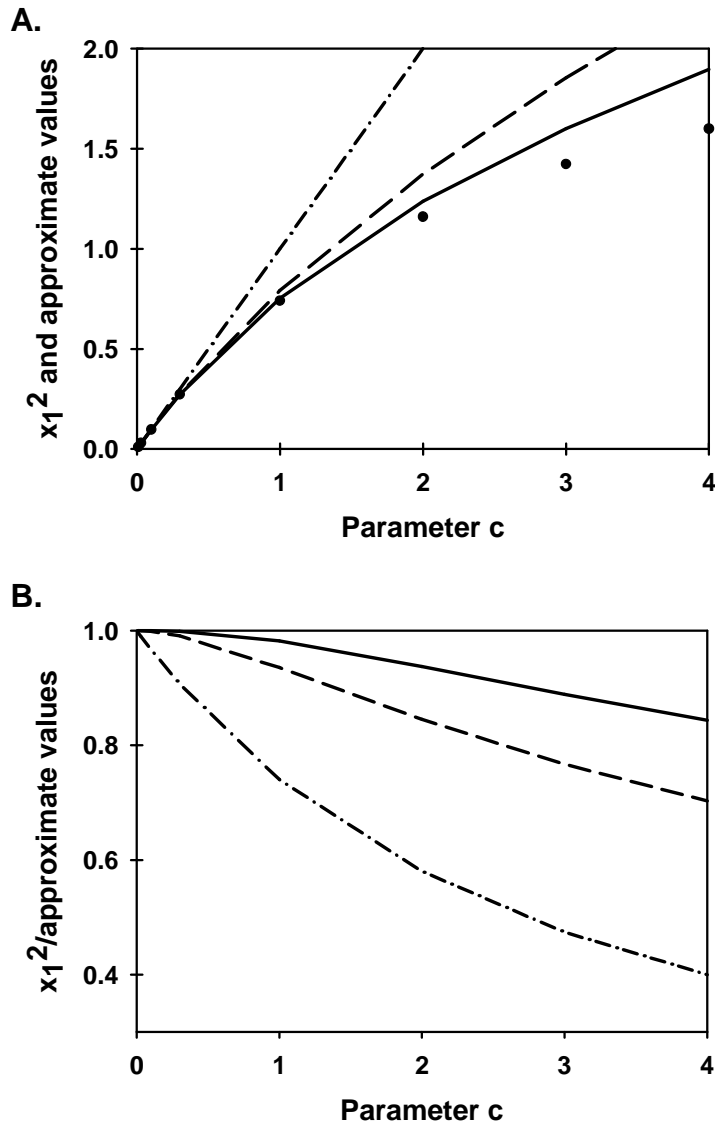
Figure 5-3. Accuracy of x_1^2 approximations. Parameters x_n and c are described in the caption to Fig. 2. (A) This plot shows x_1^2 approximated as c (- · - ·) (see text), as Eq. 10 (- - -) and as Eq. 11 (—) compared with the numerically determined values (•) as a function of the parameter c (Eq. 5-9). (B) This plot shows the corresponding ratios of the actual x_1^2 values divided by the approximations as a function of c, such that values close to one represent an accurate approximation. As shown, Eq. 11 is a good approximation for the experimental conditions used in this work ($c \leq 2$).

A.

