Summer 2023

Engineering and Characterizing Proteins: A Dual Study on Canine p53 Protein and Cyan Thermal Protein

Dhruv B. Sitapara
Missouri State University, dbs28s@MissouriState.edu

ENGINEERING AND CHARACTERIZING PROTEINS: A DUAL STUDY ON CANINE P53 PROTEIN AND CYAN THERMAL PROTEIN

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Chemistry

By

Dhruv B. Sitapara

August 2023

ENGINEERING AND CHARACTERIZING PROTEINS: A DUAL STUDY ON CANINE P53 PROTEIN AND CYAN THERMAL PROTEIN

Chemistry

Missouri State University, August 2023

Master of Science

Dhruv B. Sitapara

ABSTRACT

This research aims to explore protein development and characterization, focusing on a dual study of canine p53 protein and cyan thermal protein. For the canine p53 protein, the goal is to comprehend its structure to better understand a key component of canine cancers. This multi-step process involved expressing the canine p53 DNA-binding domain in *E. coli* and purifying it through affinity and ion-exchange chromatography. The purified protein was studied to test the binding of canine p53 protein to human DNA sequences using Surface Plasmon Resonance (SPR) experiments and an Electrophoretic Mobility Shift Assay (EMSA), visualized with a SYBR-safe stain. The protein's three-dimensional structure is obtained via X-ray diffraction following the formation of protein crystals. Simultaneously, the study investigates the modification of the thermal green protein (TGP) into a cyan version by introducing a mutation at residue 67 that alters a tyrosine (Y) residue into a tryptophan (W) residue (Y67W) in the protein's chromophore. Successful mutation led to the chromophore fluorescing in the cyan region. The mutant protein was expressed in E. coli and purified using affinity and ion-exchange chromatography. The mutant's stability was determined through chemical, thermal, and pH stability assays. Secondary mutants were constructed, including a Q66E mutation or the I199 residue exchanged for either a serine (I199S) or a threonine (I199T) to stabilize the chromophore. Protein crystals were grown in the first step towards determining a crystal structure for CTP which may aid future efforts in optimizing the protein.

KEYWORDS: canine p53 protein, cyan thermal protein, electrophoretic mobility shift assay, fluorescent, cancer

ENGINEERING AND CHARACTERIZING PROTEINS: A DUAL STUDY ON CANINE P53 PROTEIN AND CYAN THERMAL PROTEIN

By

Dhruv B. Sitapara

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Chemistry

August 2023

Approved:

Natasha DeVore, Ph.D., Thesis Committee Chair

Gautam Bhattacharyya, Ph.D., Committee Member

Tuhina Banerjee, Ph.D., Committee Member

Christopher Lupfer, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I want to thank the following people for contributing to my research and life:

- Dr. Natasha DeVore, my thesis advisor, for letting me join her research group and for being incredibly kind, thoughtful, and patient throughout the last two years. Also, for her willingness to help find a solution to every challenging problem. I admire her approach to everything and hope to emulate similar qualities throughout my career.
- My Thesis committee- Dr. Bhattacharyya, Dr. Lupfer, and Dr. Banerjee, for agreeing to be on my thesis committee and providing valuable feedback.
- Dr. Wanekaya, Dr. Schick, and Linda for always providing much-needed support and advice throughout the last two years.
- Graduate College - Dr. Hoelscher, Dean Masterson, and Lisa Taylor - for helping me find success outside of academics at MSU and for funding needs for my research and conference travel.
- Dr. Madhuri Manpadi- my undergraduate advisor at Drury, for introducing me to research and for always being there when life got hard.
- The graduate students of the chemistry program for shared adventures. Special shoutout to Carissa for showing me how to use SPR.
- My parents and my family- For always supporting and loving me unconditionally and for their upbringing.

I want to dedicate this thesis to my late grandfather, Bapuji, who sadly passed away a few days before my thesis defense.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

This thesis is divided into two different parts. The first part is about the characterization and crystallization of canine p53 protein. The second part concerns developing and engineering a cyan thermal protein (CTP) and its mutants. Despite being very different topics, there are some overlaps on the experimentation side of the projects. The general research design of both projects is shown in Figure 1.

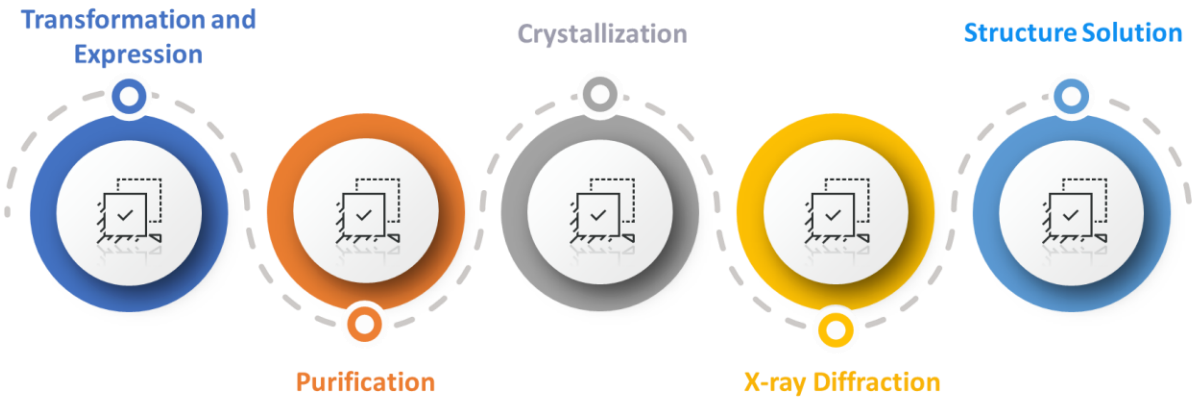

Figure 1: General research outline.

Both projects have a few differences as well. The CTP projects require the creation of mutations using site-directed mutagenesis (SDM) before the transformation and expression step so that the correct sequences can be obtained. Once the mutants are purified, they are characterized by using different assays to study their quantum yield (QY) and chemical, thermal, and pH stability. The canine p53 protein project does not require SDM but does require assays used to test the binding affinity of the canine protein to different human DNA sequences. Both projects share a similar structure biology component: solving the structure of canine p53 protein and different cyan thermal protein (CTP) mutants.

CHAPTER 1: CANINE p53 PROTEIN PROJECT

Introduction

Overview of Cancer. Cancer is a complex disease with a large variance between the over 100 types of defined cancer.[1] Cancer can be affected by a combination of genetic, immune, and environmental factors.[2] The prevalence of cancer is influenced by a myriad of factors, with age being the main risk factor due to the accumulation of successive mutations.[3] According to the American Association of Cancer, there were an estimated 18.1 million cases of cancer globally in 2020, which is set to increase to 28 million by 2040 (Figure 2).[4] Based on the same report, the estimated death toll of cancer was 9.89 million in 2020, which is set to increase to 18.1 million by 2040 (Figure 2).[4] The same report also estimated that the direct medical cost of cancer care in the US alone was $183 billion in 2015, and they are estimated to increase to $246 billion by 2030.[4] Based on these statistics alone, cancer is set to become an even bigger healthcare concern in the next few decades.

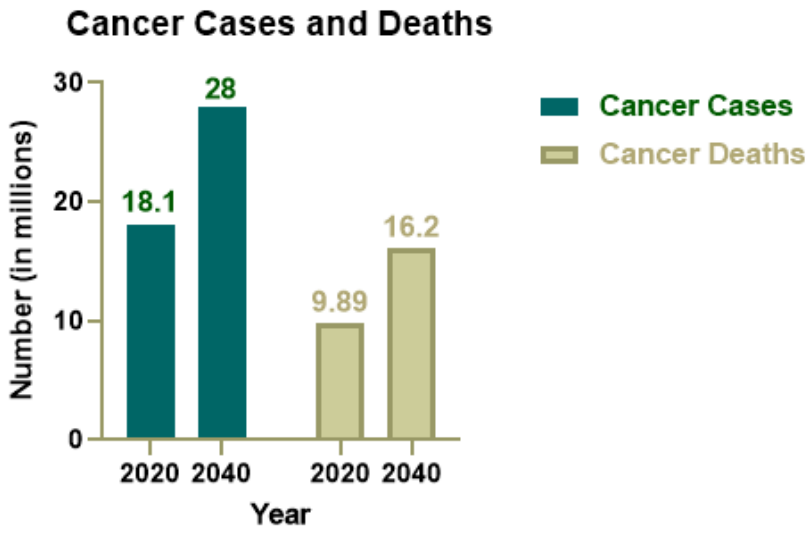

Figure 2: Estimated cancer cases and deaths in 2020 and 2040.[4]

Cancer in Dogs. More than 1 million new cases of dog cancer occur within the US, and cancer is the most common cause of death of dogs, with an estimated rate of 30%.[5,6] The most common types of dog cancer include non-Hodgkin lymphoma, malignant melanoma, osteosarcoma, and multiple brain cancer types.[5,7] According to the American Veterinary Medical Association (AVMA), one out of four dogs at some stage in their life, and over fifty percent of dogs over the age of 10, will develop cancer.[8] Dog cancers are a valuable source of information and need to be studied more as a large proportion of dogs will develop cancer over their lifetimes.[9,10] Dogs and humans share similar living environments (domesticated dogs), nutrition, cancer histology, therapeutic response, metastasis, and genetic and molecular targets.[9,11–15]

Every year, more evidence supports the study of canine cancers for clinical and translational research.[9,13,15–18] Currently, 19,000 canine genes are orthologous to human genes.[9,19] Canine and human DNA and protein sequences are more similar to mice as there are 650 megabases (Mb) of ancestral sequence overlapping between humans and dogs that are not shared with mice.[9,15,20] Tumor suppressor p53 is estimated to be mutated in over 50% of all human cancers[21], and a study done in dogs with osteosarcoma (OSA) revealed that approximately 41% of canine OSA tumors had mutations associated with p53.[22] Due to the similarities in the cancers in dogs and humans, many human cancer therapeutic agents are used for canine cancer treatments.[22] Most of these treatments are older platinum-based treatments and not targeted immunotherapies available for humans.[22] This means that dog cancer treatments have largely remained the same over the past few years, and newer therapies specifically for canine cancers are much needed. For veterinary clinics and dogs to benefit from newer therapies, it is important to understand better the genes and proteins involved in the initiation and progression of canine cancers. As of July 2023, only 151 experimentally solved 3D structures in the PDB with *Canis*

lupus familiaris (dog) as the source organism, compared to over 63,000 human-origin structures, highlighting the need for more research on canine proteins.

