UNIVERSITY OF CALIFORNIA,

Santa Barbara

The Effect of RNA Nucleic Acid Aptamers on the Aggregation of Green Fluorescent Protein

A Thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science in Biochemistry and Molecular Biology

by

Thomas Nguyen

Committee in charge:

Professor Irene Chen, Chair

Professor Kevin Plaxco

Professor Frederick Dahlquist

June 2017

The thesis of Thomas Nguyen is approved.

_____

Kevin Plaxco

_____

Frederick Dahlquist

_____

Irene Chen, Committee Chair

June 2017

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Irene Chen, for her unwavering support and mentorship throughout this project.

I would like to extend my thanks to everyone in the Chen Lab who offered me collegial guidance and support over the years. You guys truly taught me how to be an awesome scientist.

I would like to thank my roommates, Steve Lee and Kilian Komala, for being extremely friendly, smart, and fun people to hang out with. I thank you guys for keeping me sane throughout this project.

I would also like to dedicate this thesis to my family in Southern California. They have always supported and believed in me. Dad, thanks for teaching me the meaning of hard work. Mom, thanks for showing me your unconditional love. Monica, thanks for reminding me to work out and to have a good work-life balance. Steven, thanks for reminding me to follow my true goals in life.

Last but not least: Sasha Schaps- thanks for being you

Abstract

The Effect of RNA Nucleic Aptamers on the Aggregation of Green Fluorescent Protein

by

Thomas Nguyen

Protein aggregation can impact the product quality of pharmaceutical drugs in terms of efficacy, safety, and immunogenicity leading to economic and technical problems for biotechnology and pharmaceutical companies. Furthermore, protein aggregation can lead to the onset of a variety of diseases including amyloidosis, prion diseases, and other protein deposition disorders. To combat these issues, we investigated the potential of aptamers to prevent protein aggregation within our model, modified fluorescent protein. We have identified and successfully demonstrated that the previous selected AP3 aptamer has a high affinity for not only green fluorescent protein (GFP), but also to emerald GFP with a K_D of 199 nM. Our work has successfully demonstrated that protein aggregation can be inhibited upon binding of RNA aptamer.

Table of Contents

I. Introduction

A. Protein aggregation

Proteins are complex amino acid macromolecules that are present in nearly every process within cells. Proteins are responsible for the structure, function, and regulation of body tissues and organs, making them an interesting biomolecule for scientists to understand [1]. With 20 different amino acids, thousands of different amino acid sequences can be formed. The sequences of amino acids determine the protein's structure and function. However, in order for proteins to function properly, it is crucial for proteins to fold into a defined three-dimensional structure known as the native fold. Generally, protein folding starts during protein synthesis at the ribosome and goes through structural intermediates before the native state is reached (Fig 1) [2].

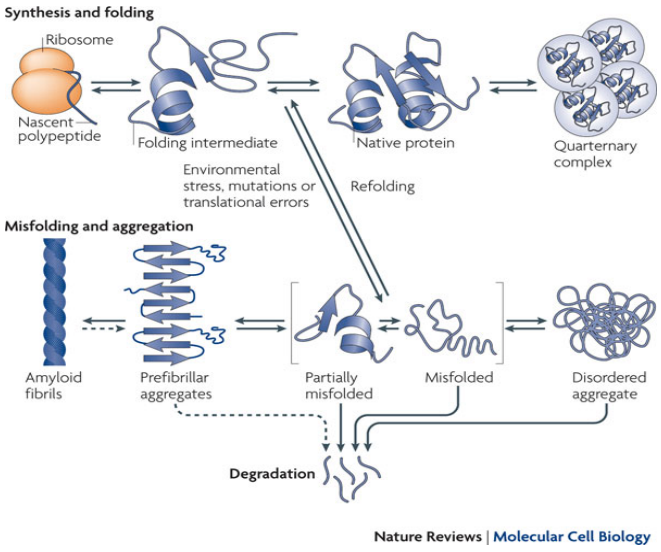

Figure 1. Overview of cellular protein aggregation. Image from [2]

Proteins have the potential to misfold and become non-functional due to different stresses, mutations in the nascent polypeptide, or translational errors [2]. Once the protein is misfolded in the cell, these intermediates can be refolded back to the native state with the help of

molecular chaperones or can be degraded by intracellular proteases in order to prevent a buildup of misfolded proteins [3]. If these networks of chaperones and proteases are overcome through persisting stress conditions, increased amounts of proteins, or in aged cells, the production of misfolded proteins will exceed the refolding or degradative capability of a cell leading to an accumulation of protein aggregates (Fig 1) [2].

B. Factors that influence protein aggregates

Many proteins have a natural propensity to aggregate due to the dynamic nature of their structures. Van der Waals forces, hydrogen bonds, disulfide linkages, and hydrophobic interactions hold protein shapes together. Environmental stresses that alter these forces and interactions could expose internal hydrophobic regions of a polypeptide chain causing an interaction of hydrophobic regions with other parts of the protein; a process that could lead to protein aggregation by stabilizing partially folded intermediates [4]. The main environmental stresses that determine the extent and rate of aggregation are the following: protein concentration, temperature, pH, ionic strength, and cosolutes [4].

1. Protein Concentration

An increase in protein concentration could result in decreased aggregation due to crowding effects, increased aggregation due to increased chance of association, and/or precipitation due to solubility limit [4]. The concept of macromolecular crowding expects a reduction in protein aggregation because crowding will favor formation of compact conformations over extended conformations of macromolecules [5]. In contrast, large concentrations of proteins would increase the degree of intermolecular protein-protein interactions, which would have a significant effect on protein aggregation [5]. Lastly, proteins can readily aggregate, precipitate, or crystallize when the protein concentration exceeds the

solubility limit – a process that is strongly dependent on solution conditions [6]. Thus, protein concentration is an important parameter in protein aggregation.

2. Temperature

Temperature is arguably the most critical environmental factor when handling proteins [5]. Increasing temperatures can theoretically reduce the activation energy of protein aggregation relative to thermal fluctuations, increase hydrophobic interaction, enhance the diffusion of proteins, and increase the frequency of collisions [5]. At decreasing temperatures, protein destabilization and denaturation may occur. Altogether, these processes will destabilize the protein and cause proteins to aggregate. However, in comparison to higher temperatures, low temperature-induced protein aggregation is often reversible and is not always an issue for proteins [5].

3. pH

pH is another important environmental factor in protein aggregation. The pH determines the surface charge of proteins, a factor that affects the intramolecular folding interactions, as well as intermolecular interactions of the protein [5]. With pH at extreme conditions, a protein's surface can become heavily charged leading to an increase solubility and a decline in the aggregation of proteins due to an increase in repulsive intramolecular and intermolecular interactions [5].

4. Ionic Strength

In close relation to pH, ionic strength is a key condition that impacts protein aggregation. Positive and negative ions can electrostatically interact with proteins and lead to altered charge-charge interactions and different conformational states, which causes proteins to

aggregate [5]. However, the overall effect of ionic strength on protein aggregation is protein-dependent and dependent on the pH of the solution. If neutralization of the protein surface charges favor protein folding/stability, reduction of such interactions by increasing ionic strength would destabilize the protein, which would open up hydrophobic patches in the protein structure and eventually lead to protein aggregation [5].

5. Cosolutes

Cosolutes are small organic molecules that affect the solubility of proteins. Kosomotropes causes water molecules to favorably interact, which stabilizes intramolecular proteins. In opposition, chaotrope agents can disrupt the hydrogen-bonding network between water molecules. These interactions could reduce the stability of the native state of proteins by weakening the hydrophobic effect and causing protein to aggregate [7].

C. Structural differences and types of protein aggregates

In general, protein aggregates takes many forms and can be classified in the following categories: *in vivo* and *in vitro*, and ordered and disordered. Amyloid fibrils (both *in vivo* and *in vitro*) are examples of highly ordered β-sheet-rich aggregates, whereas inclusion bodies are examples of *in vivo* disordered aggregates. These *in vivo* aggregates are generally formed through persisting harsh stress conditions [4]. In contrast, disordered *in vitro* aggregates are formed during the refolding of denaturant-unfolded proteins at high protein concentrations, or under weakly native conditions at high protein concentrations [4]. Intramolecular β-sheets have been recognized as a common structural component of aggregates shared by both amorphous aggregates and amyloid fibrils [2]. The degree of β-sheet-rich organization is variable between different aggregate forms, with the highest present in amyloid fibrils, in which the β-sheets run perpendicular to the fibril axis (Fig 1) [8][9]. The morphology of the

protein is influenced by different denaturing conditions, which can potentially lead to different unfolding and aggregation mechanisms and pathways [10].

D. Problems in biopharmaceuticals, biotechnology, and health from protein aggregation

1. Neurodegenerative Diseases

Amyloid protein aggregation and inclusion body formations are known to lead to the onset of neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and prion diseases [11]. According to the Institute for Neurodegenerative Diseases at the University of California, San Francisco, the global cost of Alzheimer's disease in 2010 alone was $604 billion, or 1% of the global gross domestic product. In the U.S., people suffering from both Alzheimer's and Parkinson's diseases cost the nation $200 billion annually in patient care and lost productivity. Furthermore, this amount is estimated to increase to $1.1 trillion by 2050. Due to these overwhelming costs, it is crucial to find ways to improve the treatment of these diseases.

Reasons for the toxicity of these protein aggregates in neurodegenerative diseases are still unknown. The most proposed reasons is that inclusions and aggregates represent the end stage of a molecular cascade and that the earlier steps in the cascade are linked to the pathogenesis of these diseases [11]. Many scientists today are working on understanding the molecular mechanisms of protein aggregation in order to develop therapeutics to these diseases.

2. Protein based pharmaceutics

Protein-based pharmaceuticals are among the fastest growing therapeutic agents targeting many different diseases including cancer, infectious diseases, and autoimmune diseases, and

AIDS/HIV [12]. However, protein aggregation presents a key challenge in the development and manufacturing of biologic formulations as it can impact the product quality in terms of efficacy, safety, and immunogenicity. During protein drug processing, the protein is exposed to conditions that could affect and alter the chemical and physical stability of the structure, leading to aggregation and precipitation [13]. This occurs most commonly during storage or the delivery of protein drugs. Thus, understanding how to solubilize these protein aggregates is crucial to ensure patient safety [4].