B.
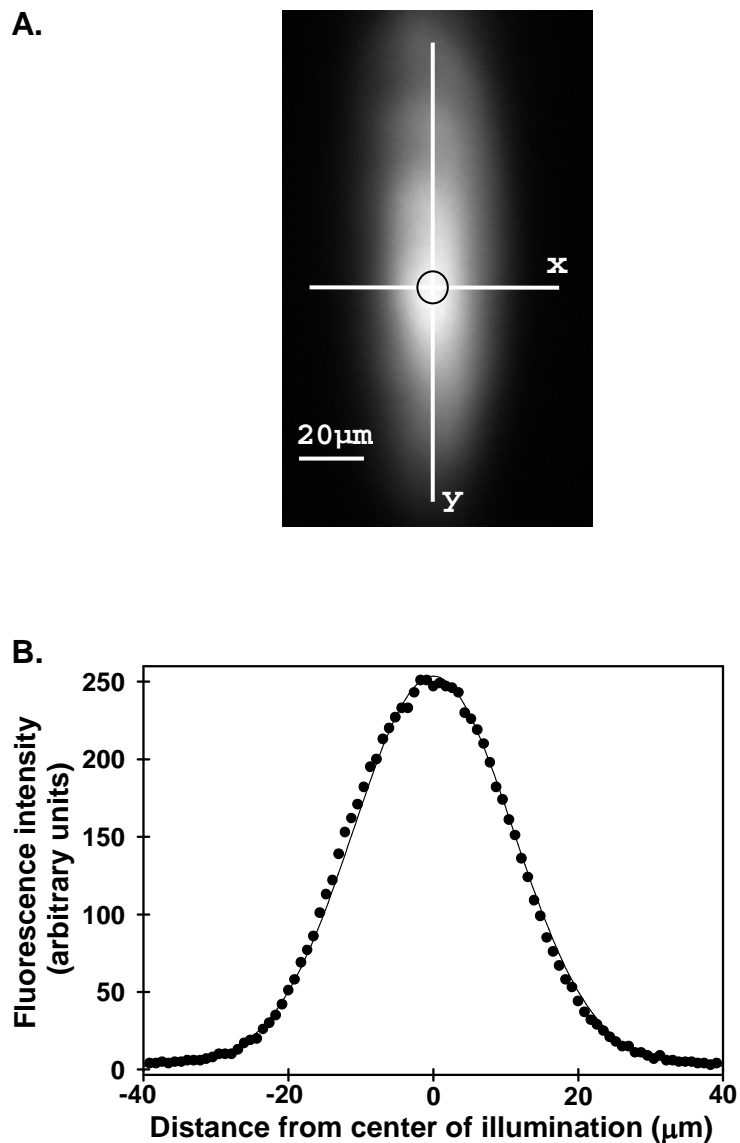
Figure 5-4. Spatial profile of the evanescent illumination in the x-y plane.
(A) Evanescently excited Alexa Fluor 488 dye was non-specifically adsorbed to the surface of fused silica. The circle indicates the observed area within the pinhole. (B) Data were obtained by slicing images with x = 0 or y = 0 and plotting the corresponding pixel intensities as a function of distance. A representative slice with y = 0 and its best fit to Eq. 24 is also shown.
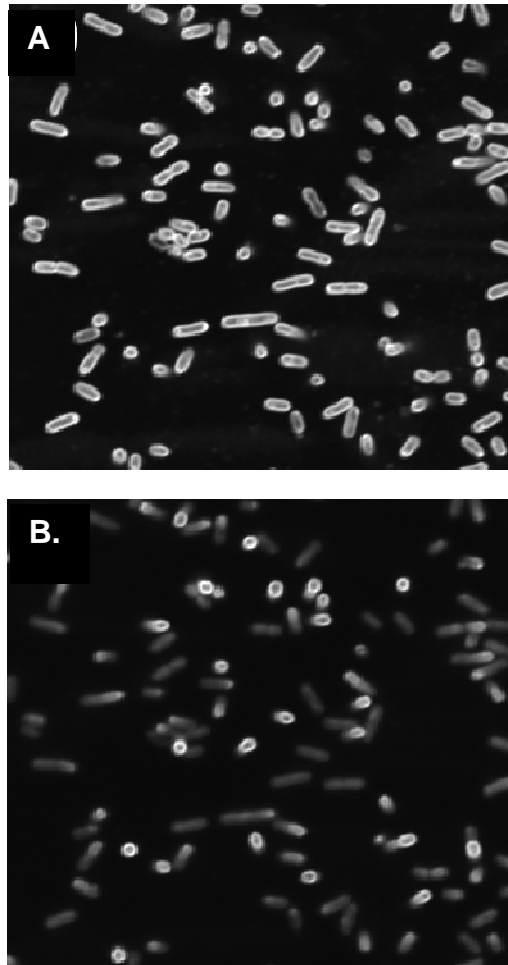
Figure 5-5. Cell length, L. *E. coli* cells expressing GFP were stained with FM1-43 dye and attached to poly-L-lysine coated coverslips. Using a confocal microscope, images were collected in the x-y plane by moving away from the coverslip in 0.2-µm increments. The images are 36.6 µm x 36.6 µm and were collected (A) 1.4 µm and (B) 2.4 µm from the coverslip.
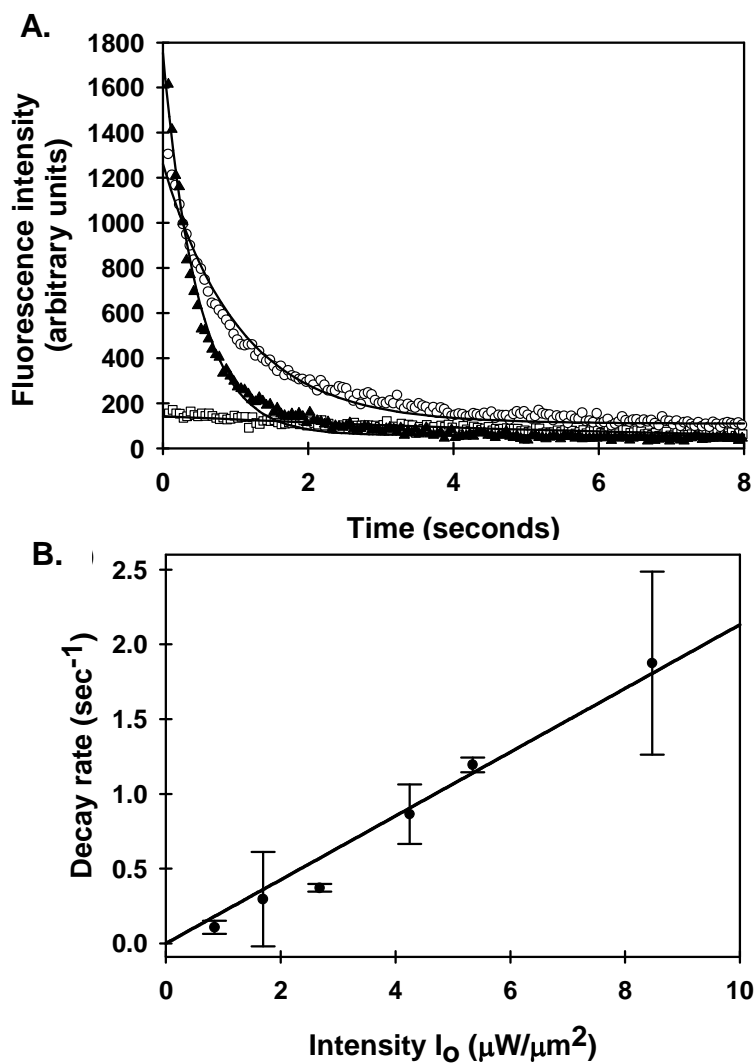
Figure 5-6. Photobleaching propensity for immobilized GFP. (A) Representative evanescently excited fluorescence decay curves are shown for purified GFP immobilized on fused silica. The excitation