Introduction to p53. Many genes and proteins are involved in pathways that lead to cancer, but the most frequently altered gene in human cancers is the TP53 tumor suppressor gene (p53 refers to a protein, TP53 refers to the gene).[23–25] TP53 is located on chromosome 17 and was initially identified as a protein that bound to simian virus 40 antigens[26] and was named p53 in the 1970s because it appeared as a 53 kD protein on Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE).[27] The actual molecular weight of p53 is 43.7 kDa, but a large number of proline residues slowed its migration.[25]

Role of p53 Protein. The p53 protein is an essential tumor suppressor that plays an important role in preventing malignant growth.[28] The activation of p53 occurs when DNA is damaged and can cause cell cycle arrest to reduce the proliferation of damaged cells.[29] The biggest regulator of p53 is the mouse double mutant 2 homolog (MDM2) which causes the inactivation of p53, promoting its degradation.[29] This cycle causes an autoregulatory feedback loop that regulates the balance between p53 and MDM2 levels, which is crucial for maintaining a good balance between cell growth and cell death.[29] A simplistic version of a pathway depicting the role of p53 and its regulation by MDM2 is shown in Figure 3. In the figure, activated p53, following DNA damage, binds to DNA to exert transcriptional control along the target genes. The binding of p53 to DNA is crucial in its role as a transcription factor.[25] This transcriptional control leads to the increase or decrease in the formation of target gene products. In the absence of DNA damage or during "normal" periods, MDM2 leads to p53 ubiquitination and proteasomal breakdown. Another key player in the regulation of p53 are microRNAs that act as negative regulators by binding to the 3'-UTR region of p53.[30]

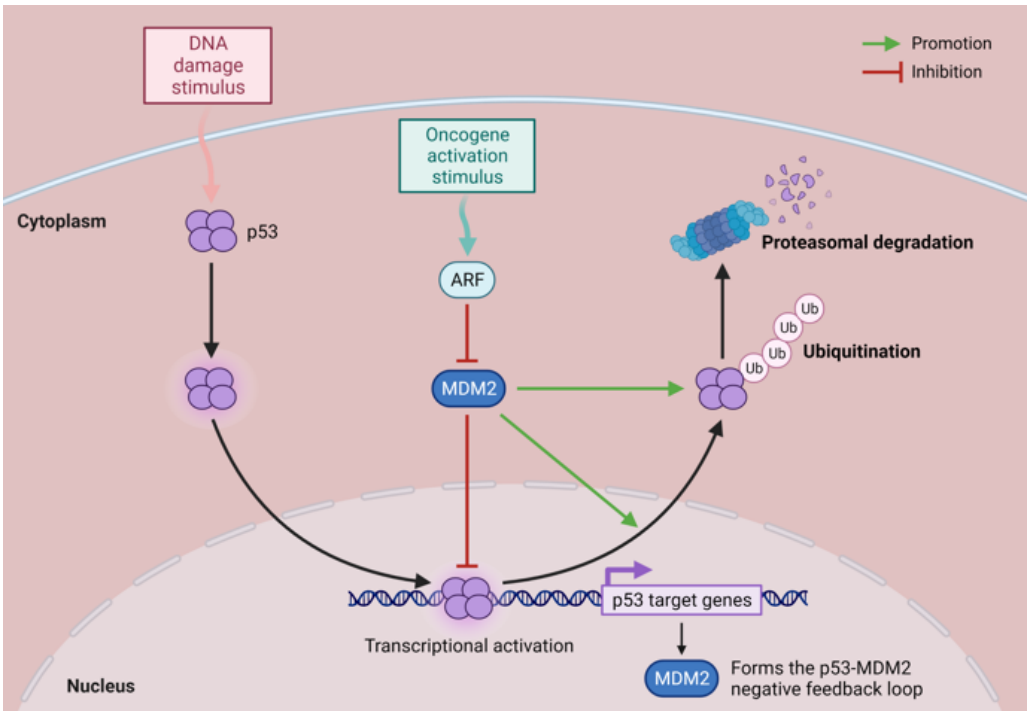

Figure 3: Regulation of p53 by MDM2. Created and adapted using Biorender.com[31]

Studies have shown that p53 mutations are widespread in many types of human cancers, such as breast, colorectal, lung, sarcoma, and others.[32–35] In colon cancer, p53 mutations are linked to worse survival outcomes and poorer prognosis.[36] There are many other roles that p53 plays in the human body, including cell cycle control, programmed cell death, DNA repair, and maintenance of genomic stability.[24] The organs with the most p53 mutations include ovaries, colon, rectum, and lungs, as shown in Figure 4.[25] Similarly, in dogs, p53 mutations have been linked to many different types of cancers such as canine mammary tumors,[37] canine hemangiosarcoma,[17] osteosarcomas,[22] and others.[38]

Structure of p53 Protein. The human version of the p53 protein contains multiple domains and 393 amino acids. It contains an N-transactivation domain, a proline-rich domain, a DNA binding domain linked to a tetramerization domain, and a C-regulatory domain.[25,40,41] More than 40% of p53 protein regions, majorly in the DNA binding and tetramerization domains, are

intrinsically disordered, allowing p53 to be modular and interact with various partner proteins.[42] The domains of canine p53, as classified by the InterPro domain identification system, are shown in Figure 5 below.

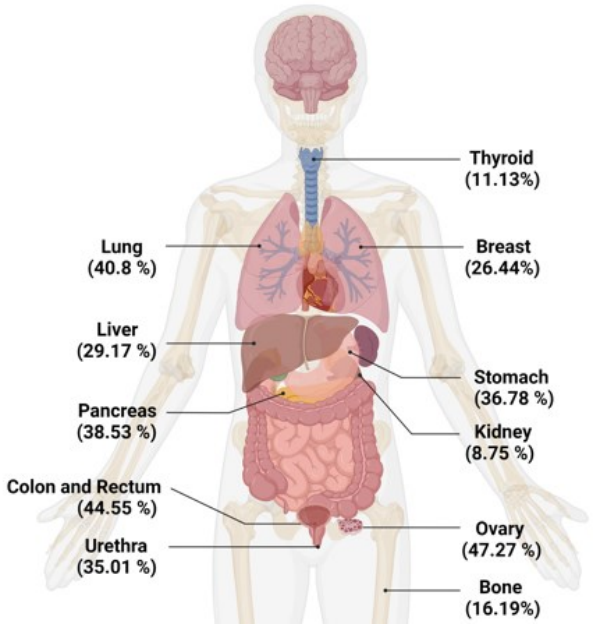

Figure 4: Mutation frequency of TP53 in different organs.[25] Created and adapted using Biorender.com[39]

The structure of the core domain of human p53 in complex with DNA was first solved by Cho et al. in 1994 (RCSB Protein Data Bank (PDB) – 1TSR, Figure 6).[44] Although hundreds of solved structures of the different domains of human p53 are available on the RCSB PDB, the same is not true for the canine version, as many important canine cancer proteins and genes have not been extensively studied.

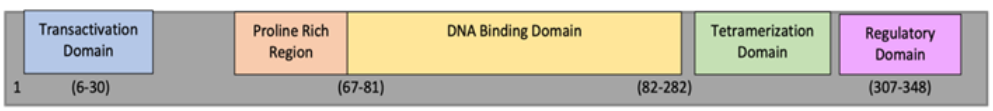

Figure 5: Domains of canine p53 based on InterPro domain identification.[43]

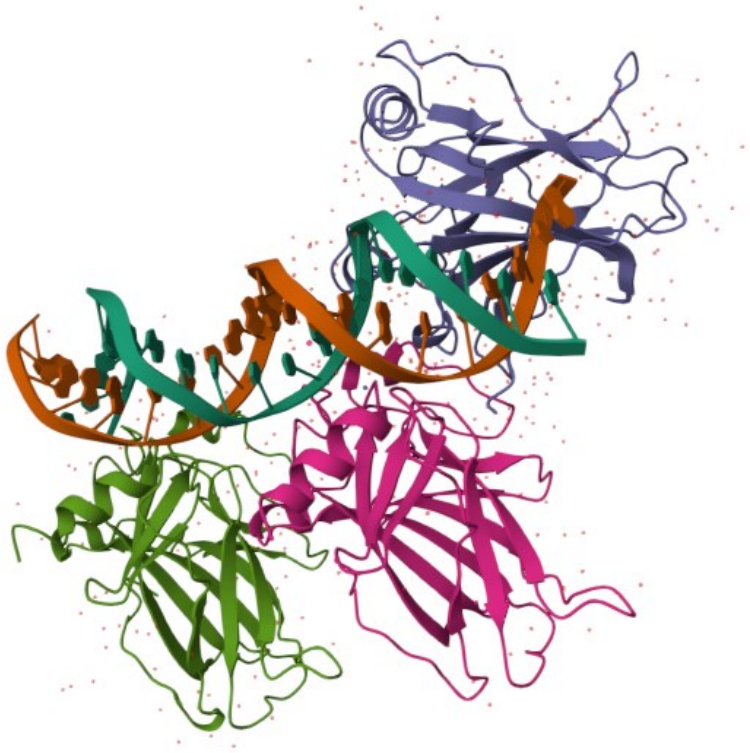

Figure 6: The structure of human p53 core domain in complex with DNA. RCSB PDB - 1TSR. Image adapted from PDB. Original structure by Cho et al., 1994[44]

Comparison of Human and Canine p53. The canine p53 protein is 80.5% identical to the native sequence of human p53 (Figure 7). The greatest variability is in the first, N-transactivation domain. Our domain of interest is the DNA binding domain, which is 89.6% identical to the human version (Figure 7). Most TP53 mutations are missense mutations (73%) where a single amino acid substitution occurs, and the most common domain for this mutation is the DNA binding domain.[45] Furthermore, the DNA-binding domain accounts for over 90% of all p53 mutations in humans.[46] Mutation within the DNA binding domain causes the protein to lose its ability to bind specific targets and act as a tumor suppressor. G245, R249, R282, and R283 are "hot spot" mutations occurring at the highest frequency in humans.[47] Among the hot spot mutations, G245 and R249 mutations cause structural changes that lead to reduced protein

thermostability, reduced folding at physiological temperatures, and the inability to bind DNA.[25] The G245 mutations seen in humans (G233 in dogs) are also found in canine mammary tumor cell line (CMT7) and canine oligodendrogliomas.[38] Studies on canine cancers show that many canine p53 mutations are similar in frequency and location to human p53 mutations,[48] although large model pan-genome studies for canine cancers have not been done in a similar volume to human cancers.