E. Current work in finding therapeutics to inhibit protein aggregation

Many approaches have been developed to target and prevent protein aggregation. One method is to reduce the protein concentration in the system. One approach is the use of chemotherapy agents and proteasome inhibitors to reduce and eliminate the cells that are being used to secrete aggregated proteins [15]. However, in these studies the patients undergoing these treatments are often too sick to tolerate chemotherapy. Another protein reduction strategy that is being pursued in Alzheimer's disease research is to inhibit or modulate the activity of specific enzymes such as β-secretases and/or γ-secretase in an effort to reduce the amount of Aβ peptides in the system [16]. However, this process can inhibit the production of other important cellular signaling proteins such as Notch, leading to serious toxicity in the human body [17]. Further knowledge about the assembly, regulation, and specificity of β-secretases and/or γ-secretase could potentially lead to a good safety profile for clinical trials. Another approach is to silence the gene by RNA-mediated interference (RNAi) to reduce the mRNA gene of the protein that is aggregated in the cell [18]. These approaches are the least mature and are explored in tissue culture and animal models [19]. There have also been data showing that proteases such as the Aβ-degrading enzyme neprilysin can degrade

monomeric and oligomeric Aβ-peptides [20]. Even so, there is a concern that the late state oligomers and amyloid fibrils will not be able to be degraded by these enzymes.

Other therapeutic techniques focus on protein quality-control strategies. This approach seeks to understand and adjust intracellular and/or extracellular proteostasis pathways involved in protein folding, trafficking, and/or degradation [14]. Studies have shown that overexpressing the molecular chaperone Hsp70 in the cell significantly reducing protein aggregation which lead to a reduction in proteotoxicity and neurodegeneration in models of Huntington's disease and Parkinson's disease [21][22] . However, there are other studies showing that overexpression of the Hsp70 chaperones failing to lower the toxic load of aggregates causing motor deficits in mouse models [23].

Another therapeutic approach is to stabilize the monomeric protein structure to prevent protein aggregation. Clinical trials results provided convincing evidence that the process of transthyretin (TTR) amyloidogenesis causes neurodegeneration and the presence of these small-molecules stabilizers slows and prevents the progression of neurological impairment [14]. Previous work has shown small molecules, like anti-inflammatory drug, diflunisal (Merck), binding and stabilizing TTR tetramer protein to inhibit protein aggregation [24].

Another strategy involves removing targeting amyloid fibrils from tissue by passive or active immunization against the amyloidogenic protein. With the use of antibodies, these biomolecules can bind to the monomeric, oligomeric, and fibrillar proteins, which can direct the degradation of these structures by phagocytic cells. This is currently being studied in Alzheimer's disease, where immunization against Aβ-peptides causes an anti-Aβ immune response and is effective in removing these aggregates [25]. However, clinical trials have

shown that immunization against Aβ-patients with late-stage Alzheimer's disease leads to deleterious outcomes, specifically aseptic meningoencephalitis in 6% of patients [26].

While these methods have been promising in inhibiting protein aggregation, there is always an interest for other effective techniques that generate results with high efficacy and low toxicity.

II. Motivation

A. Green fluorescent protein (GFP)

The extracellular and/or intracellular cross-β-sheet amyloid fibrils in patients of amyloid diseases are known as the histopathological hallmarks and serves as a basis for diagnosis. Even though human amyloid diseases consist of aggregates of diverse structure, the focus has always been on cross-β-sheet amyloid fibrils due to their well-defined structure, a characteristic in protein aggregates that has provided scientists the opportunity to target these structures [14]. Different fluorescent dyes that are known to bind to cross-β-sheet including Thioflavin T, Congo red, and other aromatic chromophores. These dyes are used extensively to understand, target, and characterize protein aggregation in the human body. However, data has shown certain dyes like Thioflavin T interacting with specific Aβ(1-40) residues and promote a rigid partially-folded conformation in solution, resulting in a more aggregation-prone ensemble of peptides [27].

To circumvent the use of dyes, we used green fluorescent protein (GFP), a well-characterized soluble globular protein that as our model for protein aggregation.

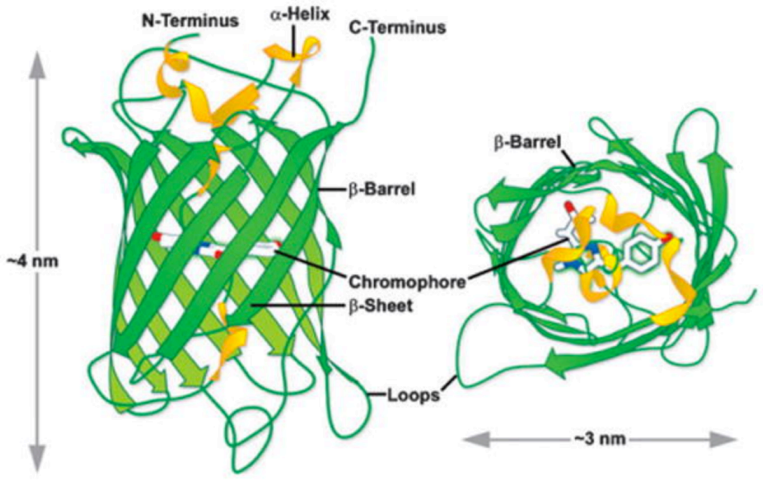

Figure 2. *A. victoria* GFP beta-barrel architecture and approximate dimension drawing from protein data bank ID:1s7s. Image from [29].

The crystal structure of this 238 amino acid residue protein was solved, revealing that the cyclic tri-peptide chromophore is buried in the center of a cylinder formed by a tightly interwoven 11-stranded β-barrel structure (Fig 2) [30]. The encapsulation of the centrally located chromophore is responsible for a high quantum yield of fluorescence, inability of O_2 to quench the excited state, and resistance of chromophore titration of external pH in solution [30]. If the protein is unfolded or aggregated, the β-barrel structure is altered causing the intrinsic fluorescence of the protein to diminish. Thus, the characteristics of GFP fluorescence serve as our marker for protein aggregation.

Over the last decade, mutagenesis strategies have been applied to the sequence encoding wild type GFP (wtGFP) to determine if different amino acid substitutions could optimize its spectral characteristics [29]. This led to the first spectral variant of *Aequorea* wtGFP known as Enhanced GFP (EGFP). EGFP had a distinct improvement over wtGFP due to its five-fold increase in fluorescence. Furthermore, EGFP matures more rapidly allowing fluorescence to be detected at earlier time points. The only issues with using EGFP is that it is slightly

sensitive to pH and has a weak tendency to dimerize, which causes an increased chance for protein aggregation [29]. In our study, we used an engineered version of EGFP, Emerald GFP (emGFP). This variant does not only contain the mutations featured in EGFP, but also includes four additional point mutations to improve the efficiency of maturation and folding at 37°C. These features would minimize the chances that would cause our GFP to aggregate on its own, an important characteristic that can help us avoid irreproducibility in aggregation studies [29].

B. Supercharging proteins and its potential to inhibit protein aggregation

Proteins are least soluble and are prone to aggregation when they are around their isoelectric point, where they have a net charge of zero [31]. In studies done by the Liu group at the Howard Hughes Medical Institute, they have explored the aggregation resistance of GFP in comparison to their supercharged variants of GFP [32]. They have created supercharged variants of GFP by identifying 29 different positions in the crystal structure of GFP that were highly solvent-exposed and mutated these amino acid residues to negatively charged residues for the negative variant and positively charged residues for the positive variant. They validated that their supercharged variants exhibited circular dichroism spectras (not shown) similar to their initial GFP control, starting GFP (stGFP), proving that these structures share similar secondary structures. In order to examine the effect of supercharging on aggregation resistance, different variants of GFP were induced to aggregate by thermal or chemical unfolding.

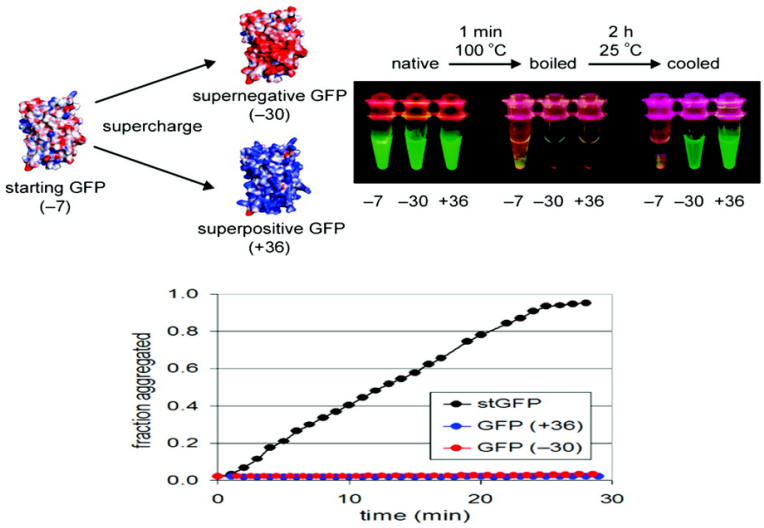

Figure 3. Aggregation of electrostatic surface potentials of starting GFP (stGFP), GFP(+36), and GFP(-30) by thermal and chemical denaturation. Thermal denaturation was induced by heating these samples for 1 min at 100°C, and subsequently cooled for 2 hours at 25°C. Chemical unfolding was induced with 40% 2,2,2-Trifluoroethanol at 25°C and monitored by right angle light scattering. Image from [32].

These results have demonstrated that protein aggregation occurred only for stGFP and not for the supercharged variants of GFP, which is shown by a loss in fluorescence for the stGFP sample and an increase in light scattering data in comparison to the supercharged samples (Fig 3). This shows that supercharged proteins lead to a significant modification in their intermolecular properties resulting in aggregation resistance. A caveat of this approach is that in order to find the amino acids that are highly solvent-exposed for supercharging, the protein's crystal structure must be well characterized [32]. Thus, alternative ways to 'supercharge' proteins is a therapeutic approach for protein aggregation that should be explored.

C. Aptamers

Nucleic acid aptamers are small single-stranded DNA or RNA oligonucleotide sequences that bind to their targets with high affinity and specificity [33]. They are synthesized using an *in vitro* selection process, also known as Systematic Evolution of Ligands by Exponential

Enrichment (SELEX) [33]. Aptamers are similar to antibodies in the fact that they can bind onto multiple and different targets including proteins, phospholipids, sugars, nucleic acids, and whole cells. Aptamers are ideal candidates for clinical applications and could be advantageous over antibodies because they are i) smaller in size and lower complexity; ii) ease of synthesis and modification *in vitro* iii) their high affinity and specificity for their targets; iv) their structural flexibility enables aptamers to bind onto hidden epitopes, which cannot be targeted by antibodies; and v) they exhibit higher stability, with the potential to be easily stored until use [34].