intensities I_0 were 0.9 (□), 4.3 (○), and 8.5 (▲) $\mu W/\mu m^2$. The solid curves show the best fits of the data to Eq. 5-28. (B) The average fluorescence decay rate R is a linear function of the excitation intensity, I_0.
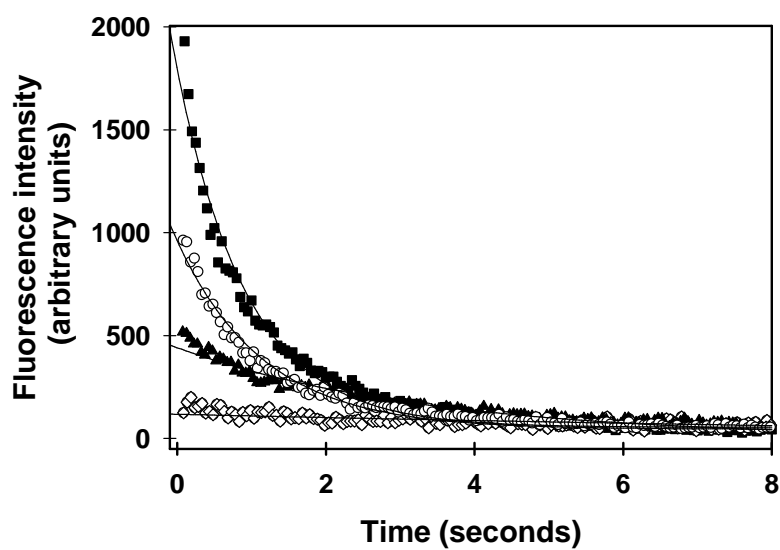
Figure 5-7. Representative evanescently excited fluorescence decay curves for GFP in _E. coli c_ells. The excitation intensities were 5 (◊), 35 (▲), 87 (○), and 220 (■) µW/ µm^2. The solid curves show the best fit of the data to Eq. 5-28.
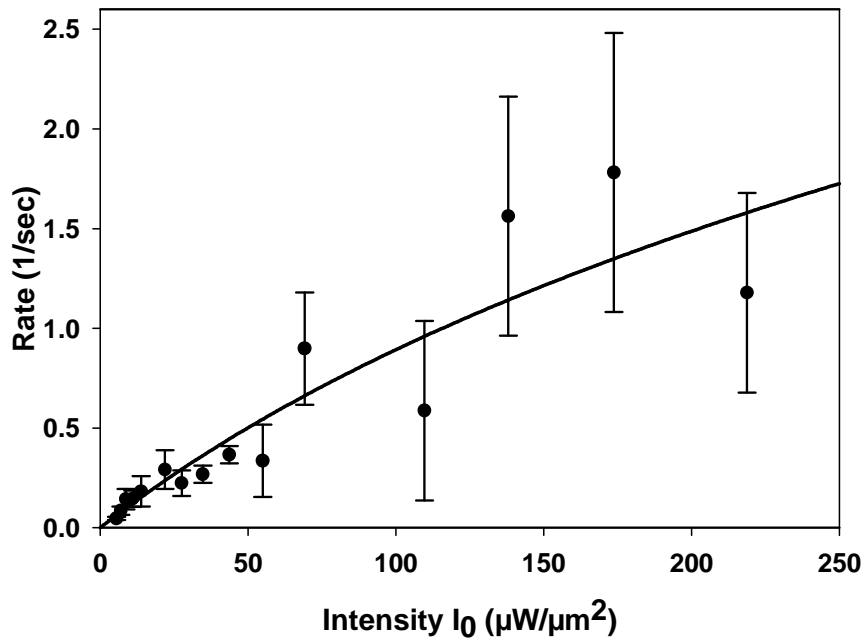
Figure 5-8. Diffusion coefficient D and parameter b for GFP in *E. coli*. Plot shows data from one data set in which the measured decay rates R are plotted as a function of the excitation intensities, I_0. Rates represent an average of at least three measurements with uncertainties given as standard deviations. The line shows the best fit to Eq. 20 with x_1^2 given by Eq. 11, L fixed at 2.2 μm, and free parameters D and b = dκ. In this case, D = 6.4 μm^{2}sec^{-1}, b = 0.025 μm^{3}μW^{-1}sec^{-1}, the maximum value of c = 2.0 (Eq. 5-9), and correlation coefficient was 0.93.