As most of the structural biology research on oncoproteins and tumor suppressor proteins has focused on the DNA binding domain, this research focused on the same for two reasons. First, it is challenging to determine a high-resolution structure of the full-length p53 due to highly intrinsically unfolded regions within p53 such that the best structure available of the full-length human p53 was solved at a 4.6 Å resolution using cryo-electron microscopy.[25,49] This resolution is significantly lower than many other protein structures with 1-2 Å resolutions. Second, the DNA binding domain is responsible for the role of p53 as a transcription factor and its interaction with other proteins,[25] and therefore the biggest contributor to its tumor suppressive functions. These two reasons form the basis of DeVore lab's focus on the DNA-binding domain of the canine p53 protein.

Research Goals. The main objective of this research was first to express and purify a sufficient quantity of the canine p53 protein, as previous studies within the DeVore lab could not express and purify the protein in sufficient quantities successfully. The purified canine p53 proteins were then used to study their binding to human DNA sequences. Human DNA sequences were used for binding studies as the identities of sequences binding to canine p53 proteins are not yet known. As the sequences of the DNA binding domains of both the canine and human p53 were 90% similar, both might be able to bind to human DNA sequences, albeit

with different binding affinities. The binding affinities would provide insights into the structural similarities between the human and canine versions of p53. Electrophoretic mobility shift assays (EMSA) and Surface Plasmon Resonance (SPR) studies were developed to study the binding affinity of purified canine p53 protein with human DNA sequences. Although it was possible that studying canine p53 protein may somewhat extend our understanding of the human counterpart, the EMSA binding experiments showed a drastic difference in binding affinity, showing a substantial structural difference between the two proteins.

The secondary aim was to crystallize the protein and solve the three-dimensional structure of the canine p53 protein using X-ray diffraction studies. There are countless structures of the human p53 version in the protein data bank, but there is a dearth of similar structures for the canine versions. The solution of the 3D structure of canine p53, the EMSA results, and SPR data should help us provide insights into the canine p53's structure. Lastly, future studies would focus on creating known mutations within the canine p53 protein (G233A, which corresponds to G245 in humans) to understand how the mutations affect the protein's structure and function.

Methods

Transformation of p53. An *Escherichia coli* (*E. coli*) codon-optimized version of the canine p53 DNA-binding domain (residues 82-280) subcloned into a pET28a vector containing an N-terminal six-histidine tag was purchased from Twist Bioscience (Figure 7). The vector also contained a thrombin cut side utilized for the removal of the six-histidine tag. Two microliters of the circular plasmid were added to 50 microliters of chemically competent *E. coli* BL21(DE3) cells. The transformation was induced using the heat-shock treatment, and a one-hour outgrowth step with NZY+ media was added to ensure a successful transformation. The cells were pelleted

and added to culture plates containing Lennox broth (LB)-kanamycin plates containing 50 µg/mL kanamycin. (see Appendix A for the full profile of canine p53 protein).

Figure 7: Protein alignment of human p53 and canine p53. The DNA-binding domain is highlighted in yellow — alignment done using Align from www.expasy.org.

Expression of p53. A single colony was used to start an overnight starter culture containing 200 milliliters of LB media with 50 µg/mL kanamycin at 37 ˚C. The next day, three equal portions of the starter culture were transferred to three one-liter flasks containing Terrific Broth (TB) media and grown for a few hours until the optical density at 600 nanometers (OD_{600}) reached an absorbance value of 0.8, following which, protein expression was induced using one mM isopropyl β-d-1-thiogalactopyranoside (IPTG). The temperature was reduced to 28 ˚C and grown for two days. Following that, the cultures were centrifuged at 5000 rpm for 20 minutes, and the pellets were resuspended in a resuspension buffer (100 mM potassium phosphate, pH 7.4, 10% glycerol, 300 mM NaCl, and 10 µM $ZnCl_2$). The resuspended cells were then lysed

using 4- 30 second sonication cycles and centrifuged at 20,000 rpm for 20 minutes to separate soluble protein from cell debris. The pellets were discarded, and the supernatant (lysate) was collected for purification.

Purification of p53 Protein. The lysate was loaded onto a Nickel-nitriloacetic acid (Ni-NTA, Gold biotechnology) affinity column for primary purification. The protein was washed with a wash buffer (100 mM potassium phosphate, pH 7.4, 10% glycerol, 100 mM NaCl, 10 mM imidazole, and 10 μM $ZnCl_2$) and eluted using an elution buffer (100 mM potassium phosphate, pH 7.4, 10% glycerol, 200 mM imidazole, 10 μM $ZnCl_2$, and 10 mM ethylenediaminetetraacetic acid - EDTA). The elution fraction from primary purification underwent a secondary purification step using ion-exchange chromatography. A carboxymethyl cellulose (CM) column was used to bind the primary purified protein diluted seven times with ion-exchange wash buffer (50 mM potassium phosphate, pH 7.4, 10% glycerol, and 10 μM $ZnCl_2$) to reduce the concentration of NaCl. The protein was then eluted using an elution buffer (50 mM potassium phosphate, pH 7.4, 10% glycerol, 10 μM $ZnCl_2$, and 500 mM NaCl). A portion of this eluent underwent tertiary purification using thrombin digestion for comparative studies using a 1 unit of thrombin per 100 mg of protein ratio. After the overnight digestion, the untagged protein was separated using a gravity Ni-NTA column. A list of buffers used is shown in Table 1.

Table 1: List of buffers used for p53 protein purification steps.

Buffer (pH 7.4)	Potassium Phosphate (mM)	Glycerol (%)	Sodium Chloride (mM)	Zinc Chloride (μM)	Imidazole (mM)	EDTA (mM)
Ni-NTA Resuspension	100	10	300	10	-	-
Ni-NTA Wash	100	10	100	10	10	-
Ni-NTA Elution	100	10	-	10	200	10
CM Wash	50	10	-	10	-	-
CM Elution	50	10	500	10	-	-

Concentration and Quantification. The purified protein was then quantified using an absorbance measurement (Shimadzu UV-2101 PC spectrophotometer) from a 700 to 250 nm range. Beer's law was used to calculate the concentration and yield of protein using the absorbance reading at 280 nm. The protein was then concentrated using Amicon ultra-4 centrifugal filter units (10 kD) until the final volume of protein was reduced to less than 500 µL. The protein was then re-quantified, aliquoted, and stored at -80 ˚C.

Crystallization. The purified and concentrated protein was then used to set up plates of crystals via the hanging drop vapor diffusion method. One microliter of protein was combined with an equivalent amount of well solution and covered using a coverslip. Over 100 different crystal screen conditions were used using commercially available kits such as Hampton Research Crystal Screen I and II. The plates were periodically checked for crystal growth, and promising conditions were further optimized to obtain better-quality crystals for future XRD experimentation.

Electrophoretic Mobility Shift Assay (EMSA). A make-shift gel assay was used to determine the binding affinity of the canine p53 protein with human hairpin DNA constructs (Figure 8). These hairpin DNA constructs were known to bind with the human p53. These constructs have two consecutive dodecamer DNA repeats joined by five cytosines.

The EMSA kept the amount of DNA in each well of the gel constant, with increasing amounts of protein in consecutive wells (Table 2). The experiments were conducted using a Tris-Borate-EDTA (TBE) buffer system and TBE gels. The reagents in Table 2 were combined using a 1x binding buffer (0.01mM HEPES, 0.15M NaCl, 3 mM EDTA, 1 mM DTT, and 10% glycerol). The gel was allowed to run at 100 V for 45 minutes before staining with SYBR-safe stain to visualize under UV light. A ChemiDoc touch gel reader was used to image the gel, and

the image was analyzed using the ImageJ software, and the results were compiled using the one-site specific binding feature in GraphPad Prism.

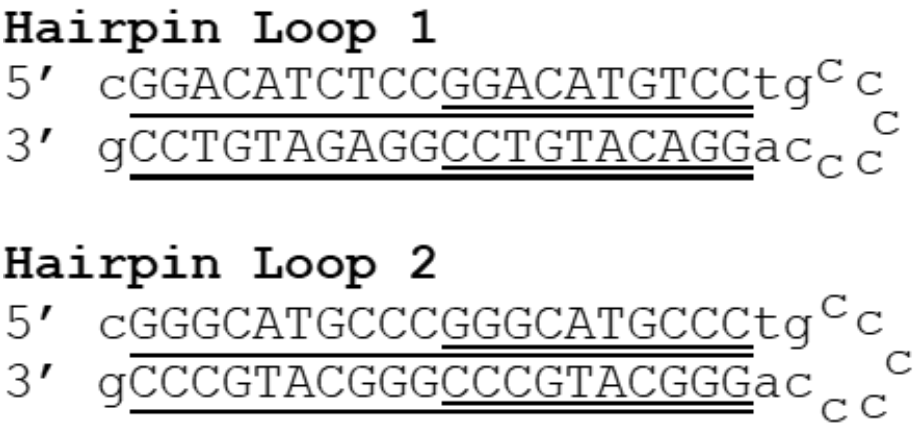

Figure 8: DNA hairpin constructs.

Surface Plasmon Resonance (SPR) Experiments. SPR experiments were conducted using a Reichert SR7000DC dual-channel instrument with a carboxymethyl dextran (CM) chip. Phosphate Buffered Saline with 0.005% Tween 20 (PBST) was used as a buffer and as a sample diluent. A mixture of 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) and N-hydroxysuccinimide (NHS) in a 4:1 ratio was used for the conjugation and activation of the chip. The p53 protein (1 mg/mL) was then immobilized on the chip, followed by ethanolamine (1M, pH 8.5) blocking to cover all remaining binding sites on the chip. After the blocking, varying concentrations of DNA (50 nM-5 μM) were used to test the binding affinity of the protein and the hairpin constructs. The chip surface was regenerated using a 10 mM glycine-HCl buffer. Conversely, a different protocol where the DNA hairpin loops were immobilized on the chip first, followed by varying protein concentrations, was also tried to compare the impact of appending the protein on the chip vs. DNA on the chip.

Table 2: Sample concentrations and volumes for an EMSA.

Final Protein Concentration (nM)	p53 amount (uL) (120 uM)	1x Binding Buffer (uL)	Hairpin DNA 500 nM stock(uL)
0	0	10	10
1500	0.25	9.75	10
3000	0.5	9.5	10
6000	1	9	10
9000	1.5	8.5	10
12000	2	8	10
15000	2.5	7.5	10
18000	3	7	10
21000	3.5	6.5	10
24000	4	6	10
27000	4.5	5.5	10
30000	5	5	10
60000	10	0	10

X-Ray Diffraction. Once crystals were sufficiently grown, they were added to a cryo solution that consisted of 50% glycerol or ethylene glycol and 50% mother liquor and looped. The loop was then affixed to the pedestal of a Rigaku Xtalab Synergy-S, and flash-frozen at 100 K. X-ray beams were passed through the crystal for 20 seconds for preliminary data collection. Unfortunately, none of the crystals diffracted enough for full data collection and processing.