Recently, the Kotlikoff group at Cornell University developed a group of RNA aptamers that bind GFP and other related fluorescent proteins with low nanomolar affinity through SELEX [36]. After their first 15 rounds of SELEX, through nitrocellulose filter-binding assays, they found the G16 aptamer was >80% enriched.

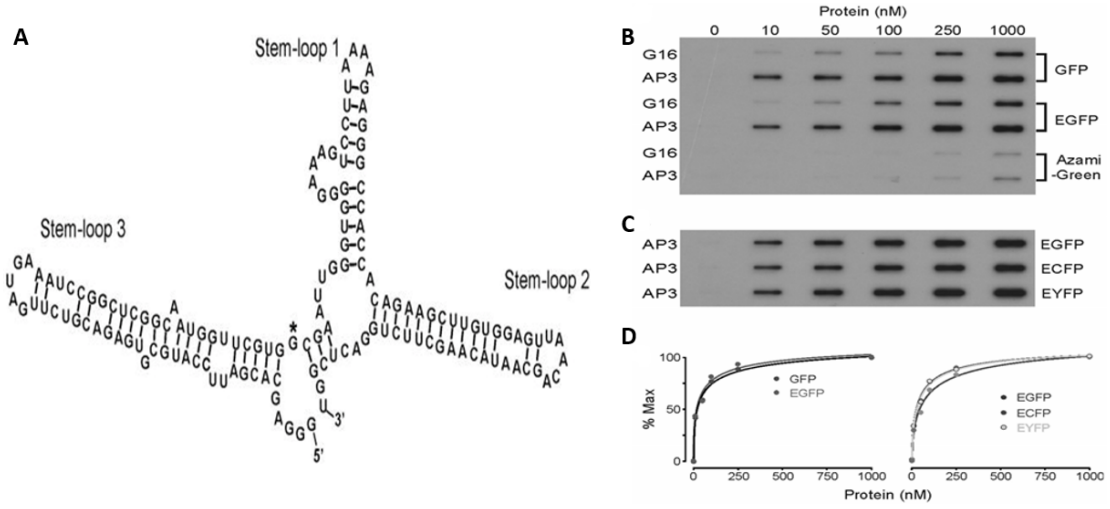

Figure 4. Structure of AP3 aptamer and binding of fluorescent binding protein aptamers to different variants of GFP. (**A**) Schematic diagram of the secondary structure of AP3. (**B**) Comparison of G16 and AP3 binding to fluorescent proteins showing that AP3 binds GFP and EGFP with higher affinity than G16. (**C**) Filter-binding assays shows similar binding of AP3 to GFP-related proteins. (**D**) Hill plot fits of data shown in (A) and (B). Image from [36].

Through further optimization and after 12 more rounds of SELEX, they found that the AP3 aptamer had a 5-fold greater affinity for both GFP and EGFP relative to G16 aptamer and also bound to YFP, and CFP (Fig 4B,C,D). Moreover, they confirmed that the AP3 and G16 aptamers only has a high affinity for fluorescent proteins by comparing the aptamer binding to the coral protein Azami-green (Fig 4B).

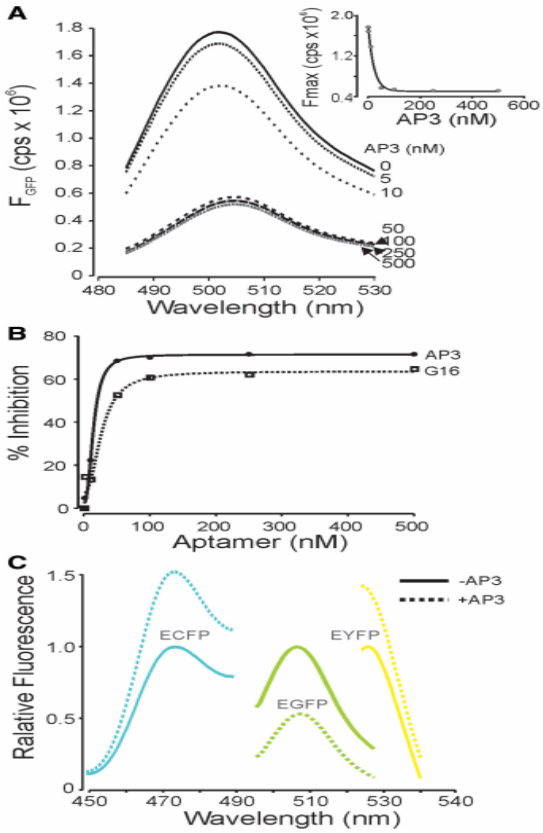

Figure 5. Fluorescence protein binding aptamers modify protein fluorescence. (**A**) Emission scan (475 nm excitation) of 25 nM GFP in the presence of increasing concentration of AP3. (**B**) Inhibition of GFP fluorescence by G16 and AP3 aptamers (475/505) (**C**) Fluorescence emission scans of 10 nM enhanced fluorescent proteins in the presence and absence of 250 nM AP3 aptamer. Image from [36].

In addition, they showed that upon binding of AP3 aptamer, there was a decrease in the fluorescence of GFP in a dose-dependent manner of its original level (Fig 5A). Interestingly, the effect on fluorescence was not equivalent for all GFP derivatives; the AP3 aptamer increased ECFP and EYFP fluorescence (Fig 5C).

Nucleic acid aptamer sequences are covalently linked through phosphodiester bonds, resulting in a hydrophilic backbone that is highly negative charged. Upon binding to its selected protein, these highly negative aptamers can potentially induce a 'supercharge' effect towards the protein of interest, resulting in a resistance towards aggregation. In addition to supercharging the protein aggregation is prevented due to an increase in steric hindrance or steric repulsions upon glycosylation of the protein [35]. With aptamers, proteins can have this same increase in steric hindrance upon aptamer binding leading to an inhibition in protein aggregation. Furthermore, with an increase of aptamer binding complexes, there would be a decrease in free monomeric protein in the system, which could also inhibit aggregation of the protein. Due to the properties of nucleic acid aptamers, the goal of this thesis is to investigate if the AP3 RNA aptamer inhibits emGFP aggregation by using the fluorescence of GFP as a marker for protein aggregation.

III. Experimental Setup

A. EmGFP Expression and Purification

Recombinant emGFP expression and purification was achieved through the transformation of the pRSET-EmGFP expression vector (Invitrogen) in chemically competent *Escherichia coli* (*E. coli*) BL21(DE3) (Novagen) by traditional heat shock transformation methods. High-level expression of the soluble mature form of emGFP was achieved by growing *E. coli* cultures in LB broth at 37 °C with the induction of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). In order to prepare soluble extracts, the cells were collected by centrifugation and lysed using a Branson Sonifier in a lysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl and 10 mM Imidazole at pH 8.0. Protein purification was carried out in an IMAC column loaded with HisPur Ni-NTA resin

(ThermoFisher) using 50 mM NaH_2PO_4, 300 mM NaCl, and 20-50 mM imidazole at pH 8 a sthe wash buffer and 50 mM NaH_2PO_4, 300 mM NaCl, and 175 mM imidazole at pH 8 as the elution buffer. Other batches of emGFP was purified through Fast Protein Liquid Chromatography (FPLC) with a 5 mL HisTrap HP column (GE Healthcare Life Sciences), using the same wash and elution buffer. All GFP containing fractions were pooled and dialyzed against 1X PBS. Aliquots were frozen in liquid nitrogen and then stored at -80 °C until use. The purity of GFP was confirmed using SDS polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Brilliant Blue R-250 (Sigma Aldrich). The following gels were imaged using Foto/Analyst Express (Fotodyne) and the Amersham Typhoon 5 Biomolecular Imager (GE Healthcare Life Sciences).

B. In Vitro Synthesis of RNA Aptamer

In order to generate the template for *in vitro* transcription, a PCR product was amplified from the DNA template sequence with the corresponding primer pairs, which contained the T7 promoter. The AP3 was made by *in vitro* transcription with T7 RNA polymerase (Kapa Biosystems) using the PCR products as the template. The RNAs were extracted by phenol-chloroform method. The quality of the RNAs was verified using 40:1 8% Native PAGE and stained with SYBR-Gold and imaged using a Foto/Analyst Express (Fotodyne). The purified RNA was stored in -80 °C. Prior to using these aptamer stocks, the RNA was heated at 80°C for 5 minutes and cooled at room temperature for 30 minutes in order to refold the aptamer into its correct structure.

C. Aptamer-Binding Assays

For electrophoretic mobility shift assays (EMSA), all binding assays were prepared in 1X PBS 5 mM $MgCl_2$ (pH 7.5). A typical binding reaction contained different amount of emGFP with a constant concentration of AP3 aptamer incubated at 37°C for 30-60 min before loading on a 79:1 12% Native polyacrylamide gels. Gels ran at 250 V for 3 hours at 4°C then stained using SYBR-Gold (Invitrogen). The exposed screens were scanned with the Amersham Typhoon 5 Biomolecular Imager (GE Healthcare Life Sciences) and analyzed with ImageQuant software. Other images were taken using the Foto/Analyst Express upon excitation using a Dark Reader transilluminator (Clare Chemical).

For fluorescence anisotropy assays, binding assays were prepared in 1X PBS 5 mM $MgCl_2$ (pH 7.5). Binding reactions consist of a serial dilution of aptamer with constant emGFP incubated at 37°C for 30-60 min. Anisotropy measurements were carried out on the FluoroMax-4 Spectrofluorometer (Horiba).

D. Fluorescence measurements and melting curves

Single fluorescence measurements and emission curves were carried out on the FluoroMax-4 (Horiba). For high throughput data sets, fluorescence measurements and emission curves were produced using the Infinite M200 PRO Multimode Microplate Reader (Tecan). Melting curves were also carried on using the FluoroMax-4 (Horiba) with a temperature heat controller.

E. Methods of aggregation

Before inducing the protein, stock emGFP aggregates were pelleted centrifugation at 13,000 rpm by a Symphony 2417R centrifuge (VWR). The supernatant was collected and the concentration of protein was determined using the Nanophotometer P300 (Implen).