References

1. Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol.* **192**, 189-221.

2. Cayley, D. S., Guttman, H. J. & Record, M. T., Jr. (2000). Biophysical characterization of changes in amounts and activity of *Escherichia coli* cell and compartment water and turgor pressure in response to osmotic stress. *Biophys. J.* **78**, 1748-64.

3. Minton, A. P. (2001). The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* **276**, 10577-80.

4. Minton, A. P. (2005). Influence of macromolecular crowding upon the stability and state of association of proteins: predictions and observations. *J. Pharm. Sci.* **94**, 1668-75.

5. Lavalette, D., Tetreau, C., Tourbez, M. & Blouquit, Y. (1999). Microscopic viscosity and rotational diffusion of proteins in a macromolecular environment. *Biophys. J.* **76**, 2744-51.

6. Hatters, D. M., Minton, A. P. & Howlett, G. J. (2002). Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J. Biol. Chem.* **277**, 7824-30.

7. Munishkina, L. A., Cooper, E. M., Uversky, V. N. & Fink, A. L. (2004). The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.* **17**, 456-64.

8. del Alamo, M., Rivas, G. & Mateu, M. G. (2005). Effect of macromolecular crowding agents on human immunodeficiency virus type 1 capsid protein assembly in vitro. *J. Virol.* **79**, 14271-81.

9. Zorrilla, S., Rivas, G., Acuna, A. U. & Lillo, M. P. (2004). Protein self-association in crowded protein solutions: a time-resolved fluorescence polarization study. *Protein Sci.* **13**, 2960-9.

10. Ghahghaei, A., Rekas, A., Price, W. E. & Carver, J. A. (2007). The effect of dextran on subunit exchange of the molecular chaperone alphaA-crystallin. *Biochim. Biophys. Acta.* **1774**, 102-11.

11. Bolis, D., Politou, A. S., Kelly, G., Pastore, A. & Temussi, P. A. (2004). Protein stability in nanocages: a novel approach for influencing protein stability by molecular confinement. *J Mol. Biol.* **336**, 203-12.

12. Srere, P. A. & Ovadi, J. (1990). Enzyme-enzyme interactions and their

metabolic role. *FEBS Lett.* **268**, 360-4.

13. Ovadi, J., Orosz, F. & Hollan, S. (2004). Functional aspects of cellular microcompartmentation in the development of neurodegeneration: mutation induced aberrant protein-protein associations. *Mol. Cell. Biochem.* **256-257**, 83-93.

14. Ovadi, J. & Saks, V. (2004). On the origin of intracellular compartmentation and organized metabolic systems. *Mol. Cell. Biochem.* **256-257**, 5-12.

15. Dedmon, M. M., Patel, C. N., Young, G. B. & Pielak, G. J. (2002). FlgM gains structure in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12681-4.

16. McNulty, B. C., Young, G. B. & Pielak, G. J. (2006). Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder. *J Mol. Biol.* **355**, 893-7.

17. Moran-Zorzano, M. T., Alonso-Casajus, N., Munoz, F. J., Viale, A. M., Baroja-Fernandez, E., Eydallin, G. & Pozueta-Romero, J. (2007). Occurrence of more than one important source of ADPglucose linked to glycogen biosynthesis in *Escherichia coli* and *Salmonella*. *FEBS Lett.* **581**, 4423-9.

18. Homchaudhuri, L., Sarma, N. & Swaminathan, R. (2006). Effect of crowding by dextrans and Ficolls on the rate of alkaline phosphatase-catalyzed hydrolysis: a size-dependent investigation. *Biopolymers* **83**, 477-86.

19. Sasaki, Y., Miyoshi, D. & Sugimoto, N. (2007). Regulation of DNA nucleases by molecular crowding. *Nucleic Acids Res.* **35**, 4086-93.

20. Olsen, S. N., Ramlov, H. & Westh, P. (2007). Effects of osmolytes on hexokinase kinetics combined with macromolecular crowding: test of the osmolyte compatibility hypothesis towards crowded systems. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **148**, 339-45.

21. Jiang, M. & Guo, Z. (2007). Effects of macromolecular crowding on the intrinsic catalytic efficiency and structure of enterobactin-specific isochorismate synthase. *J. Am. Chem. Soc.* **129**, 730-1.

22. Kozer, N., Kuttner, Y. Y., Haran, G. & Schreiber, G. (2007). Protein-protein association in polymer solutions: from dilute to semidilute to concentrated. *Biophys. J.* **92**, 2139-49.

23. Kuttner, Y. Y., Kozer, N., Segal, E., Schreiber, G. & Haran, G. (2005). Separating the contribution of translational and rotational diffusion to protein association. *J. Am. Chem. Soc.* **127**, 15138-44.

24. Luby-Phelps, K., Castle, P. E., Taylor, D. L. & Lanni, F. (1987). Hindered diffusion of inert tracer particles in the cytoplasm of mouse 3T3 cells. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4910-3.

25. Seksek, O., Biwersi, J. & Verkman, A. S. (1997). Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J. Cell Biol.* **138**, 131-42.

26. Williams, S. P., Haggie, P. M. & Brindle, K. M. (1997). 19F NMR measurements of the rotational mobility of proteins *in vivo. Biophys. J.* **72**, 490-8.