Results

Expression and Purification of Canine p53. A total of eight rounds of expression and purification were conducted for canine p53 in JM109 and BL-21(DE3) *E. Coli* cells. Over the different rounds of expression, it was noted that the canine p53 expresses much better in the BL-21 cell line than the JM-109 cell line. The initial rounds of expression were done in tris hydrochloride buffers, but the later rounds were switched to potassium phosphate buffers. It was also noted that potassium phosphate buffers led to a significant increase in protein yield. During

the first two rounds of expression and purification with tris buffers in the JM-109 cell line, the net p53 protein yield after purification was less than 1 mg/L which was not ideal. After switching to potassium phosphate buffers with BL-21(DE3) cells and optimizing for temperature and duration of growth, the net yield was improved to 18 mg/L. The best protocol consisted of a 3 liter/2-day growth in Lennox Broth (LB) media with the growth temperature set to 28 ˚C after induction. In addition to the native canine p53 protein, a G233A mutant (homologous to G245 hotspot mutation in humans) was also expressed, but the yield after purification was negligible. Due to the low yield of the mutant, the project focused on the native canine p53 protein. In the future, it may be interesting to study more mutants to understand the impact of different mutations on the protein's overall structure.

Primary Purification. The first purification step was done using a Ni-NTA affinity column which operates on the principle of the six-histidine tag binding with the Ni-NTA column with strong affinity. After binding, the protein can be washed using a higher concentration of imidazole in the elution buffer, displacing the protein and binding to the column. An absorbance spectrum of the Ni-NTA purification step is shown in Figure 9 below.

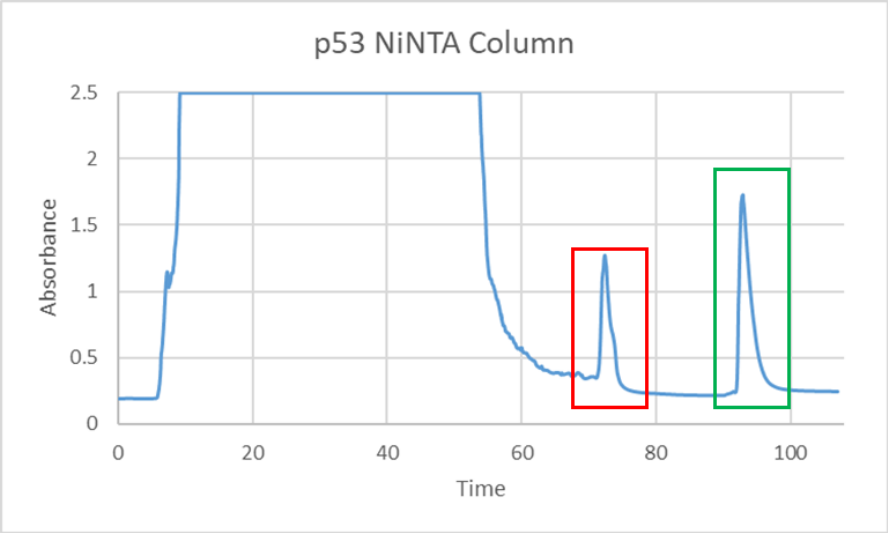

Figure 9: Canine p53 purification using a Ni-NTA column. The red and green boxes denote the wash and elution fractions, respectively.

Secondary Purification. The secondary purification was done using a negatively charged carboxymethyl-cellulose (CM) column. As the p53 protein is positively charged, it binds with the column and can be eluted using a high concentration of sodium chloride. An absorbance spectrum of the CM purification step is shown in Figure 10. The purity of the protein can then be verified using a Sodium Dodecyl Sulfate- Poly Acrylamide Gel Electrophoresis (SDS-PAGE) step (figure 11).

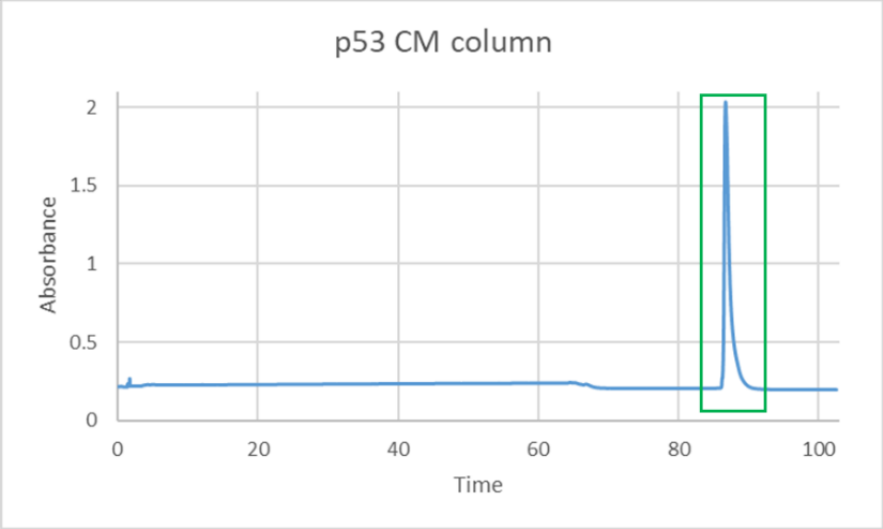

Figure 10: Canine p53 purification using a CM column. The green box denotes the elution fraction.

Tertiary Purification. The tertiary purification step involved using a thrombin digestion procedure to cleave the six-histidine tag from the p53 protein. Following an overnight thrombin digest, the resulting protein solution was separated using a reverse gravity Ni-NTA column. The cleaved protein did not include a six-histidine tag, so it flowed through the column without binding, whereas the free histidine tags bound to the column. Although the tertiary purification step results in a purer protein, over two-thirds of the protein was lost during this step. Due to this,

most of the p53 protein preparations were done with only the first two purification steps (without the thrombin digestion procedure).

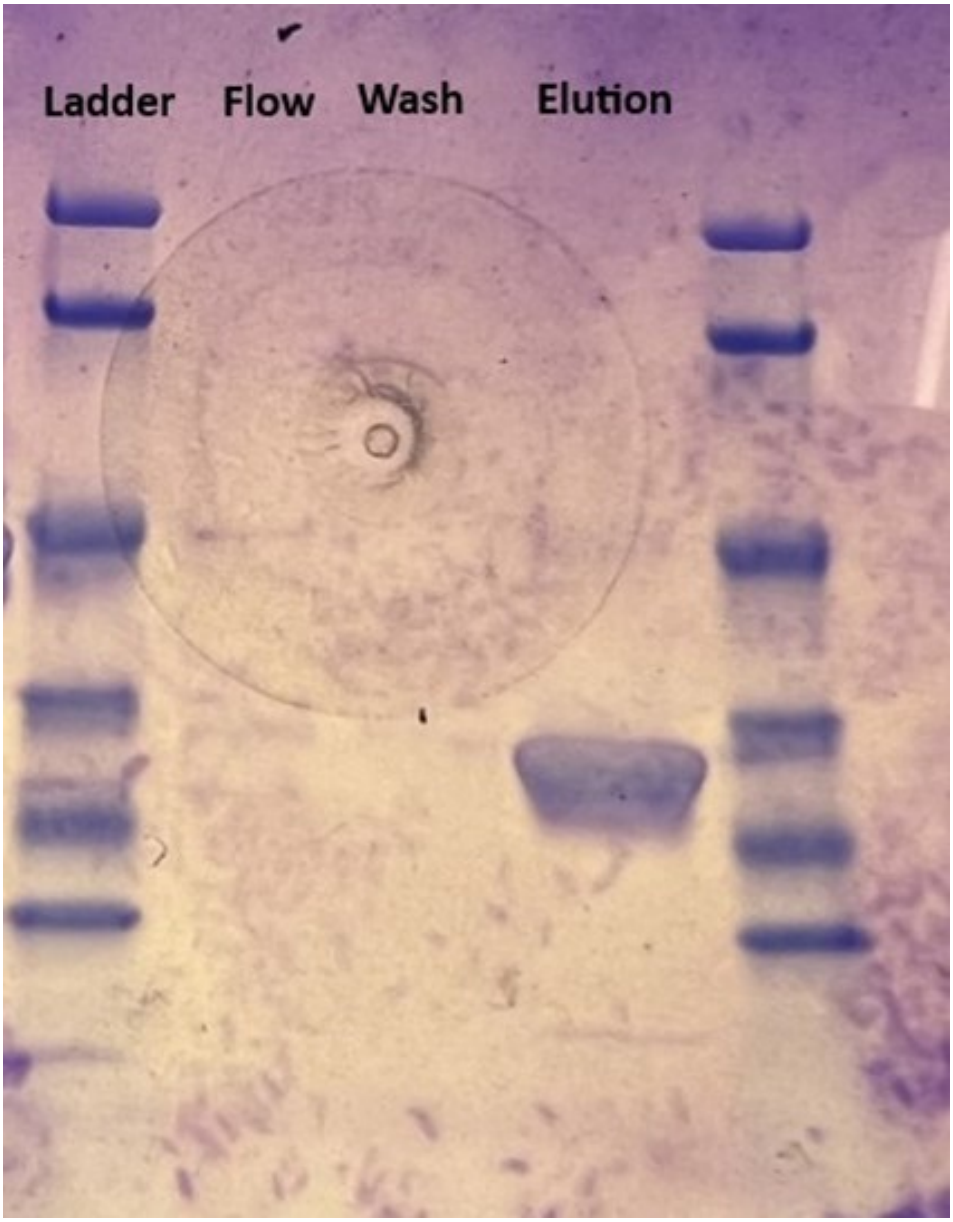

Figure 11: SDS-PAGE of canine p53 protein after ion-exchange chromatography.

Electrophoretic Mobility Shift Assay (EMSA). A total of eight EMSA runs were done to create a novel protocol based on the research by Seo et al., 2019.[50] The first four runs were

done on 0.5-1% Agarose gels to test the idea of canine p53 protein binding to target DNA. As seen in Figure 12, the binding can be seen on Agarose gels. In Figure 12, the green boxes show protein bound to DNA (heavier) which did not move far from the initial position, and the red boxes indicate free DNA, as protein concentration for those wells is lower, and there is not enough protein to bind to the hairpin loops. The first two wells show hairpin loop #1, and the last two show hairpin loop #2.

Agarose does not provide the best conditions for proteins to move, and therefore, 5% pre-cast Tris-Borate-EDTA (TBE) gels from Bio-Rad were used (Figures 13 and 15). After a few runs to optimize the best concentrations for the EMSA, the TBE gels were visualized using a ChemiDoc touch gel reader, and the images were processed using ImageJ to find the raw integrated density (RID) of the free DNA bands.