For thermal unfolding induced aggregation, highly concentrated emGFP samples denatured by heating the solution at 95°C for at different time points using a digital heat block. Samples were then cooled at room temperature for an hour before fluorescence measurements using Tecan Microplate Reader.

For chemical unfolding induced aggregation, emGFP samples were denatured using different concentrations of 2,2,2-trifluoroethanol (TFE). Aggregated samples were analyzed using dynamic light scattering at 173º from Zetasizer Nano (Malvern) and fluorescence measurements using Tecan Microplate Reader.

For lyophilization and shaking/shearing induced aggregation, 500 µL emGFP samples were first lyophilized using Freeze Zone 4.5 (Labconco) in 2 mL eppendorf tubes.

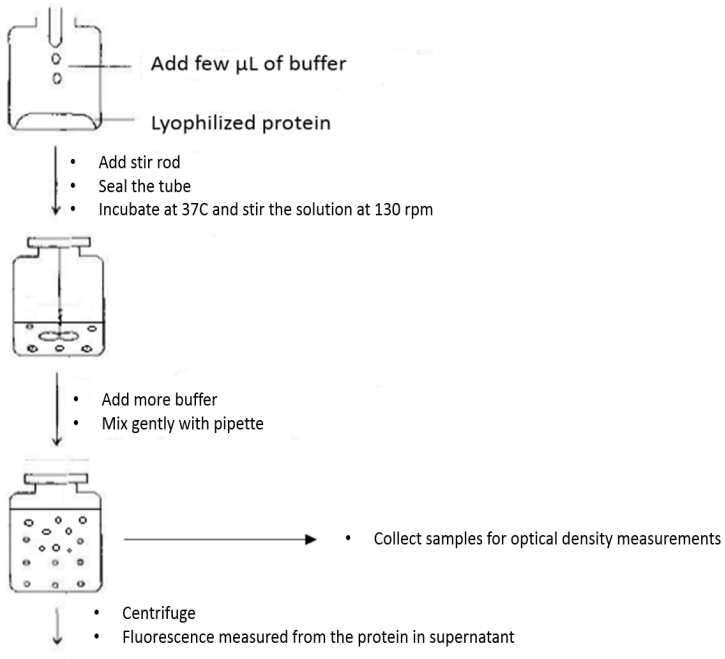

Figure 6. Schematic of lyophilization and shaking/shearing induced aggregation method.

After resuspension of the lyophilized powder using different volumes of 1X PBS, an egg-shaped teflon spin bar 3/8”X3/16” (VWR) stirred the solution at 120 rpm using a stir plate in

a 37°C warm room for 48 hours. Eventually, aggregation through shearing/shaking was completed by adding 3mm polystyrene glass beads (Fisher Scientific) into 96-well plates and shaken in the Tecan Microplate Reader at 40°C orbitally for 90 hours at 432 rpm. Fluorescence measurements and optical density measurements were done using the Tecan Microplate Reader and Nanophotometer P300 (Implen). Fluorescence measurements were carried out on the supernatant of the samples after pelleting the aggregates through centrifugation.

F. Statistical Analysis

Results are expressed as mean ± SD. Student's unpaired t-test was used to determine the statistical differences of aggregation results using GraphPad Software. (*P values <0.05, **P values <0.01, and ***P values <0.001. NS = not significant)

IV. Results

A. Expression and purification of emGFP

Initial expression was done in small scales. Thus, purification was first performed using an IMAC column loaded with HisPur Ni-NTA resin. Following purification, SDS-PAGE was used to determine the purity of the protein.

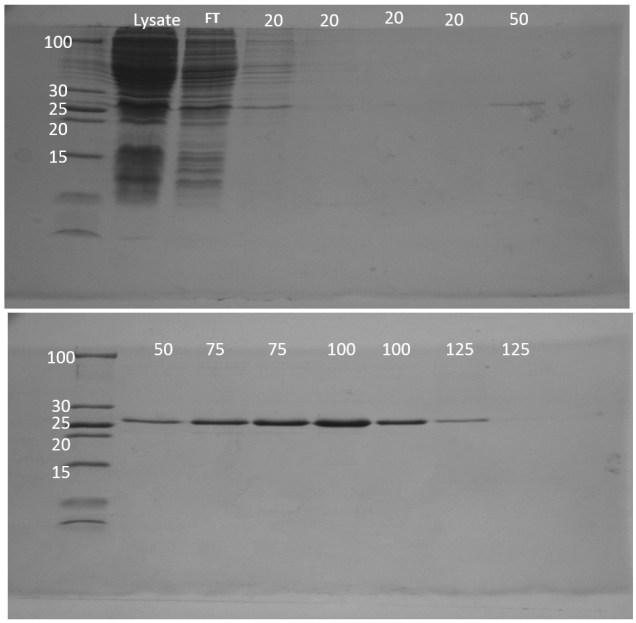

Figure 7. SDS-PAGE of emGFP fractions after purification from IMAC column loaded with HisPur Ni-NTA resin. The following concentrations of imazadole in mM concentrations are provided above in the gel image. The protein was washed with 11 column volumes of 20 mM imizadole and eluted through a gradient of different imizadole concentrations of up to 125 mM imizadole.

Purification using the IMAC column was effective shown by the single band at the molecular weight of emGFP (~27 kDa) (Fig 7). Additionally, expression and purification produced a high yield of 450 µL of 10 mg/mL emGFP or 376 µM emGFP.

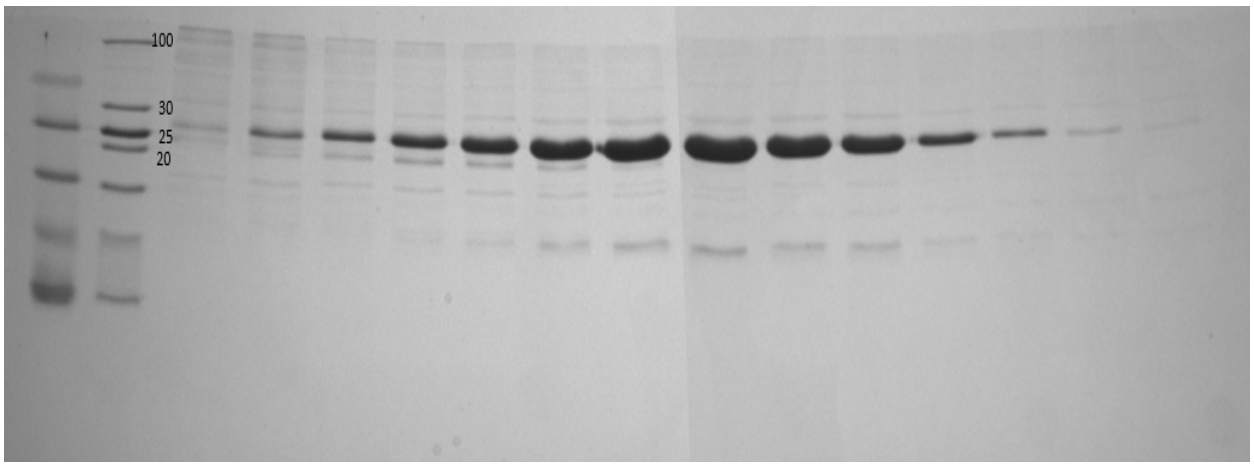

Figure 8. SDS-PAGE of emGFP fractions after purification from FPLC using a gradient of different imizadole concentrations. The protein was washed with 11 column volumes of 20 mM imizadole and eluted through a gradient of different imizadole concentrations up to 175 mM imizadole.

When purifying emGFP at larger batches, purification was completed using the FPLC using a 5 mL HisTrap HP column. Due to the impurities present from the SDS-PAGE image produced from FPLC purification (Fig 8), following batches of protein were washed longer at 20 mM imizadole, and the protein was eluted at a lower concentration of imizadole to ensure purity of emGFP (not shown). Overall, these gel images have shown that emGFP was successfully expressed and purified.

B. Purity of AP3 aptamer

After synthesizing the AP3 aptamer through *in vitro* transcription, the purity of the RNA aptamer was tested by running the samples onto a 40:1 8% native polyacrylamide gel.

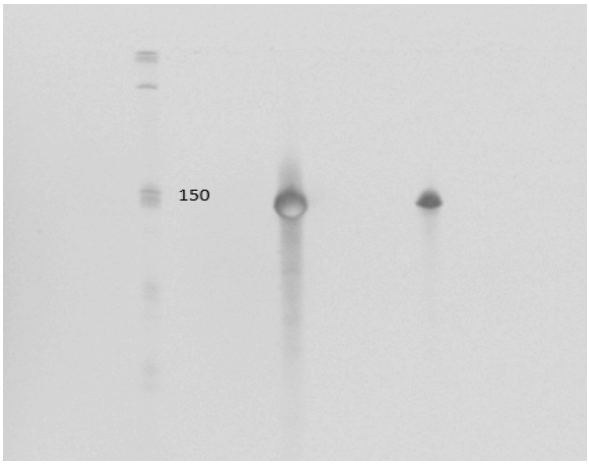

Figure 9. 29:1 8% Native polyacrylamide gel of purified AP3 Aptamer after staining with SYBR-Gold.

A single high intensity band at around 143 nt proved that the AP3 aptamer was synthesized and purified with a high yield of 450 µL of 2.190 mM of RNA (Figure 9). This same aptamer was used for future tests.

C. EmGFP-AP3 binding affinity determination

The AP3 aptamer has only been tested to bind to wtGFP and EGFP [37]. However, there has not been any data showing whether the AP3 aptamer has an affinity towards emGFP. In order to confirm the binding of the aptamer, EMSA and fluorescence anisotropy assays were carried out.