27. Williams, S. P., Fulton, A. M. & Brindle, K. M. (1993). Estimation of the intracellular free ADP concentration by 19F NMR studies of fluorine-labeled yeast phosphoglycerate kinase *in vivo. Biochemistry* **32**, 4895-902.

28. Haggie, P. M. & Brindle, K. M. (1999). Mitochondrial citrate synthase is immobilized *in vivo. J. Biol. Chem.* **274**, 3941-5.

29. Serber, Z., Keatinge-Clay, A. T., Ledwidge, R., Kelly, A. E., Miller, S. M. & Dotsch, V. (2001). High-resolution macromolecular NMR spectroscopy inside living cells. *J. Am. Chem. Soc.* **123**, 2446-7.

30. Serber, Z. & Dotsch, V. (2001). In-cell NMR spectroscopy. *Biochemistry* **40**, 14317-23.

31. Pielak, G. J., Li, C., Miklos, A. C., Schlesinger, A. P., Slade, K. M., Wang, G. & Zigoneanu, I. G. (2009). Protein nuclear magnetic resonance under physiological conditions. *Biochemistry* **48**, 226-34.

32. Reckel, S., Lohr, F. & Dötsch, V. (2005). In-cell NMR spectroscopy. *Chembiochem.* **6**, 1601-6.

33. Dix, J. A. & Verkman, A. S. (2008). Crowding effects on diffusion in solutions and cells. *Annu. Rev. Biophys.* **37**, 247-63.

34. Golding, I. & Cox, E. C. (2006). Physical nature of bacterial cytoplasm. *Phys. Rev. Lett.* **96**, 098102.

35. Lavalette, D., Hink, M. A., Tourbez, M., Tetreau, C. & Visser, A. J. (2006). Proteins as micro viscosimeters: Brownian motion revisited. *Eur. Biophys. J.* **35**, 517-22.

36. Zhou, H. X., Rivas, G. & Minton, A. P. (2008). Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37**, 375-97.

37.   Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B. & Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli. J. Bacteriol.* **181**, 197-203.

38.   Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* **16**, 1055-69.

39.   Reits, E. A. & Neefjes, J. J. (2001). From fixed to FRAP: measuring protein mobility and activity in living cells. *Nat. Cell Biol.* **3**, E145-7.

40.   Tsien, R. Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509-44.

41.   Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392-5.

42.   Liu, H. S., Jan, M. S., Chou, C. K., Chen, P. H. & Ke, N. J. (1999). Is green fluorescent protein toxic to the living cells? *Biochem. Biophys. Res. Commun.* **260**, 712-7.

43.   Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995). Improved green fluorescence. *Nature* **373**, 663-4.

44.   Yang, T. T., Cheng, L. & Kain, S. R. (1996). Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* **24**, 4592-3.

45.   Beckwith, J. (1987). The Lactose Operon. In Escherichia coli *and* Salmonella typhimurium*: Cellular and Molecular Biology* (Neidhardt, F., ed.), Vol. 2, pp. 1444-1452. 2 vols. American Society for Microbiology, Washington D.C.

46.   Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol. Biol.* **189**, 113-30.

47.   Harley, C. B. & Reynolds, R. P. (1987). Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**, 2343-61.

48.   Ikeda, R. A., Ligman, C. M. & Warshamana, S. (1992). T7 promoter contacts essential for promoter activity *in vivo. Nucleic Acids Res.* **20**, 2517-24.

49.   Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**, 2237-55.

50. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. *J. Bacteriol.* **177**, 4121-30.

51. Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J. & Tsien, R. Y. (2002). New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and *in vivo*: synthesis and biological applications. *J. Am. Chem. Soc.* **124**, 6063-76.

52. Cavagnero, S. & Jungbauer, L. M. (2005). Painting protein misfolding in the cell in real time with an atomic-scale brush. *Trends Biotechnol.* **23**, 157-62.

53. Griffin, B. A., Adams, S. R., Jones, J. & Tsien, R. Y. (2000). Fluorescent labeling of recombinant proteins in living cells with FlAsH. *Methods Enzymol.* **327**, 565-78.

54. Griffin, B. A., Adams, S. R. & Tsien, R. Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269-72.

55. Luebke, K. J. (1998). A FLASH of insight into cellular chemistry: genetically encoded labels for protein visualization *in vivo. Chem. Biol.* **5**, R317-22.

56. Ignatova, Z. & Gierasch, L. M. (2004). Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 523-8.

57. Kay, L. E. (1995). Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. *Prog. Biophys. Mol. Biol.* **63**, 277-99.

58. Khan, F., Kuprov, I., Craggs, T. D., Hore, P. J. & Jackson, S. E. (2006). 19F NMR studies of the native and denatured states of green fluorescent protein. *J. Am. Chem. Soc.* **128**, 10729-37.

59. Jackson, J. C., Hammill, J. T. & Mehl, R. A. (2007). Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **129**, 1160-6.

60. Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C. & Mehl, R. A. (2007). Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat. Protoc.* **2**, 2601-7.

61. McNulty, B. C., Tripathy, A., Young, G. B., Charlton, L. M., Orans, J. & Pielak, G. J. (2006). Temperature-induced reversible conformational

134

change in the first 100 residues of alpha-synuclein. *Protein Sci.* **15**, 602-8.