Once the RIDs were quantified, % binding was calculated by subtracting the smallest RID from all RID values, resulting in relative RIDs. Each RID was then divided by the largest relative RID and multiplied by 100, giving a percentage RID. Each percentage RID was then subtracted from 100, resulting in % binding. The % binding values were then plotted against protein concentrations, and the binding affinity (K_d) between the canine p53 and hairpin loops was calculated using a one-site specific binding function on GraphPad Prism. The binding affinity plot for canine p53 protein and hairpin loop #1 is shown in Figure 14.

The EMSA gel for the interaction between canine p53 protein and DNA hairpin loop #2 is shown in Figure 15. The K_d value for the interaction between the canine p53 protein and hairpin loop #2 was 12.54 μM (Figure 16), indicating that the canine p53 protein had a slightly greater affinity for hairpin loop #2 than hairpin loop #1. Efforts to understand the difference in K_d between the two hairpin loops are ongoing through further experimentation.

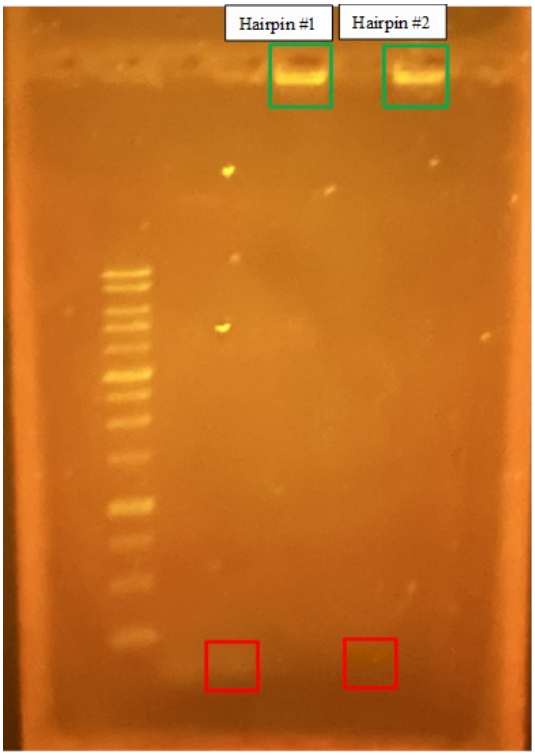

Figure 12: EMSA with 1% agarose gel visualized under a dark reader. The green boxes indicate high protein concentration, and the red boxes indicate lower protein concentration. The concentration of DNA is kept constant between all wells.

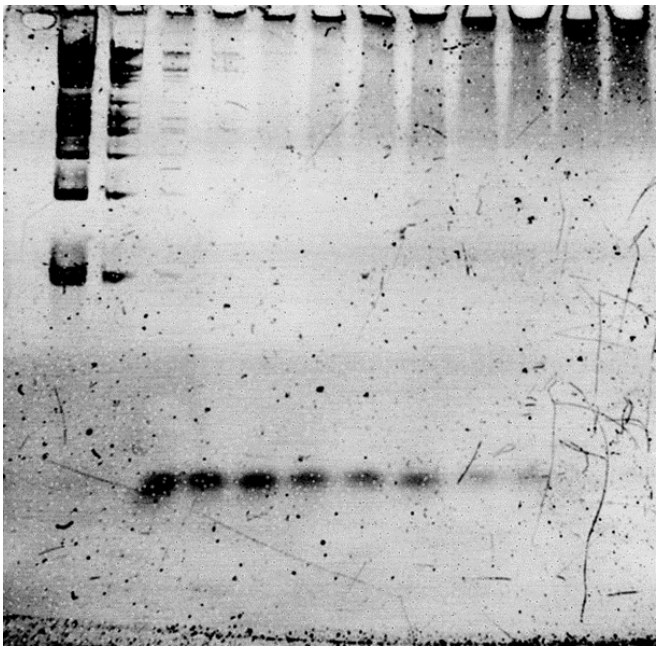

Figure 13: EMSA of canine p53 and human DNA hairpin loop #1 using 5% TBE gel. Gel visualized using a ChemiDoc Touch gel reader. The concentration of protein increases from the left to the right with constant DNA concentration throughout.

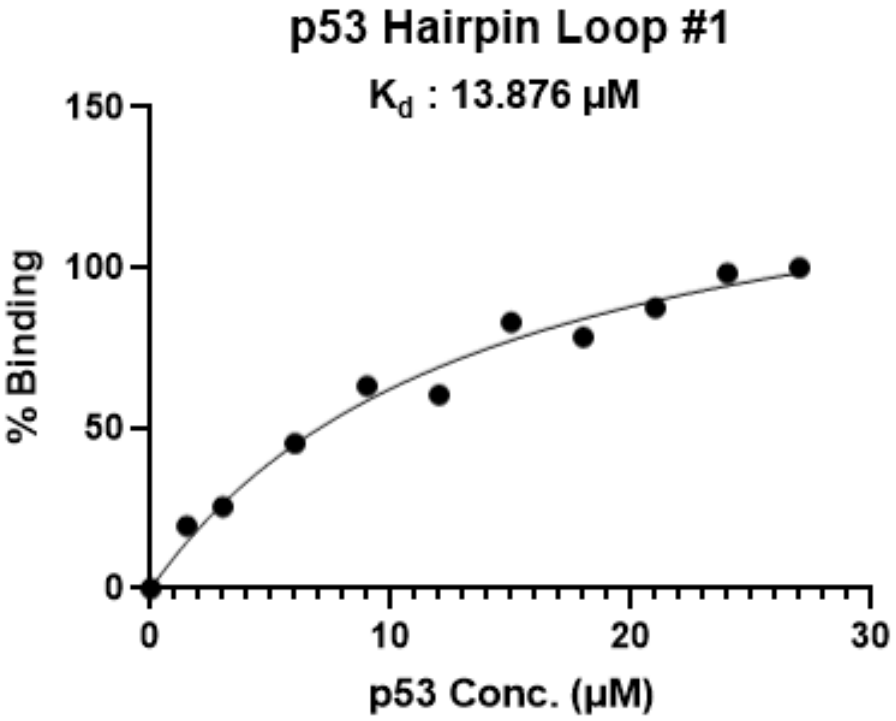

Figure 14: K_d value for binding between canine p53 and hairpin loop #1.

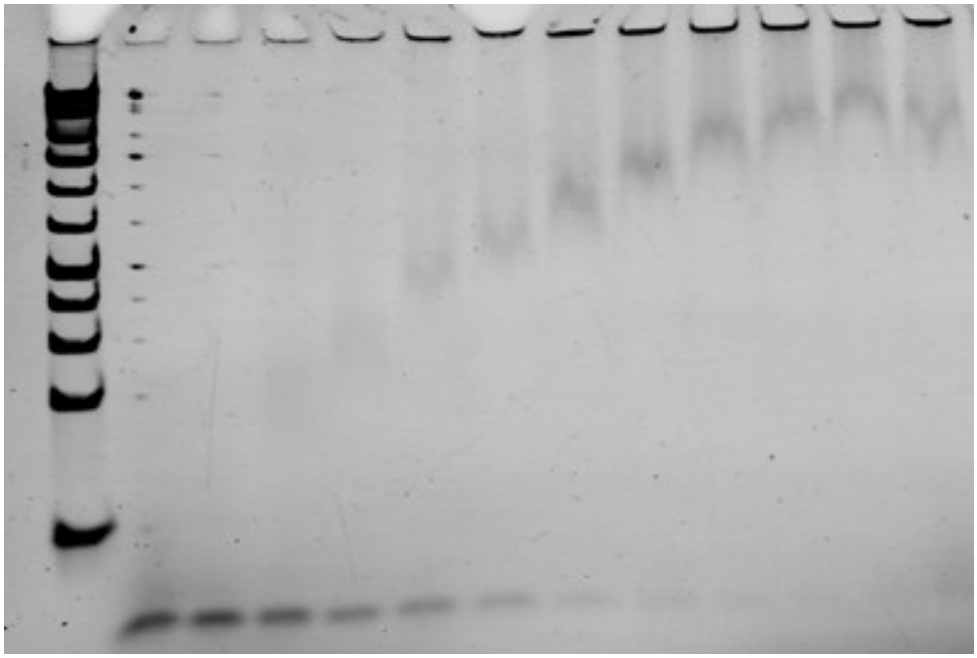

Figure 15: EMSA of canine p53 and human DNA hairpin loop #2 using 5% TBE gel. Gel visualized using a ChemiDoc Touch gel reader. The concentration of protein increases from the left to the right with constant DNA concentration throughout.

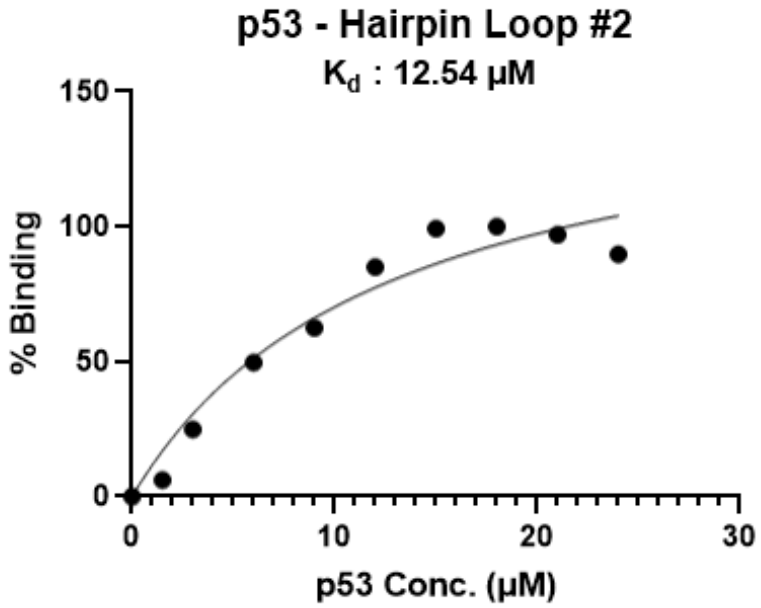

Figure 16: K_d value for binding between canine p53 and hairpin loop #2.

Surface Plasmon Resonance (SPR) Experiments. Nine SPR experiments were conducted to study the binding affinity between the DNA hairpin loops and the canine p53 protein. Eight experiments focused on binding the p53 protein to the CM-dextran chip first, followed by the DNA hairpin loops to study binding.