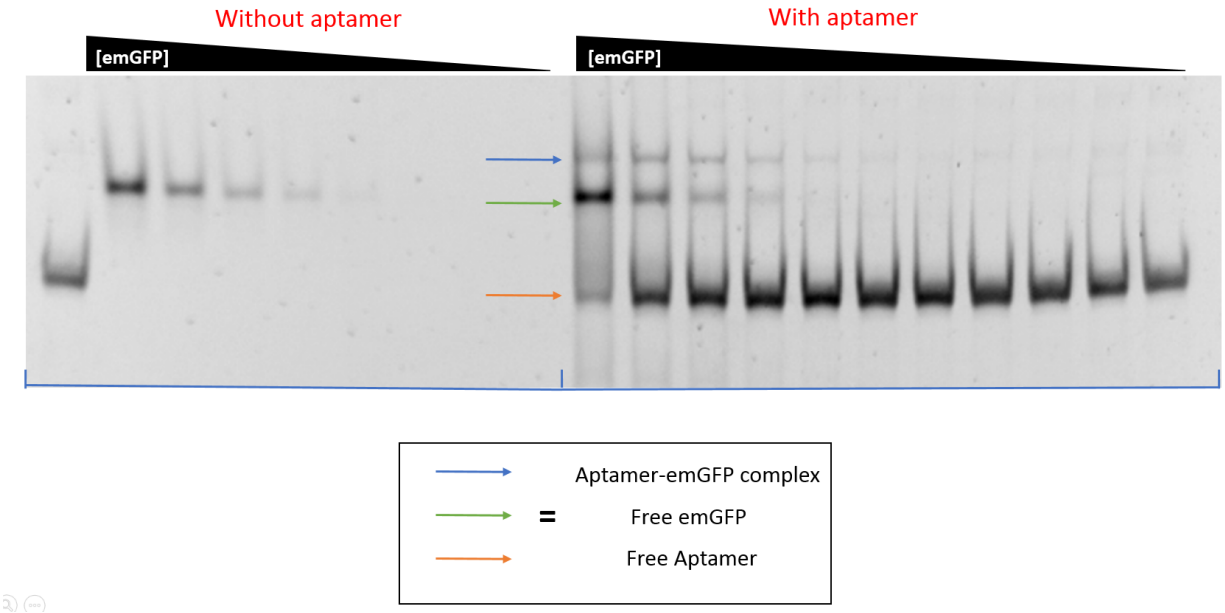

Figure 10. EMSA assay of AP3 aptamer with emGFP. EMSA images were taken using Foto/Analyst Express and excited using a Dark Reader Transilluminator. The first 9 lanes correspond to the free aptamer and emGFP controls. The following 11 lanes represent a constant concentration of 0.24 µM AP3 with a 2X serial dilution of emGFP starting from 1.7 µM emGFP.

The dark reader excites both emGPP and SYBR-Gold at a broad range of wavelengths at pure visible blue light. Through this process of excitation, a third band at a higher molecular weight is produced (shown by blue arrow) and is expected to be the AP3-emGFP complex (Fig 10).

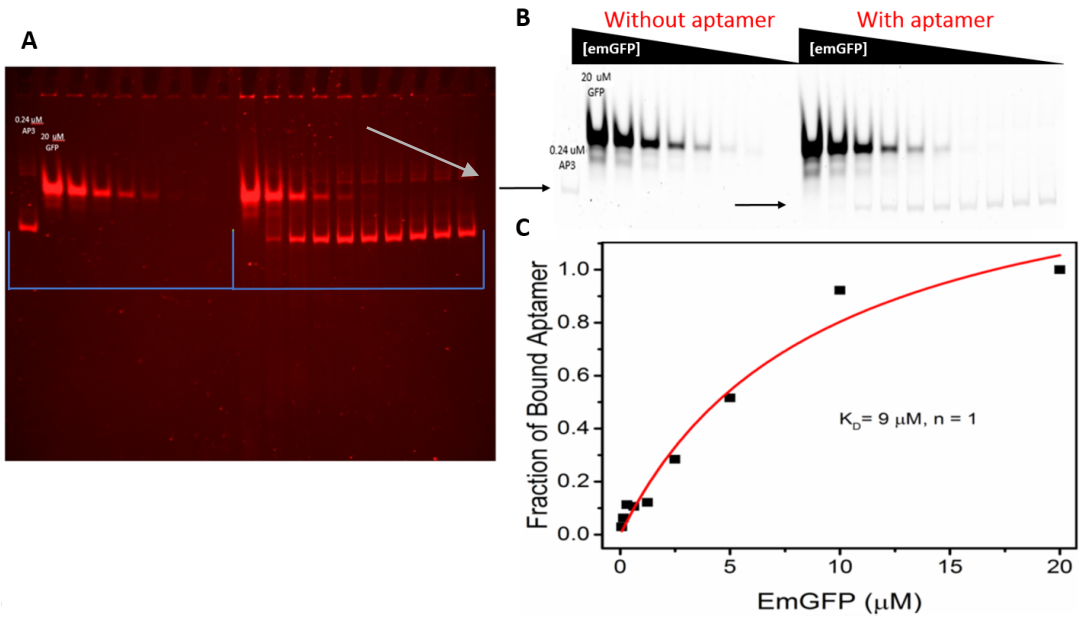

Figure 11. EMSA assay of AP3 aptamer with emGFP. The first 9 lanes correspond to the free aptamer and emGFP controls. The following 11 lanes represent a constant concentration of 0.24 μM AP3 with a 2X serial dilution of emGFP starting from 20 μM emGFP. (**A**) Image of SYBR-Gold stained EMSA assay using Foto/Analyst Express (Fotodyne), which was excited using the Dark Reader transilluminator. (**B**) Typhoon image of SYBR-Gold stained EMSA assay (**C**) Hill plot equation fit of data shown in (B) using OriginPro 2015.

From the typhoon image (Fig 11B), it is difficult to detect the bound complex (gray arrow in Fig 11A) since the peak excitation wavelength from typhoon (488 nm) is close to that of emGFP (487 nm) causing overexposure of the emGFP band (Fig 11B). However, the typhoon image has shown that increasing concentrations of emGFP results in lower aptamer band intensities. By comparing the ratio of the band intensities of free aptamer (the lowest band produced on the gel in Fig 11B) in the aptamer bound complexes to the aptamer control band intensity (lane 1 in Fig 11B), the fraction of the free aptamer of the EMSA assay over a broad range of emGFP concentrations produced a K_D of 9 μM for AP3 (Fig 11C).

To further confirm the binding of the AP3 aptamer, the changes in fluorescence anisotropy were analyzed upon AP3 aptamer binding.

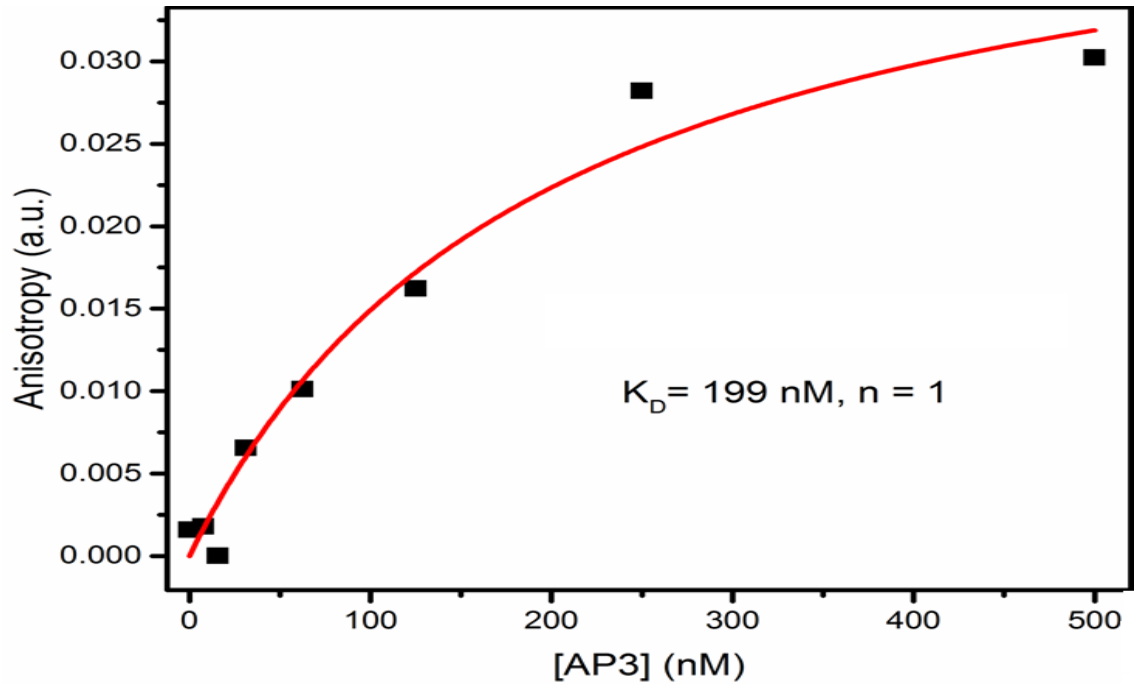

Figure 12. Fluorescence Anisotropy measurements of AP3 bound emGFP. These anisotropy readings were produced from a constant concentration of 25 nM emGFP with different concentrations of AP3 aptamer. Fluorescence anisotropy was fitted through the Hill equation using OriginPro 2015.

Fluorescence anisotropy measurements of the aptamer-emGFP complex in solution increased with higher concentrations of AP3, producing a K_D of 199 nM for AP3 (Fig 12). The results produced from the EMSA and fluorescence anisotropy assays have proven that the AP3 binds to emGFP.

D. The effect of AP3 aptamer on emGFP fluorescence

From previous studies, the AP3 aptamer is known to alter protein fluorescence and causing an increase in ECFP and EYFP fluorescence and a decrease in EGFP fluorescence [36]. Here, we test the effect of AP3 on the emGFP fluorescence.

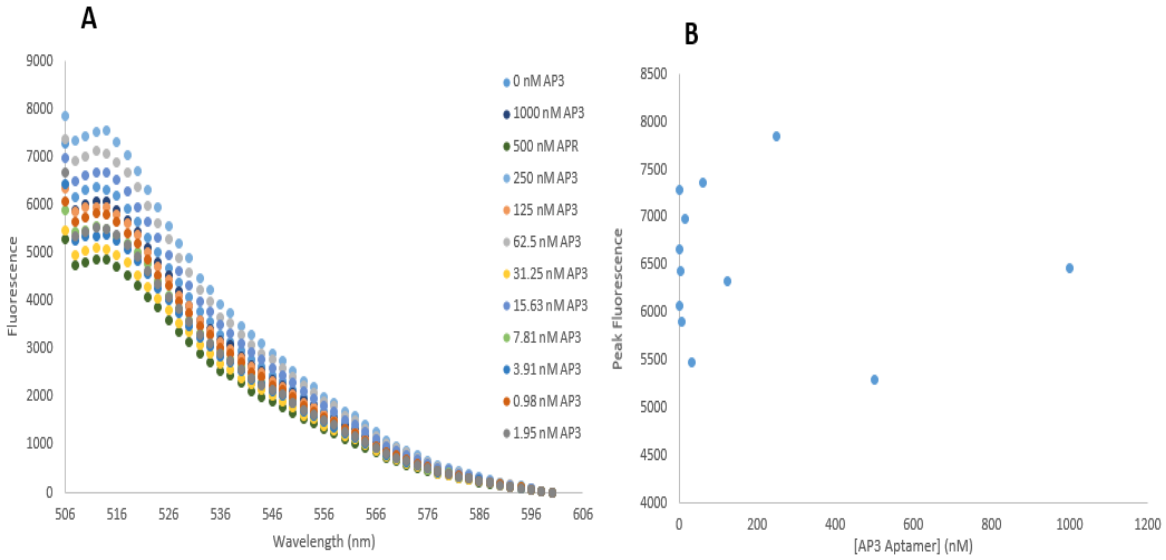

Figure 13. Fluorescent protein binding aptamers effect on protein fluorescence. (**A**) Emission curves of 25 nM emGFP with various concentrations of AP3 Aptamer was measured and excited at 475 nm using the Tecan Microplate Reader. (**B**) Peak fluorescence at different concentrations of AP3 aptamer shown in (A).