62. Charlton, L. M., Barnes, C. O., Li, C., Orans, J., Young, G. B. & Pielak, G. J. (2008). Residue-level interrogation of macromolecular crowding effects on protein stability. *J. Am. Chem. Soc.* **130**, 6826-30.

63. Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T. & Dahlquist, F. W. (1997). The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. *Nat. Struct. Biol.* **4**, 285-91.

64. Ishizuka, H., Hanamura, A., Inada, T. & Aiba, H. (1994). Mechanism of the down-regulation of cAMP receptor protein by glucose in *Escherichia coli*: role of autoregulation of the crp gene. *EMBO J.* **13**, 3077-82.

65. Ishizuka, H., Hanamura, A., Kunimura, T. & Aiba, H. (1993). A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli. Mol. Microbiol.* **10**, 341-50.

66. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition state of folding. *Biochemistry* **30**, 10436-43.

67. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-35.

68. Itzhaki, L. S., Neira, J. L. & Fersht, A. R. (1997). Hydrogen exchange in chymotrypsin inhibitor 2 probed by denaturants and temperature. *J Mol. Biol.* **270**, 89-98.

69. Neira, J. L., Itzhaki, L. S., Ladurner, A. G., Davis, B., de Prat Gay, G. & Fersht, A. R. (1997). Following co-operative formation of secondary and tertiary structure in a single protein module. *J Mol. Biol.* **268**, 185-97.

70. Li, C., Charlton, L. M., Lakkavaram, A., Seagle, C., Wang, G., Young, G. B., Macdonald, J. M. & Pielak, G. J. (2008). Differential dynamical effects of macromolecular crowding on an intrinsically disordered protein and a globular protein: implications for in-cell NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 6310-1.

71. Surguchov, A. (2008). Molecular and cellular biology of synucleins. *Int. Rev. Cell. Mol. Biol.* **270**, 225-317.

72. Nordstrom, K. & Austin, S. J. (1989). Mechanisms that contribute to the stable segregation of plasmids. *Annu. Rev. Genet.* **23**, 37-69.

73.   Austin, S. & Nordstrom, K. (1990). Partition-mediated incompatibility of bacterial plasmids. *Cell* **60**, 351-4.

74.   Dzivenu, O. K., Park, H. H. & Wu, H. (2004). General co-expression vectors for the overexpression of heterodimeric protein complexes in *Escherichia coli. Protein Expr. Purif.* **38**, 1-8.

75.   Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P. & Bradbury, A. R. (2007). Plasmid incompatibility: more compatible than previously thought? *Protein Eng. Des. Sel.* **20**, 309-13.

76.   Voet, D. & Voet, J. (2004). *Biochemistry.* 3rd edit (Harris, D., Ed.), John Wiley & Sons, Danvers, MA.

77.   Steen, R., Dahlberg, A. E., Lade, B. N., Studier, F. W. & Dunn, J. J. (1986). T7 RNA polymerase directed expression of the *Escherichia coli* rrnB operon. *EMBO J.* **5**, 1099-103.

78.   Donachie, W. & Robinson, A. (1987). Cell Division: Paramter Values and the Process. In Escherichia coli *and* Salmonella typhimurium*: Cellular and Molecular Biology* (Neidhardt, F., ed.), Vol. 2, pp. 1578-1591. 2 vols. American Society for Microbiology, Washington D.C.

79.   Koch, A. L. (1987). The Variability and Individuality of the Bacterium. In Escherichia coli *and* Salmonella typhimurium*: Cellular and Molecular Biology* (Neidhardt, F., ed.), Vol. 2, pp. 1606-1614. 2 vols. American Society for Microbiology, Washington D.C.

80.   Lin-Chao, S., Chen, W. T. & Wong, T. T. (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Mol. Microbiol.* **6**, 3385-93.

81.   Burz, D. S., Dutta, K., Cowburn, D. & Shekhtman, A. (2006). In-cell NMR for protein-protein interactions (STINT-NMR). *Nat. Protoc.* **1**, 146-52.

82.   Mullineaux, C. W., Nenninger, A., Ray, N. & Robinson, C. (2006). Diffusion of green fluorescent protein in three cell environments in *Escherichia coli. J. Bacteriol.* **188**, 3442-8.

83.   Konopka, M. C., Shkel, I. A., Cayley, S., Record, M. T. & Weisshaar, J. C. (2006). Crowding and confinement effects on protein diffusion *in vivo. J. Bacteriol.* **188**, 6115-23.

84.   van den Bogaart, G., Hermans, N., Krasnikov, V. & Poolman, B. (2007). Protein mobility and diffusive barriers in *Escherichia coli:* consequences of osmotic stress. *Mol. Microbiol.* **64**, 858-71.

85.   Boos, W. & Shuman, H. (1998). Maltose/maltodextrin system of

Escherichia coli: transport, metabolism, and regulation. *Microbiol. Mol. Biol. Rev.* **62**, 204-29.

86.   Stevens, F. C. (1983). Calmodulin: an introduction. *Can. J. Biochem. Cell Biol.* **61**, 906-10.

87.   Uversky, V. N. (2008). Alpha-synuclein misfolding and neurodegenerative diseases. *Curr. Protein Pept. Sci.* **9**, 507-40.

88.   Huang, C., Ren, G., Zhou, H. & Wang, C. C. (2005). A new method for purification of recombinant human alpha-synuclein in *Escherichia coli*. *Protein Expr. Purif.* **42**, 173-7.