The EDC-NHS activation of the dextran chip results in the formation of N-hydrosuccinimide esters that are highly reactive.[51] When a protein is flown on this chip, the esters react with the amine groups on the protein to form covalent amide linkages resulting in the appending of the protein onto the chip surface.[51] The large increase can see this protein binding in the blue line (sample channel). As the blue channel increased by 20,000 µRIUs without any change in the red channel (reference channel), it meant that 20,000 micro refractive index units (µRIUs) of protein bound to the chip surface. Concentrations ranging from 0.5 mg/mL to 2 mg/mL of protein were used to append the protein to the chip. Then, 50 nm to 5 µM of hairpin loops were run through the chip.

As seen in Figure 17, upon the addition of DNA to the surface of the chips, there is an equivalent increase in both the sample and reference channels, meaning that no DNA was bound to the appended protein. The full chronological order of an SPR experiment is shown in Figure 17, starting from the EDC-NHS activation of the chip to the addition of DNA. Unfortunately, despite several protocols and consulting literature changes, the experiments failed to show DNA binding to the canine p53 protein.

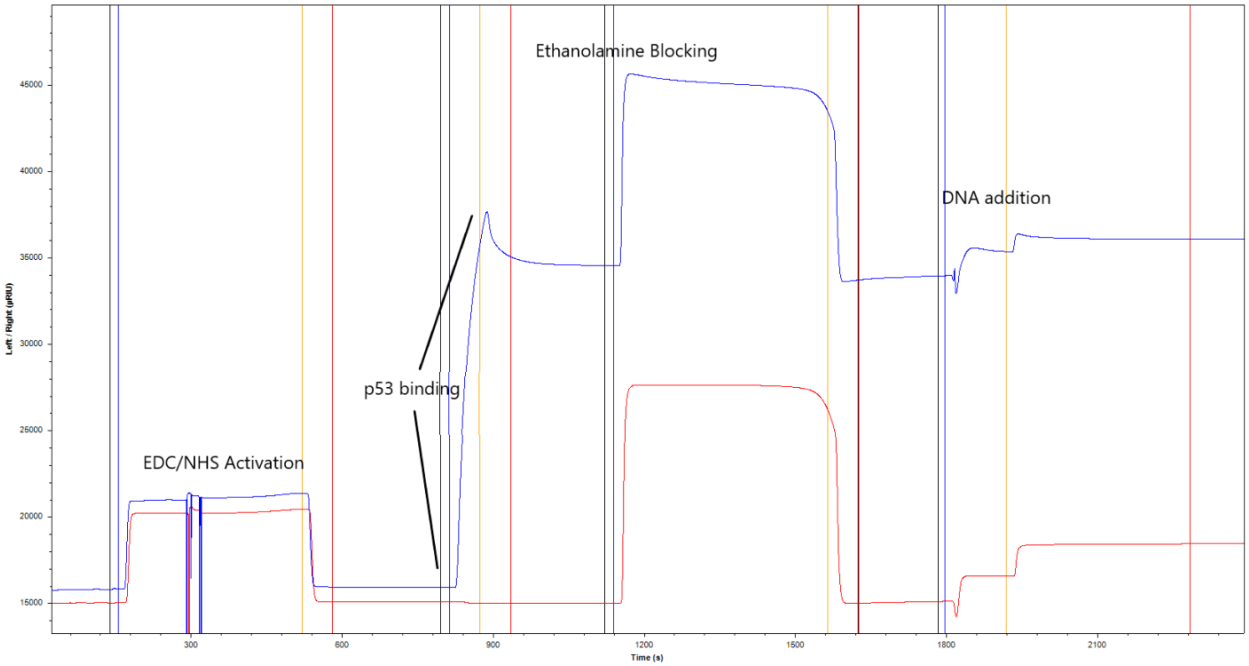

Figure 17: SPR data. Blue indicates the sample channel, and red indicates the reference channel. Note: the anomaly during EDC/NHS activation, as shown by dramatic dips reflects the presence of gas bubbles in the line, and is resolved on its own within a few seconds.

The last SPR experiment consisted of trying to bind the DNA first, followed by the protein. There was no relative change in the µRIU values (both sample and reference channel moved the same amount) upon flowing the DNA on the chip, meaning DNA could not bind to the chip first (Figure 18). The rationale behind first binding the DNA to the chip was to ensure

that the conformation of the DNA binding domain of canine p53 was not altered upon binding to the chip, resulting in the unavailability of the exposed binding sites on the protein for DNA to bind. Unfortunately, the SPR experiments did not give us any meaningful data. It is possible that the concentration of DNA used for experimentation was not high enough to bind to the chip successfully. Some other potential reasons for the failure of the experiments include incompatible buffers and surface contamination of the chip. As the experiment to bind DNA was only conducted once, future experimentation is needed with different protocols to improve the chances of success.

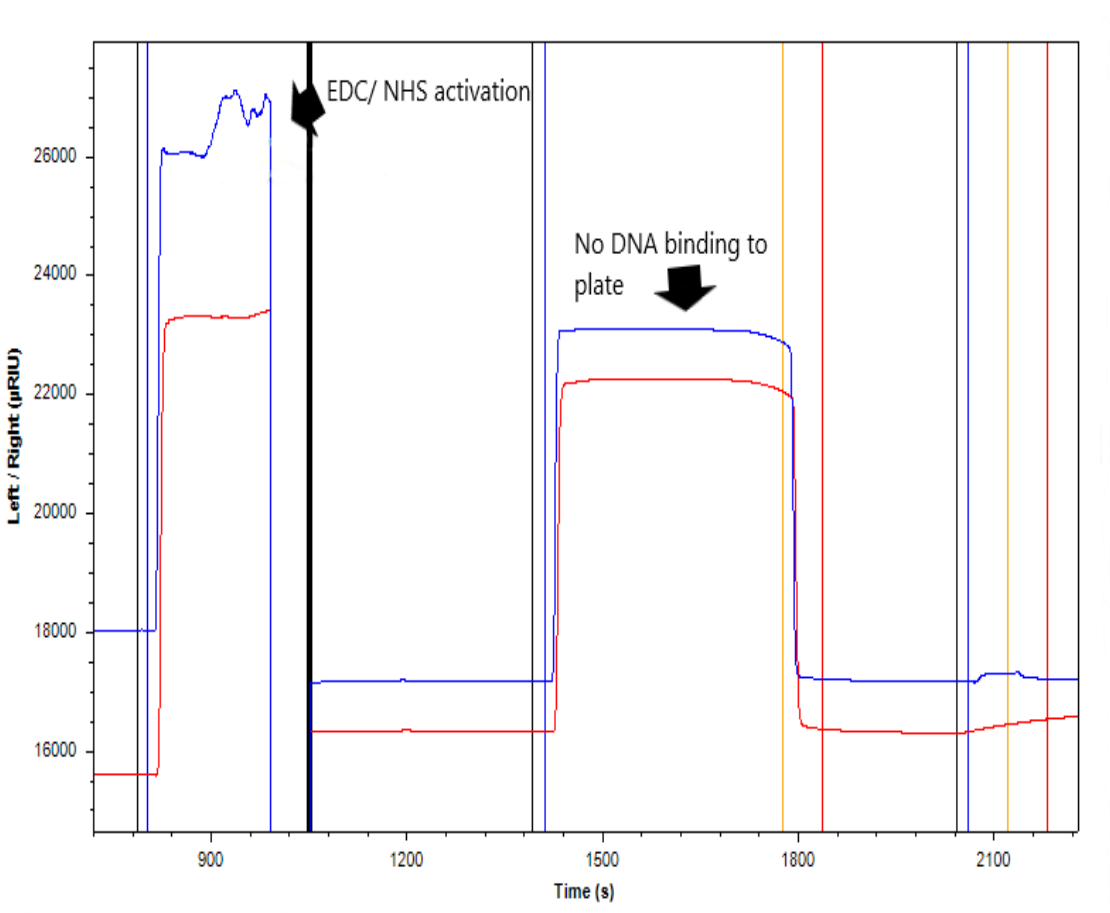

Figure 18: SPR data with DNA binding to the chip first. No relative change in the blue channel (sample) relative to the red channel (reference), indicating no DNA binding.

Crystallography. After purification and concentration, the proteins were screened using Hampton Research Crystal Screen I and II to check for crystal hits (Appendix B and C). The conditions that led to crystals were optimized to create better quality crystals. Over 200 unique conditions were used to promote crystal growth, and the best crystals were used for X-ray diffraction (XRD) experiments.

The conditions that provided crystals used for screening were Hampton Research Crystal Screen I (CSI) condition 18 (0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, and 20% w/v Polyethylene glycol 8,000), and condition 24 (0.2 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, and 20% w/v 2-Propanol) as well as Hampton Research Crystal Screen II (CSII) condition 41 (0.01 M Nickel (II) hexahydrate, 0.1 M Tris pH 8.5, and 1.0 M Lithium sulfate monohydrate). These were optimized to create better-quality crystals. A full list of screens used is in Appendix B and C. The best-optimized conditions that led to crystals for p53 are were:

- 0.1 M sodium cacodylate trihydrate pH 6.5, 0.25 M Magnesium acetate tetrahydrate, and 25% polyethylene glycol 8,000
- 0.1 M sodium cacodylate trihydrate pH 6.5, 0.275 M Magnesium acetate tetrahydrate, and 25% polyethylene glycol 8,000

Unfortunately, none of the crystals diffracted well enough to collect a full dataset and solve the structure of the proteins. The efforts to crystallize and optimize more conditions are ongoing. An example of a p53 crystal in the well (Figure 19-A) and a looped p53 crystal (Figure 19-B) are shown below. Some crystals were also grown in the presence of DNA substrate to see if the substrate stabilizes the protein and makes it easier to crystallize. An example of such a crystal is shown below in Figure 20. The best condition that led to crystals with DNA substrate

was CSI- condition 18 (0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, and 20% w/v Polyethylene glycol 8,000). Interestingly, the addition of substrate led to the quicker formation of crystals, potentially meaning that DNA did stabilize the protein to crystallize better. Adding a substrate may cause the protein to stabilize in a lower energy conformation, explaining the reduction in the time taken to form crystals.