In comparison to previous studies, it is unclear how the aptamer effects the fluorescence of emGFP. Shown from the peak values from the emission curve after baseline subtraction, there is no clear trend of a dose-dependent decrease or increase in the fluorescence of emGFP (Fig 13B). Due to this, fluorescence data should not be fully trusted unless there is a clear indication that the protein has aggregated or has lost most of its fluorescence.

E. Finding a method to induce aggregation

1. Thermal and Chemical induced aggregation

Unfolded proteins can easily lead to the aggregated state during the refolding process since they can undergo different intermediate states [37]. Hence, thermal denaturation and chemical denaturation was utilized to aggregate emGFP.

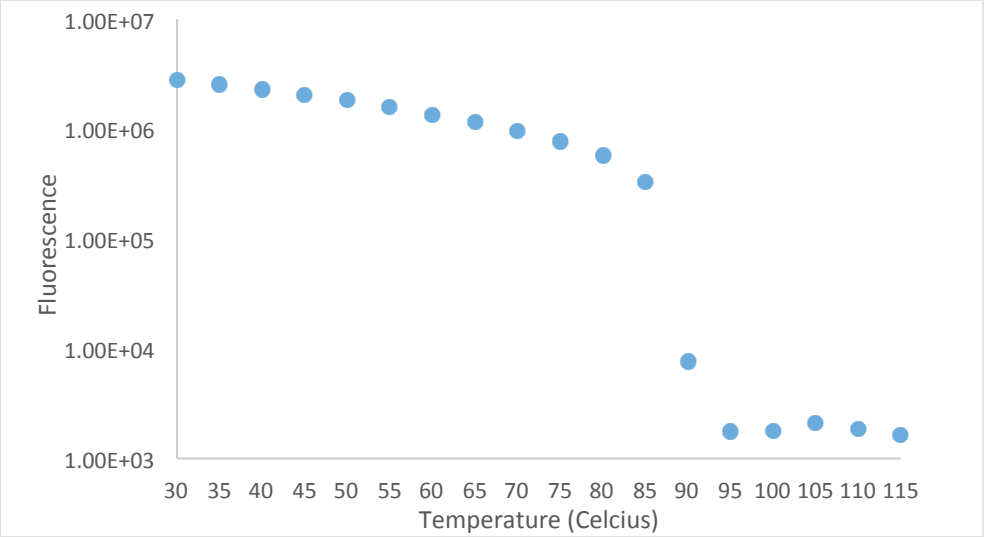

Figure 14. Melting curve of 80 µM emGFP from peak fluorescence measurements using FluorMax-4 (475/514 nm).

Before testing the effects of thermal denaturation on the aggregation of emGFP, a melting curve of the protein shows full denaturation at ~95ºC shown by the steady baseline fluorescence after this temperature (Fig 14). This temperature was used for further thermal induced aggregation steps.

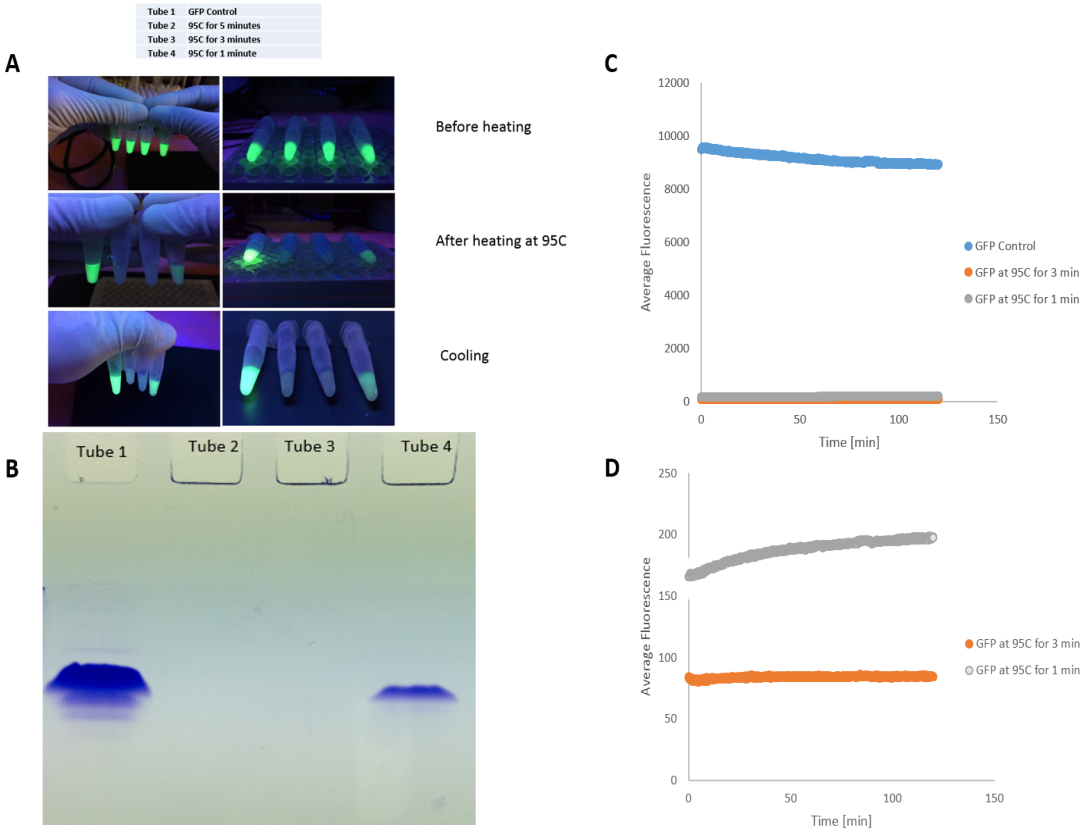

Figure 15. Thermal denaturation induced aggregation of 70uM emGFP. (**A**) Images of eppendorf tube samples of emGFP after excitation using UV-Lamp (254 nm) from thermally induced aggregation. (**B**) 29:1 8% Native Gel of tube samples in (A). (**C**) Fluorescence of different thermally induced aggregation conditions (475/514 nm) in comparison to the control using Tecan Microplate Reader. (**D**) Zoomed-in fluorescence curve from (C).

Thermal treatment induced protein aggregation shown by the aggregates stuck in the wells of the gel in comparison to the control (Fig 15B). Also, the fluorescence of the protein dramatically decreased proving that the emGFP has misfolded, lost its function of fluorescence, and aggregated (Fig 15A,C,D).

Protein unfolding induced aggregation was also carried out using a common protein chemical denaturant, 2,2,2-trifluoroethanol (TFE) [38].

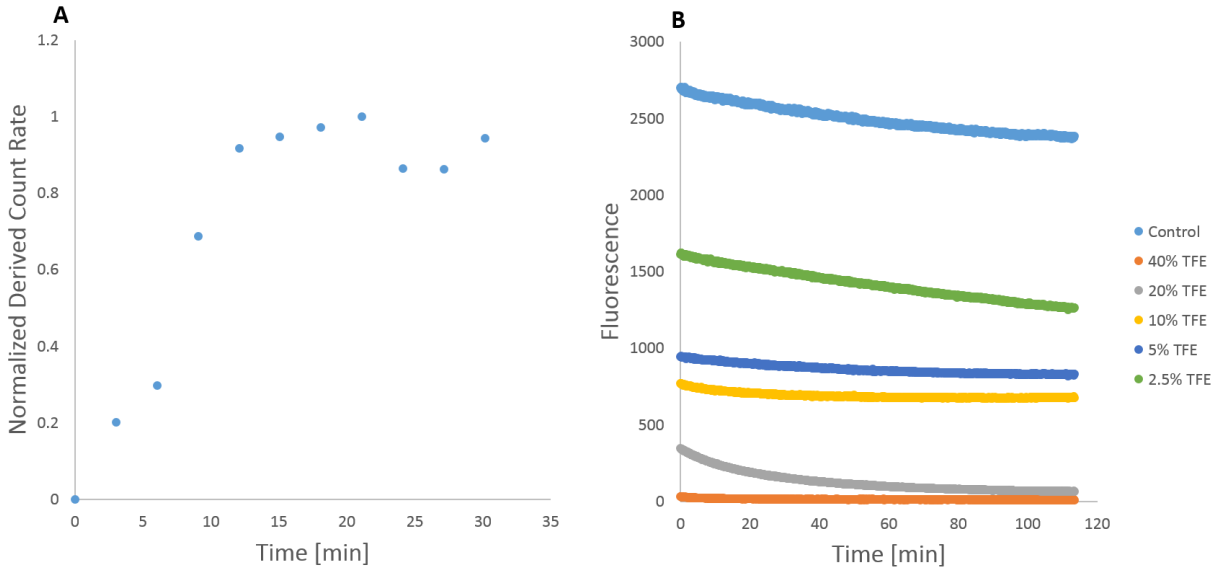

Figure 16. Chemical denaturation induced aggregation. (**A**) Dynamic Light Scattering of 60 uM emGFP aggregated using 40% TFE (**B**) Fluorescence of different chemically induced aggregation conditions (475/514 nm) in comparison to the control using Tecan Microplate Reader.