89.   Sahara, N., Maeda, S. & Takashima, A. (2008). Tau oligomerization: a role for tau aggregation intermediates linked to neurodegeneration. *Curr. Alzheimer Res.* **5**, 591-8.

90.   Yang, F., Moss, L. G. & Phillips, G. N., Jr. (1996). The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246-51.

91.   Gopalakrishna, R. & Anderson, W. B. (1982). Ca2+-induced hydrophobic site on calmodulin: application for purification of calmodulin by phenyl-Sepharose affinity chromatography. *Biochem. Biophys. Res. Commun.* **104**, 830-6.

92.   Conway, K. A., Lee, S. J., Rochet, J. C., Ding, T. T., Harper, J. D., Williamson, R. E. & Lansbury, P. T., Jr. (2000). Accelerated oligomerization by Parkinson's disease linked alpha-synuclein mutants. *Ann. N. Y. Acad. Sci.* **920**, 42-5.

93.   Medintz, I. L., Goldman, E. R., Lassman, M. E. & Mauro, J. M. (2003). A fluorescence resonance energy transfer sensor based on maltose binding protein. *Bioconjugate Chem.* **14**, 909-18.

94.   Franz, C., Durussel, I., Cox, J. A., Schafer, B. W. & Heizmann, C. W. (1998). Binding of Ca^{2+} and Zn^{2+} to human nuclear S100A2 and mutant proteins. *J. Biol. Chem.* **273**, 18826-34.

95.   Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-75.

96.   Bainer, R., Park, H. & Cluzel, P. (2003). A high-throughput capillary assay for bacterial chemotaxis. *J. Microbiol. Methods* **55**, 315-9.

97.   Birdsell, D. C. & Cota-Robles, E. H. (1967). Production and ultrastructure of lysozyme and ethylenediaminetetraacetate-lysozyme spheroplasts of *Escherichia coli*. *J. Bacteriol.* **93**, 427-37.

98.	Thorstenson, Y. R., Zhang, Y., Olson, P. S. & Mascarenhas, D. (1997). Leaderless polypeptides efficiently extracted from whole cells by osmotic shock. *J. Bacteriol.* **179**, 5333-9.

99.	Shevchik, V. E., Condemine, G. & Robert-Baudouy, J. (1994). Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity. *EMBO J.* **13**, 2007-12.

100.	Ren, G., Wang, X., Hao, S., Hu, H. & Wang, C. C. (2007). Translocation of alpha-synuclein expressed in *Escherichia coli. J. Bacteriol.* **189**, 2777-86.

101.	Swaminathan, R., Bicknese, S., Periasamy, N. & Verkman, A. S. (1996). Cytoplasmic viscosity near the cell plasma membrane: translational diffusion of a small fluorescent solute measured by total internal reflection-fluorescence photobleaching recovery. *Biophys. J.* **71**, 1140-51.

102.	Sinnecker, D., Voigt, P., Hellwig, N. & Schaefer, M. (2005). Reversible photobleaching of enhanced green fluorescent proteins. *Biochemistry* **44**, 7085-94.

103.	Muramatsu, N. & Minton, A. P. (1988). Tracer diffusion of globular proteins in concentrated protein solutions. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2984-8.

104.	Perham, M., Stagg, L. & Wittung-Stafshede, P. (2007). Macromolecular crowding increases structural content of folded proteins. *FEBS Lett.* **581**, 5065-9.

105.	Slade, K. M., Steele, B. S., Pielak, G. J. & Thompson, N. L. (2009). Quantifying GFP diffusion in *Escherichia coli* by using continuous photobleaching with evanescent illumination. *J. Phys. Chem. B* **In press**.

106.	Ghaemmaghami, S. & Oas, T. G. (2001). Quantitative protein stability measurement *in vivo. Nat. Struct. Biol.* **8**, 879-82.

107.	Konopka, M. C., Weisshaar, J. C. & Record, M. T., Jr. (2007). Methods of changing biopolymer volume fraction and cytoplasmic solute concentrations for *in vivo* biophysical studies. *Methods Enzymol.* **428**, 487-504.

108.	Cayley, S. & Record, M. T., Jr. (2003). Roles of cytoplasmic osmolytes, water, and crowding in the response of Escherichia coli to osmotic stress: biophysical basis of osmoprotection by glycine betaine. *Biochemistry* **42**, 12596-609.

109.	Siegele, D. A. & Hu, J. C. (1997). Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8168-72.

110. Stock, J. B., Rauch, B. & Roseman, S. (1977). Periplasmic space in Salmonella typhimurium and Escherichia coli. *J. Biol. Chem.* **252**, 7850-61.

111. Thomas, J. D., Daniel, R. A., Errington, J. & Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli. Mol. Microbiol.* **39**, 47-53.

112. Terry, B. R., Matthews, E. K. & Haseloff, J. (1995). Molecular characterisation of recombinant green fluorescent protein by fluorescence correlation microscopy. *Biochem. Biophys. Res. Commun.* **217**, 21-7.

113. Mullineaux, C. W. (2007). Localization and mobility of bacterial proteins by confocal microscopy and fluorescence recovery after photobleaching. *Methods in Molecular Biology* **390**, 3-16.