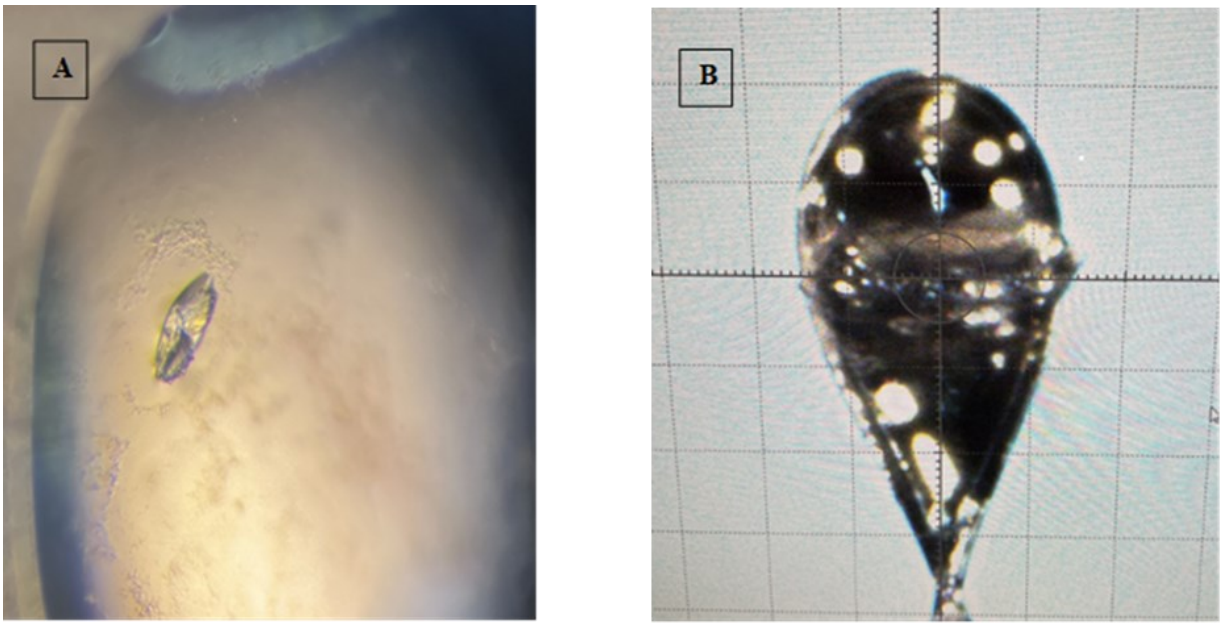

Figure 19: Canine p53 crystals. A – crystal in a well plate. B – looped crystal on XRD pedestal.

Roughly 25 p53 crystals were used for X-ray diffraction experiments. Unfortunately, none of them diffracted sufficiently to collect a full dataset for a structure solution. The ones that did diffract were salt crystals and had smeared patterns (Figure 21). It is likely that a majority of the crystals screened were salt crystals and not protein crystals. Unfortunately, no diffraction was seen in crystals produced using a DNA substrate. More hanging drop experiments need to be conducted to increase the odds of getting a diffracting crystal.

Figure 20: Canine p53 crystal with DNA substrate. Condition: 0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, and 20% w/v Polyethylene glycol 8,000

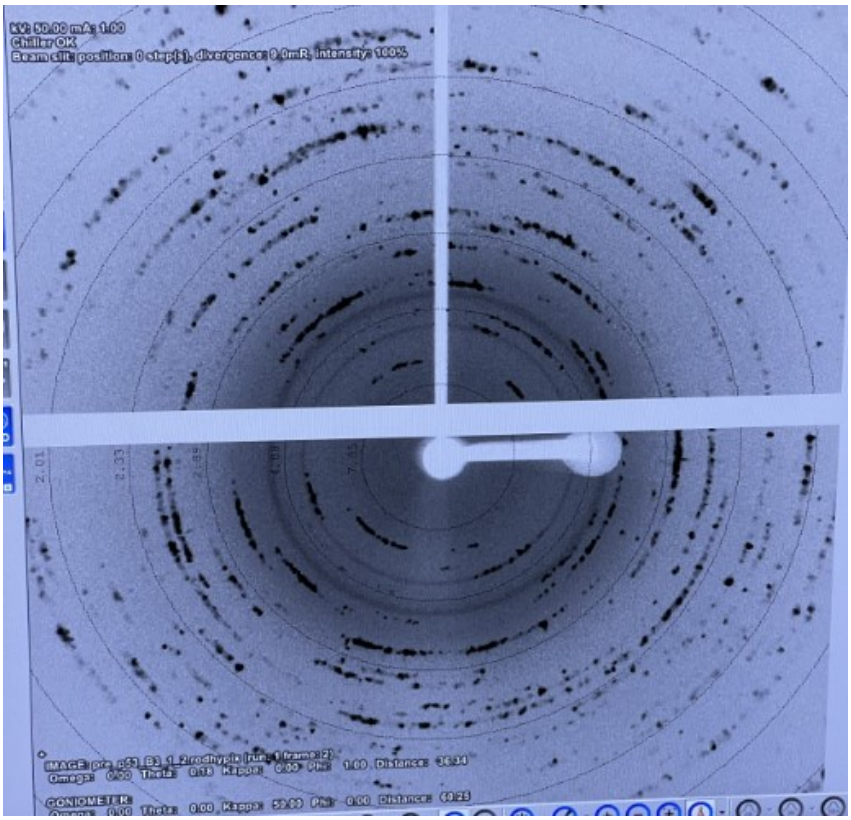

Figure 21: Canine p53 X-ray diffraction. The smeared pattern suggests salt crystals and not protein crystals.

Discussion

There were eight rounds of expression and purification of the canine p53 protein. It was discovered that the BL-21 cell line and the potassium phosphate system are the best conditions for a higher protein yield. The optimization of conditions surrounding the expression led to significant improvement in protein yield from less than 1 mg/L to 18 mg/L. One attempt to purify the G233A mutant resulted in negligible amounts of protein. As G233A is a structural mutation that leads to folding and thermostability issues at room temperature,[25] it is possible that such a protein is less stable in conventional buffers and leads to aggregation, which could explain the lower protein yield for canine G233A mutant. Future experiments need to be repeated with different conditions to grow the G233A mutant and other mutations prevalent in canine cancers. Such work may provide valuable insights into the impact of different mutations on the structure and function of the protein.

The purification steps were very successful despite the thrombin digestion step leading to substantial losses in the protein yield. In the future, the digestion conditions could be optimized by increasing the amount of thrombin and the digestion timeframe to ensure proper cleavage of the six histidine tags to reduce losses and improve protein purity. Thrombin-digested proteins are the purer form of protein due to the absence of the six-histidine tag and are a better representative model for canine p53 protein.

The EMSA showed that DNA hairpin loop #1 binds to the canine p53 with a K_d value of 13.876 μM. The binding affinity experiments of hairpin loop #2 showed a K_d value of 12.54 μM. The EMSA experiments show that hairpin loop #2 binds to the canine p53 with a slightly greater affinity. Experiments with the same hairpin loops with human p53 protein reveal significantly different binding affinity. Although the experiments were performed using slightly different

protocols, the K_d value for the binding affinity of hairpin loop #1 to human p53 was 14 nM, and the K_d value for the binding affinity of hairpin loop #2 to human p53 was 20 nM.[52] The EMSA comparison shows a roughly 1000x difference in the binding affinity of human and canine p53 proteins to human DNA sequences, indicating significant structural differences between the two proteins. Without a solved structure of the canine protein, it is challenging to accurately compare the structures as there is a 10% sequence difference and even a single mutation at the correct position can render a protein non-functional. Despite this, the EMSA experiments help us conclude the significant structural differences between the human and canine versions of p53, indicating that the canine p53 protein may not offer a suitable structural model for human comparisons. More efforts are ongoing to crystallize the protein for XRD experiments better so that a three-dimensional structure of the canine p53 can be obtained.

The SPR experiments should explore different chipsets and buffer systems to study the binding affinity between the canine p53 protein and human DNA sequences. The only SPR study conducted with human p53 and DNA sequences also included antibodies and Enzyme Linked Immuno Sorbent Assay (ELISA) comparisons to study p53 binding to DNA.[53] Future work may need to adopt similar approaches for SPR experiments.

Crystal conditions must be optimized further, and more crystals must be screened to collect a complete dataset. The addition of DNA along with protein in hanging drop vapor diffusion experiments led to a decrease in the time it took for the protein to crystallize. This could mean the protein was stabilized by adding DNA to crystallize faster. Only a few such experiments were conducted, and more experiments are needed to grow more crystals. Further optimization of growth conditions such as temperature, protein concentration, and a wider range of well solutions should also be explored to improve the quality of the crystals. Despite the many

challenges and failures, this work laid the foundation for future experiments. Further efforts in optimizing different crystal conditions will certainly bear fruit. As my research mentor often says, "It only takes one good crystal to solve the structure."

References

(1) Van Dam, P.-J.; Van Laere, S. Molecular Profiling in Cancer Research and Personalized Medicine. In *Oxford Textbook of Cancer Biology*; Pezzella, F., Tavassoli, M., Kerr, D. J., Eds.; Oxford University Press, 2019; pp 347–362. https://doi.org/10.1093/med/9780198779452.003.0024.

(2) Alam, S.; Nasreen, S.; Ahmad, A.; Darokar, M. P.; Khan, F. Detection of Natural Inhibitors against Human Liver Cancer Cell Lines through QSAR, Molecular Docking and ADMET Studies. *Curr. Top. Med. Chem.* **2021**, *21* (8), 686–695. https://doi.org/10.2174/1568026620666201204155830.

(3) Vercelli, M.; Quaglia, A.; Parodi, S.; Crosignani, P.; Itapreval Working Group:; Micheli, A.; Gatta, G.; Sant, M.; Rossi, A. G.; Francisci, S.; Saltarelli, S.; Dell'Era, L.; Gasparre, N.; Verdecchia, A.; Capocaccia, R.; Mariotto, A.; Dally, A.; Corazziari, I.; Crosignani, P.; Tagliabue, G.; Falcini, F.; Milandri, C.; Vattiato, R.; Conti, E.; Ramazzotti, V.; Caperle, M.; Zanetti, R.; Rosso, S.; Patriarca, S.; Federico, M.; Mangone, L.; Santacroce, M.; Barchielli, A.; Balzi, D.; Crocetti, E.; Paci, E.; De Lisi, V.; Serventi, L.; Barili, A.; Vercelli, M.; Casella, C.; Parodi, S.; De Leon, M. P.; Roncucci, L.; Benatti, P.; Gafà, L.; Tumino, R.; La Rosa, E.; Guzzinati, S.; Simonato, L.; Bovo, E. Cancer Prevalence in the Elderly. *Tumori J.* **1999**, *85* (5), 391–399. https://doi.org/10.1177/030089169908500506.

(4) *Cancer in 2022 - CPR22*. Cancer Progress Report. https://cancerprogressreport.aacr.org/progress/cpr22-contents/cpr22-cancer-in-2022/ (accessed 2023-07-04).

(5) Thamm, D. H. Canine Cancer: Strategies in Experimental Therapeutics. *Front. Oncol.* **2019**, *9*, 1257. https://doi.org/10.3389/fonc.2019.01257.