When analyzing the dynamic light scattering results, it is known that a higher derived count rate originates from higher concentration of particles, larger particles, or a combination of both. Taking advantage of these properties, the effectiveness of this chemical denaturant was tested by analyzing the derived count rate and fluorescence of our fluorescent protein. After the addition of 40% TFE to highly concentrated emGFP, there was a significant increase in derived count rate over time (Fig 16A). Furthermore, the optical density (A_{600}) of the solution before and after the addition of the TFE was recorded resulting in an increase of 0.001 to 0.435 respectively. Both of these results show that emGFP successfully aggregated through chemical unfolding. Further, the aggregation of emGFP upon different concentrations of TFE was tested. The fluorescence of emGFP increased with the amount of TFE, showing that lower concentration of TFE led to less emGFP aggregates in solution (Fig 16B). Even though these methods of unfolding induced aggregation were successful, these denaturation

conditions could also affect the folding and structure of the AP3 aptamer. As a result, other methods were explored to aggregate emGFP.

2. Agitation induced aggregation

Shaking and shearing could potentially induce protein aggregation through enhanced air/water interfaces. During this process, the hydrophobic property of the air relative to water induces protein alignment at the interface, maximizing exposure of the hydrophobic patches in the protein to the air, leading to misfolding and aggregation of the protein [39]. However, the impact of agitation in protein aggregation only occurs upon high shear forces [40]. To confirm if this method would work for our protein, agitation induced aggregation was tested.

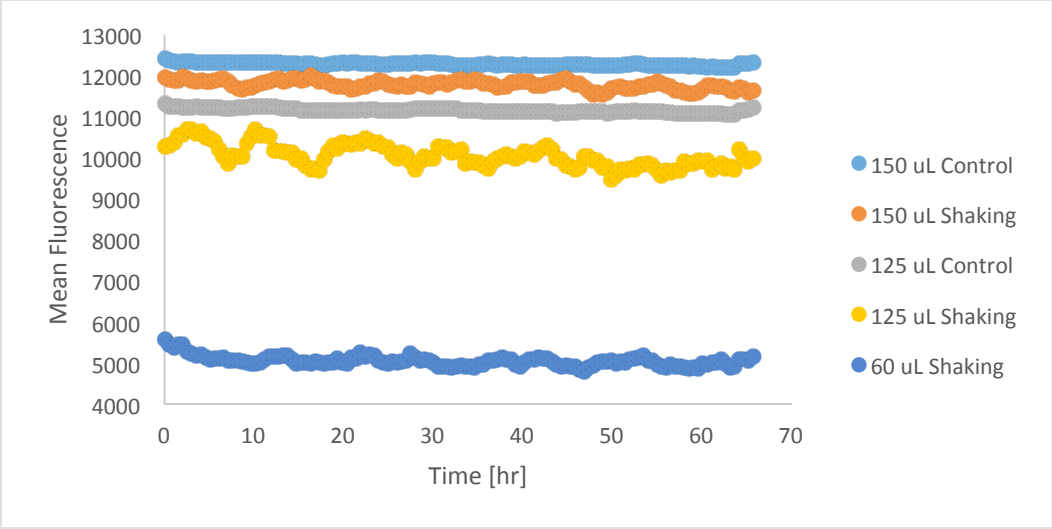

Figure 17. Aggregation kinetics of 60 µM emGFP at different volumes. Fluorescence measurements (475/514 nm) were agitated on a Tecan Microplate Reader at 40°C with polystyrene beads.

Different volumes of high protein concentrated solution were added to the tecan platereader to help us determine which volume was best for this method of aggregation. These results have shown a constant slope of fluorescence in both the control and shaking samples over a 65 hour time frame (Fig 17). In addition, the fluorescence dramatically decreased for the

samples with polystyrene beads. A possible explanation would be that the polystyrene beads and potential aggregates in solution could scatter the fluorescence, leading to a decrease in fluorescence. However, the optical density was measured between each of the samples and there were no differences between the measurements of the samples (not shown) showing that protein aggregation did not occur.

3. Lyophilization and agitation induced aggregation

Lyophilization is a common technique for the preparation of protein products. They can promote protein aggregation, both non-convalently and covalently, due to both freezing and drying stresses [41]. We combined lyophilization with shaking/shearing stresses in order to aggregate emGFP.

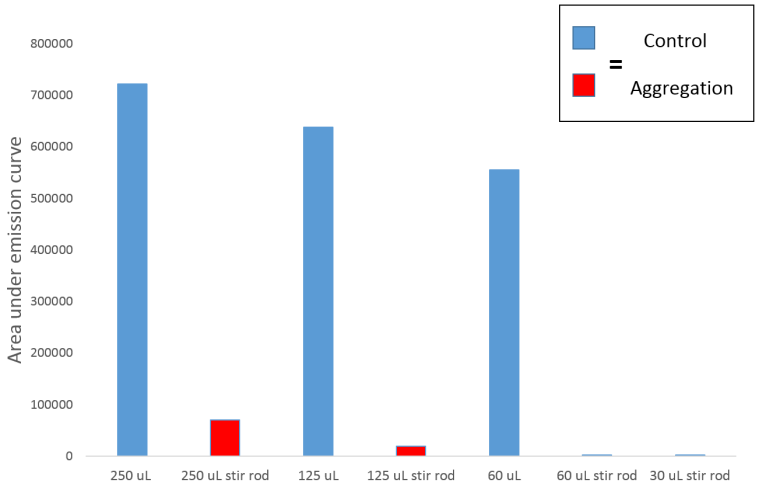

Figure 18. Aggregation of emGFP from the combination of lyophilization and shearing/shaking. 500 µL of 6 µM of emGFP was first lyophilized then resuspended at the volumes used above. The samples were then aggregated at 37°C using a stir rod. The control samples was lyophilized emGFP samples without stir rods. Samples were agitated in 1X PBS, 5 mM $MgCl_2$ before lyophilization.

The combination of both lyophilization and shaking/shearing led to aggregation shown by the dramatic decrease in fluorescence in the samples in comparison to the controls (fig 18). In

addition, a lower amount of protein was needed for aggregation making this method of aggregation approachable since there is a limited amount of AP3 aptamer. As a result, this method of protein aggregation was used to understand protein aggregation for future tests.

F. Effect of AP3 aptamer on the aggregation of emGFP

The effect of AP3 aptamer on emGFP aggregation was evaluated by the combination of lyophilization and agitation as the method of aggregation with the use of polystyrene beads instead of stir rods to overcome irreproducibility issues. The process of lyophilization would further help RNA aptamers to retain its structure even in the aggregation condition of the protein, avoiding the problem of RNA aptamer misfolding.

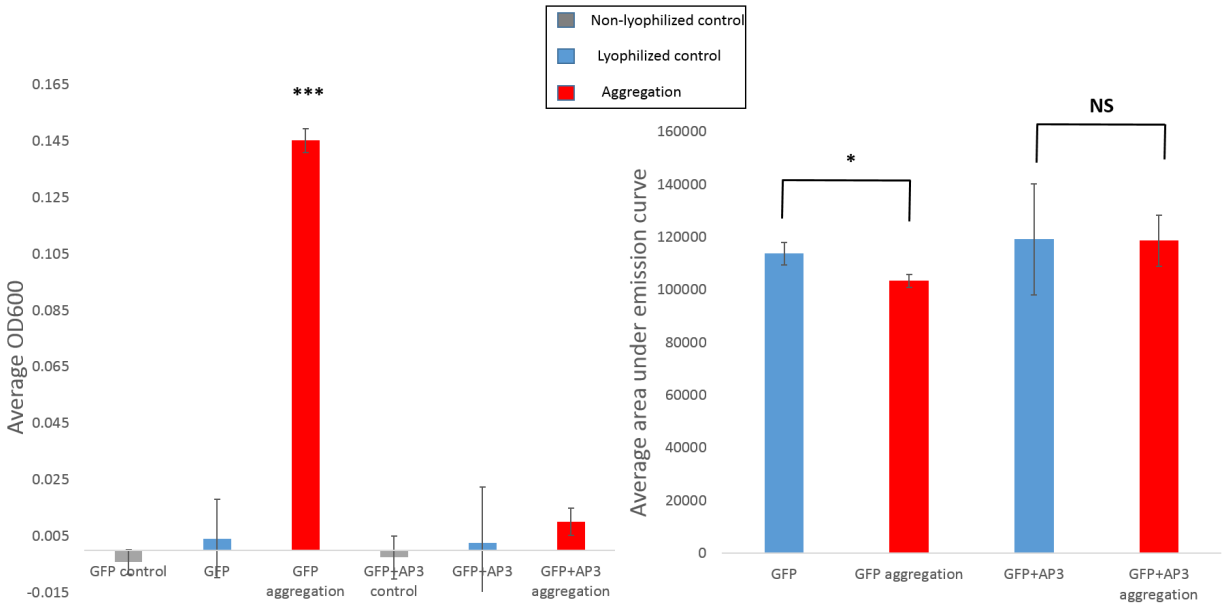

Figure 20. Effect of AP3 Aptamer on the aggregation of emGFP from samples. (**A**) Monitoring the amount of aggregates of emGFP with and without AP3 aptamer through optical density measurements of non-lyophilized (gray), lyophilized (blue), and a combination of lyophilized and shearing/shaking induced aggregation samples (red). (**B**) Observing the fluorescence of the aggregation of lyophilized emGFP with and without AP3 aptamer. The error bars represent the standard of the mean of three independent experiments with the sample size being N=3 (*P values <0.05, **P values <0.01, and ***P values <0.001. NS = not significant)

Our results showed that the process of lyophilization as a method of aggregation alone does not lead to a significant amount of emGFP aggregation (Fig 20A). Due to these issues, we've decided to measure only the emission of the lyophilized samples with and without aggregation (Fig 20B). Through lyophilization and agitation, emGFP aggregation occurred without the AP3 aptamer shown by a significant increase in OD_{600} as well as a significant decrease in fluorescence. Further, the samples of emGFP aggregation with AP3 aptamer did not aggregate shown had no significant changes in OD_{600} and fluorescence values. This data shows that the protein aggregation can be inhibited upon binding of the AP3 aptamer.

2. Discussion

We have demonstrated that the previously selected AP3 aptamer has high binding affinity not only for GFP, but also for emGFP, the modified version of this protein. This aptamer displays an affinity for emGFP with an apparent K_D of 200 nM shown from anisotropy and EMSA assays. We further revealed that upon binding of the AP3 aptamer, the emitted fluorescence could either increase or decrease, which was different in comparison to GFP and EGFP. We successfully devised a reproducible technique to aggregate emGFP by combining lyophilization and shaking/shearing the protein with a polystyrene glass bead. Upon binding of the AP3 aptamer, our results showed an inhibition in aggregation analyzed from the optical density and emission of emGFP.