114. Thompson, N. L., Lieto, A. M. & Allen, N. W. (2002). Recent advances in fluorescence correlation spectroscopy. *Current Opinion in Structural Biology* **12**, 634-41.

115. Rigler, R. & Elson, E. S., Eds. (2001). Fluorescence Correlation Spectroscopy, Theory and Applications. Springer Series in Chemical Physics, No. 65,. Springer Series in Chemical Physics. Heidelberg, Germany: Springer-Verlag.

116. Lippincott-Schwartz, J., Snapp, E. & Kenworthy, A. (2001). Studying protein dynamics in living cells. *Nature Reviews Molecular Cell Biology* **2**, 444-56.

117. Cluzel, P., Surette, M. & Leibler, S. (2000). An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science* **287**, 1652-5.

118. Peters, R., Brunger, A. & Schulten, K. (1981). Continuous fluorescence microphotolysis: A sensitive method for study of diffusion processes in single cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 962-966.

119. Brünger, A., Peters, R. & Schulten, K. (1984). Continuous fluorescence microphotolysis to observe later diffusion in membranes. Theoretical methods and applications. *J. Chem. Phys* **82**, 2147-2160.

120. Delon, A., Usson, Y., Derouard, J., Biben, T. & Souchier, C. (2006). Continuous photobleaching in vesicles and living cells: a measure of diffusion and compartmentation. *Biophysical Journal* **90**, 2548-62.

121. Wachsmuth, M., Weidemann, T., Müller, G., Hoffmann-Rohrer, U. W., Knoch, T. A., Waldeck, W. & Langowski, J. (2003). Analyzing intracellular

binding and diffusion with continuous fluorescence photobleaching. *Biophysical Journal* **84**, 3353-63.

122. Arkhipov, A., Hüve, J., Kahms, M., Peters, R. & Schulten, K. (2007). Continuous fluorescence microphotolysis and correlation spectroscopy using 4Pi microscopy. *Biophysical Journal* **93**, 4006-17.

123. van den Bogaart, G., Hermans, N., Krasnikov, V. & Poolman, B. (2007). Protein mobility and diffusive barriers in Escherichia coli: consequences of osmotic stress. *Molecular Microbiology* **64**, 858-71.

124. Swaminathan, R., Bicknese, S., Periasamy, N. & Verkman, A. S. (1996). Cytoplasmic viscosity near the cell plasma membrane: translational diffusion of a small fluorescent solute measured by total internal reflection-fluorescence photobleaching recovery. *Biophysical Journal* **71**, 1140-51.

125. Burghardt, T. & Thompson, N. (1984). Evanescent intensity of a focused Gaussian light beam undergoing total internal reflection in a prism. *Optical Engineering* **23**, 62-67.

126. Thompson, N. L., Burghardt, T. P. & Axelrod, D. (1981). Measuring surface dynamics of biomolecules by total internal reflection fluorescence with photobleaching recovery or correlation spectroscopy. *Biophysical Journal* **33**, 435-54.

127. Yang, F., Moss, L. G. & Phillips, G. N., Jr. (1996). The molecular structure of green fluorescent protein. *Nature Biotechnology* **14**, 1246-51.

128. Koch, A. L. (1987). The variability and individuality of the bacterium. In Escherichia coli *and* Salmonella typhimurium*: Cellular and Molecular Biology* (Neidhardt, F., ed.), Vol. 2, pp. 1608-1614. 2 vols. American Society for Microbiology, Washington D.C.

129. Cullum, J. & Vicente, M. (1978). Cell growth and length distribution in *Escherichia coli. Journal of Bacteriology* **134**, 330-7.

130. Baldwin, W. W., Sheu, M. J., Bankston, P. W. & Woldringh, C. L. (1988). Changes in buoyant density and cell size of *Escherichia coli* in response to osmotic shocks. *Journal of Bacteriology* **170**, 452-5.

131. Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B. & Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli. Journal of Bacteriology* **181**, 197-203.

132. Konopka, M. C., Shkel, I. A., Cayley, S., Record, M. T. & Weisshaar, J. C. (2006). Crowding and confinement effects on protein diffusion *in vivo. Journal of Bacteriology* **188**, 6115-23.

133. Mullineaux, C. W., Nenninger, A., Ray, N. & Robinson, C. (2006). Diffusion of green fluorescent protein in three cell environments in *Escherichia coli. Journal of Bacteriology* **188**, 3442-8.

134. Thompson, N. L. & Pero, J. (2005). Total internal reflection fluorescence microscopy: Applications in biophysics, in Fluorescence Spectroscopy in Biology: Advanced Methods and Their Applications to Membranes, Proteins, DNA and Cells. (Wolfbeis, O., Hof, M., Hutterer, R. & Fidler, V., eds.), pp. 79-103. Springer-Verlag, Berlin.

135. Starr, T. E. & Thompson, N. L. (2002). Fluorescence pattern photobleaching recovery for samples with multi-component diffusion. *Biophysical Chemistry* **97**, 29-44.

136. Pero, J. K., Haas, E. M. & Thompson, N. L. (2006). Size dependence of protein diffusion very close to membrane surfaces: measurement by total internal reflection with fluorescence correlation spectroscopy. *Journal of Physical Chemistry B* **110**, 10910-8.