Due to the discrepancy in K_D from EMSA and anisotropy measurements, other tests to obtain a more accurate measurement of K_D could be useful to fully confirm the binding of the aptamer. When using SYBR-Gold, uneven staining of the nucleic acid sequences could occur, leading to a large inaccurate calculation of K_D. In order to clearly detect the aptamer

and protein complex in the native gel and to find a more precise measurement of K_D, use of fluorophore-labeled RNA aptamer with an emission peak wavelength far from emGFP would be useful. Isothermal titration calorimetry (ITC) experiments would also be useful by providing a more accurate measurement of K_D of the RNA aptamer towards emGFP.

Multiple groups have shown that nucleic acid aptamers can be selected against different peptides that can inhibit the aggregation of proteins like Mutant Huntington in Huntington's disease [42], Amyloid Beta and Tau in Alzheimer's disease [43][44], and even infectious prion proteins [45]. Our work also confirms that nucleic acid aptamers can inhibit the aggregation of model fluorescent protein, emGFP. Currently, our lab focuses on elucidating the sequence-structure-function relationship of nucleic acid aptamers in order to rationally design aptamers *de novo*. Future directions focuses on building a database for therapeutic aptamers that can inhibit protein aggregation in effort to combat neurodegenerative diseases, as well as finding solutions to overcome the difficulties arising from aggregated-protein drugs in pharmaceutical industries.

Timeline

A table of the dates which various parts of this setup were complete is given in Table A.1.

Step	Date
Expression and Purification of emGFP	June 2017-Sep 2017
Reproducible method of aggregation	Sep 2017-March 2017
Synthesis of AP3 Aptamer	Sep 2017-Oct 2017
Aptamer-Binding Assays	Oct 2017-March 2017
AP3 aptamer on the aggregation of emGFP	March 2017-May 2017

Table A.1. Timeline of this thesis

References

1. "What are Proteins and What Do They Do? – Genetics Home Reference." *U.S. National Library of Medicine.* National Institute of Health, 30 May 2017. Web. 01 June 2017
2. Tyedmers, J. et al. Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology* **11**, 777-788 (November 2010)
3. Goldberg, A. L. Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895–899 (2003)
4. Fink, A. Protein aggregation : folding aggregates, inclusion bodies and amyloid. *Folding and Design* **3**, R9-R23 (1998).
5. Wang, W. et al. Protein aggregation-Pathways and influencing factors. *International Journal of Pharmaceutics.* **390 (2)**, 89-99 (2010)
6. Adachi, K., Ding, M., Asakura, T., Surrey, S., 2009. Relationship between beta4 hydrogen bond and beta6 hydrophobic interactions during aggregate, fiber or crystal formation in oversaturated solutions of hemoglobin A and S. Arch. *Biochem. Biophys*. **481**, 137–144.
7. Salvi G et al. Effective interactions between chaotropic agents and proteins. *Proteins*. **61(3)**. 429-499 (2005)
8. Fandrich, M. On the structural definition of amyloid fibrils and other polypeptide aggregates. *Cell. Mol. Life Sci.* **64**, 2066–2078 (2007).
9. Maji, S. K., Wang, L., Greenwald, J. & Riek, R. Structure-activity relationship of amyloid fibrils. FEBS Lett. 583, 2610–2617 (2009).
10. Wang, L., Schubert, D., Sawaya, M. R., Eisenberg, D., & Riek, R. (2010). A multi-dimensional Structure-Activity Relationship of a protein in its aggregated states. *Angewandte Chemie (International Ed. in English)*, ***49*(23)**, 3904–3908.
11. Ross, C. et al. Protein aggregation and neurodegenerative disease. *Nature Medicine* **10**, S10-S17 (2004)
12. Holmer, A. F. *Survey: Medicines In Development for HIV/AIDS* (Pharmaceutical Research and Manufacturers Association, Washington DC, 2004).
13. Wang, W. Protein aggregation and its inhibition in biopharmaceutics. *Int. J. Pharm.* (in the press).
14. Eisele, Y. S. et al. Targeting protein aggregation for the treatment of degenerative diseases. *Nature Reviews Drug Discovery* **14.** 759-780 (2015)
15. Kastritis, E. *et al*. Bortezomib with or without dexamethasone in primary systemic (light chain) amyloidosis. *J. Clin. Oncol.* **28**, 1031–1037 (2010).
16. De Strooper, B., Vassar, R. & Golde, T. The secretases: enzymes with therapeutic potential in Alzheimer disease. *Nat. Rev. Neurol.* **6**, 99–107 (2010).
17. De Strooper, B. & Gutierrez, L. C. Learning by failing: ideas and concepts to tackle γ-secretases in Alzheimer disease and beyond. *Annu. Rev. Pharmacol. Toxicol.* **55**, 419–437 (2014).

18. Coelho, T. *et al*. Safety and efficacy of RNAi therapy for transthyretin amyloidosis. *N. Engl. J. Med.* **369**, 819–829 (2013).
19. Hovey, B. M. *et al*. Preclinical development of siRNA therapeutics for AL amyloidosis. *Gene Ther.* **18**, 1150–1156 (2011)
20. Kanemitsu, H., Tomiyama, T. & Mori, H. Human neprilysin is capable of degrading amyloid β peptide not only in the monomeric form but also the pathological oligomeric form. *Neurosci. Lett.* **350**, 113–116 (2003)
21. Warrick, J. M. et al. Suppression of polyglutaminemediated neurodegeneration in Drosophila by the molecular chaperone HSP70. *Nat. Genet*. **23**, 425–428 (1999). 166.
22. Auluck, P. K., Chan, H. Y., Trojanowski, J. Q., Lee, V. M. & Bonini, N. M. Chaperone suppression of α-synuclein toxicity in a Drosophila model for Parkinson's disease. *Science* **295**, 865–868 (2002).
23. Shimshek, D. R., Mueller, M., Wiessner, C., Schweizer, T. & van der Putten, P. H. The HSP70 molecular chaperone is not beneficial in a mouse model of α-synucleinopathy. *PLoS ONE* **5**, e10014 (2010).
24. Sekijima, Y., Dendle, M. A. & Kelly, J. W. Orally administered diflunisal stabilizes transthyretin against dissociation required for amyloidogenesis. *Amyloid* **13**, 236–249 (2006)
25. Schenk, D. et al. Immunization with amyloid-β attenuates Alzheimer disease-like pathology in the PDAPP mouse. *Nature* **400**, 173–177 (1999).
26. Check, E. Nerve inflammation halts trial for Alzheimer's drug. *Nature* **415**, 462 (2002).
27. MG, D. et al. Thioflavin T templates amyloid β(1-40) conformation and aggregation pathway. *Biophys Chem*. **206.** 1-11 (2015)
28. Hiroyuki Fukuda,Munehito Arai, and, and Kunihiro Kuwajima*. Folding of Green Fluorescent Protein and Cycle3 Mutant. *Biochemistry* **2000.** 39 (39), 12025-12032
29. Day, R. N., & Davidson, M. W. (2009). The fluorescent protein palette: tools for cellular imaging. *Chemical Society Reviews*, ***38***(10), 2887–2921
30. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., & al, e. (1996). Crystal structure of the aequorea victoria green fluorescent protein. *Science, **273***(5280), 1392. Loeb, J. *J. Gen. Physiol.* 1921, 4, 547-555
31. Loeb, J. *J*. *Gen*. *Physiol*. **1921**, 4, 547-555
32. Lawrence, M. S., Phillips, K. J., & Liu, D. R. (2007). Supercharging Proteins Can Impart Unusual Resilience. *Journal of the American Chemical Society*, ***129***(33), 10110.
33. Mukhopadhyay, R. Aptamers are ready for the spotlight. *Analytical Chemistry*. 115A-118A (2005)
34. Pei, X., Zhang, J., & Liu, J. (2014). Clinical applications of nucleic acid aptamers in cancer (Review). *Molecular and Clinical Oncology*, **2,** 341-348.
35. SOLÁ, R. J., & GRIEBENOW, K. (2009). Effects of Glycosylation on the Stability of Protein Pharmaceuticals. *Journal of Pharmaceutical Sciences*, ***98***(4), 1223–1245.

36. Bo Shui, Abdullah Ozer, Warren Zipfel, Nevedita Sahu, Avtar Singh, John T. Lis, Hua Shi, Michael I. Kotlikoff; RNA aptamers that functionally interact with green fluorescent protein and its derivatives. *Nucleic Acids Res* 2012; **40** (5).
37. Yazdanparast, R., Esmaeili, M.A., Khodagholi, F., 2007. Control of aggregation in protein refolding: cooperative effects of artificial chaperone and cold temperature. Int. J. Biol. Macromol. **40**, 126–133.
38. Shimizu, S. et al. Alcohol Denaturation: Thermodynamic Theory of Peptide Unit Solvation. *J. Am. Chem.* 1999, **121,** 2387-2394
39. Wang, W. et al. Protein aggregation-Pathways and influencing factors. *International Journal of Pharmaceutics.* **390 (2)**, 89-99 (2010)
40. Bee, JS. et al. Monoclonal antibody interactions with micro- and nanoparticles: adsorption, aggregation, and accelerated stress studies. *J Pharm Sci.* **98** (9). 3218-3238 (2009)
41. Giehm, L. et al. Assays for a-synuclein aggregation. *Methods*. **53**(3). 295-305 (2011)
42. Chaudhary, R. K. et al. Inhibition of Aggregation of Mutant Huntingtin by Nucleic Acid Aptamers *in vitro* and in a yeast model of Huntington's disease.
43. Ji Hyeon Kim, Eunkyoung Kim, Won Hoon Choi, Jeeyoung Lee, Jung Hoon Lee, Hyojin Lee, Dong-Eun Kim, Young Ho Suh, and Min Jae Lee. Inhibitory RNA Aptamers of Tau Oligomerization and their neuro protective roles against preteotoxic stress. *Molecualr Pharmaceutics.* **13 (6)** . 2039-2048 (2016)
44. Takahasi T. et al. RNA aptamers selected against amyloid beta-peptide inhibit the aggregation of Abeta. *Mol Biosyst.* **5**(9) / 986-91 (2009)
45. Proske, D. et al. Prion-protein-specific aptamer reduces PrP^{sc} formation. *ChemBioChem* **3.** 717-725 (2002)