(6) Fleming, J. M.; Creevy, K. E.; Promislow, D. E. L. Mortality in North American Dogs from 1984 to 2004: An Investigation into Age-, Size-, and Breed-Related Causes of Death: Mortality of Dogs in North America. *J. Vet. Intern. Med.* **2011**, *25* (2), 187–198. https://doi.org/10.1111/j.1939-1676.2011.0695.x.

(7) LeBlanc, A. K.; Mazcko, C. N.; Khanna, C. Defining the Value of a Comparative Approach to Cancer Drug Development. *Clin. Cancer Res.* **2016**, *22* (9), 2133–2138. https://doi.org/10.1158/1078-0432.CCR-15-2347.

(8) *Cancer in Pets*. American Veterinary Medical Association. https://www.avma.org/resources/pet-owners/petcare/cancer-pets (accessed 2023-03-08).

(9) Wu, K.; Rodrigues, L.; Post, G.; Harvey, G.; White, M.; Miller, A.; Lambert, L.; Lewis, B.; Lopes, C.; Zou, J. Analyses of Canine Cancer Mutations and Treatment Outcomes Using Real-World Clinico-Genomics Data of 2119 Dogs. *Npj Precis. Oncol.* **2023**, *7* (1), 8. https://doi.org/10.1038/s41698-023-00346-3.

(10) Pang, L. Y.; Argyle, D. J. Using Naturally Occurring Tumours in Dogs and Cats to Study Telomerase and Cancer Stem Cell Biology. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **2009**, *1792* (4), 380–391. https://doi.org/10.1016/j.bbadis.2009.02.010.

(11) MacEwen, E. G. Spontaneous Tumors in Dogs and Cats: Models for the Study of Cancer Biology and Treatment. *CANCER METASTASIS Rev.* **1990**, *9* (2), 125–136. https://doi.org/10.1007/BF00046339.

(12) Vail, D. M.; Macewen, E. G. Spontaneously Occurring Tumors of Companion Animals as Models for Human Cancer. *Cancer Invest.* **2000**, *18* (8), 781–792. https://doi.org/10.3109/07357900009012210.

(13) Paoloni, M.; Davis, S.; Lana, S.; Withrow, S.; Sangiorgi, L.; Picci, P.; Hewitt, S.; Triche, T.; Meltzer, P.; Khanna, C. Canine Tumor Cross-Species Genomics Uncovers Targets Linked to Osteosarcoma Progression. *BMC Genomics* **2009**, *10* (1), 625. https://doi.org/10.1186/1471-2164-10-625.

(14) Mochizuki, H.; Breen, M. Comparative Aspects of BRAF Mutations in Canine Cancers. *Vet. Sci.* **2015**, *2* (3), 231–245. https://doi.org/10.3390/vetsci2030231.

(15) Sultan, F.; Ganaie, B. A. Comparative Oncology: Integrating Human and Veterinary Medicine. *Open Vet. J.* **2018**, *8* (1), 25. https://doi.org/10.4314/ovj.v8i1.5.

(16) Davis, B. W.; Ostrander, E. A. Domestic Dogs and Cancer Research: A Breed-Based Genomics Approach. *ILAR J.* **2014**, *55* (1), 59–68. https://doi.org/10.1093/ilar/ilu017.

(17) Megquier, K.; Turner-Maier, J.; Swofford, R.; Kim, J.-H.; Sarver, A. L.; Wang, C.; Sakthikumar, S.; Johnson, J.; Koltookian, M.; Lewellen, M.; Scott, M. C.; Schulte, A. J.; Borst, L.; Tonomura, N.; Alfoldi, J.; Painter, C.; Thomas, R.; Karlsson, E. K.; Breen, M.; Modiano, J. F.; Elvers, I.; Lindblad-Toh, K. Comparative Genomics Reveals Shared Mutational Landscape in Canine Hemangiosarcoma and Human Angiosarcoma. *Mol. Cancer Res.* **2019**, *17* (12), 2410–2421. https://doi.org/10.1158/1541-7786.MCR-19-0221.

(18) Gardner, H. L.; Sivaprakasam, K.; Briones, N.; Zismann, V.; Perdigones, N.; Drenner, K.; Facista, S.; Richholt, R.; Liang, W.; Aldrich, J.; Trent, J. M.; Shields, P. G.; Robinson, N.; Johnson, J.; Lana, S.; Houghton, P.; Fenger, J.; Lorch, G.; Janeway, K. A.; London, C. A.; Hendricks, W. P. D. Canine Osteosarcoma Genome Sequencing Identifies Recurrent

Mutations in DMD and the Histone Methyltransferase Gene SETD2. *Commun. Biol.* **2019**, *2* (1), 266. https://doi.org/10.1038/s42003-019-0487-2.

(19)Rodrigues, L.; Watson, J.; Feng, Y.; Lewis, B.; Harvey, G.; Post, G.; Megquier, K.; Lambert, L.; Miller, A.; Lopes, C.; Zhao, S. *Shared Hotspot Mutations in Spontaneously Arising Cancers Position Dog as an Unparalleled Comparative Model for Precision Therapeutics*; preprint; Cancer Biology, 2021. https://doi.org/10.1101/2021.10.22.465469.

(20)Rowell, J. L.; McCarthy, D. O.; Alvarez, C. E. Dog Models of Naturally Occurring Cancer. *Trends Mol. Med.* **2011**, *17* (7), 380–388. https://doi.org/10.1016/j.molmed.2011.02.004.

(21)Zhang, C.; Liu, J.; Xu, D.; Zhang, T.; Hu, W.; Feng, Z. Gain-of-Function Mutant P53 in Cancer Progression and Therapy. *J. Mol. Cell Biol.* **2020**, *12* (9), 674–687. https://doi.org/10.1093/jmcb/mjaa040.

(22)Gustafson, D. L.; Duval, D. L.; Regan, D. P.; Thamm, D. H. Canine Sarcomas as a Surrogate for the Human Disease. *Pharmacol. Ther.* **2018**, *188*, 80–96. https://doi.org/10.1016/j.pharmthera.2018.01.012.

(23)Levine, A. J. P53: 800 Million Years of Evolution and 40 Years of Discovery. *Nat. Rev. Cancer* **2020**, *20* (8), 471–480. https://doi.org/10.1038/s41568-020-0262-1.

(24)Kastenhuber, E. R.; Lowe, S. W. Putting P53 in Context. *Cell* **2017**, *170* (6), 1062–1078. https://doi.org/10.1016/j.cell.2017.08.028.

(25)Wang, H.; Guo, M.; Wei, H.; Chen, Y. Targeting P53 Pathways: Mechanisms, Structures, and Advances in Therapy. *Signal Transduct. Target. Ther.* **2023**, *8* (1), 92. https://doi.org/10.1038/s41392-023-01347-1.

(26)Linzer, D. I. H.; Levine, A. J. Characterization of a 54K Dalton Cellular SV40 Tumor Antigen Present in SV40-Transformed Cells and Uninfected Embryonal Carcinoma Cells. *Cell* **1979**, *17* (1), 43–52. https://doi.org/10.1016/0092-8674(79)90293-9.

(27)Lane, D. P.; Crawford, L. V. T Antigen Is Bound to a Host Protein in SY40-Transformed Cells. *Nature* **1979**, *278* (5701), 261–263. https://doi.org/10.1038/278261a0.

(28)Lohrum, M. A. E.; Woods, D. B.; Ludwig, R. L.; Bálint, É.; Vousden, K. H. C-Terminal Ubiquitination of P53 Contributes to Nuclear Export. *Mol. Cell. Biol.* **2001**, *21* (24), 8521–8532. https://doi.org/10.1128/MCB.21.24.8521-8532.2001.

(29)Freedman, D. A.; Wu, L.; Levine, A. J. Functions of the MDM2 Oncoprotein. *Cell. Mol. Life Sci. CMLS* **1999**, *55* (1), 96–107. https://doi.org/10.1007/s000180050273.

(30)Hu, W.; Chan, C. S.; Wu, R.; Zhang, C.; Sun, Y.; Song, J. S.; Tang, L. H.; Levine, A. J.; Feng, Z. Negative Regulation of Tumor Suppressor P53 by MicroRNA MiR-504. *Mol. Cell* **2010**, *38* (5), 689–699. https://doi.org/10.1016/j.molcel.2010.05.027.

(31)Adapted From Regulation of P53 by MDM2, 2023. https://app.biorender.com/biorender-templates.

(32)Liu, X.; Holstege, H.; Van Der Gulden, H.; Treur-Mulder, M.; Zevenhoven, J.; Velds, A.; Kerkhoven, R. M.; Van Vliet, M. H.; Wessels, L. F. A.; Peterse, J. L.; Berns, A.; Jonkers, J. Somatic Loss of BRCA1 and P53 in Mice Induces Mammary Tumors with Features of Human *BRCA1* -Mutated Basal-like Breast Cancer. *Proc. Natl. Acad. Sci.* **2007**, *104* (29), 12111–12116. https://doi.org/10.1073/pnas.0702969104.

(33)Conlin, A. The Prognostic Significance of K-Ras, P53, and APC Mutations in Colorectal Carcinoma. *Gut* **2005**, *54* (9), 1283–1286. https://doi.org/10.1136/gut.2005.066514.

(34)Chang, Y.-L.; Wu, C.-T.; Shih, J.-Y.; Lee, Y.-C. Unique P53 and Epidermal Growth Factor Receptor Gene Mutation Status in 46 Pulmonary Lymphoepithelioma-like Carcinomas. *Cancer Sci.* **2011**, *102* (1), 282–287. https://doi.org/10.1111/j.1349-7006.2010.01768.x.

(35)Hwang, S.-J.; Lozano, G.; Amos, C. I.; Strong, L. C. Germline P53 Mutations in a Cohort with Childhood Sarcoma: Sex Differences in Cancer Risk. *Am. J. Hum. Genet.* **2003**, *72* (4), 975–983. https://doi.org/10.1086/374567.

(36)Samowitz, W. S.; Curtin, K.; Ma, K.; Edwards, S.; Schaffer, D.; Leppert, M. F.; Slattery, M. L. Prognostic Significance Ofp53 Mutations in Colon Cancer at the Population Level. *Int. J. Cancer* **2002**, *99* (4), 597–602. https://doi.org/10.1002/ijc.10405.

(37)Abdelmegeed, S.; Mohammed, S. Canine Mammary Tumors as a Model for Human Disease (Review). *Oncol. Lett.* **2018**. https://doi.org/10.3892/ol.2018.8411.

(38)York, D.; Higgins, R. J.; LeCouteur, R. A.; Wolfe, A. N.; Grahn, R.; Olby, N.; Campbell, M.; Dickinson, P. J. *TP53* Mutations in Canine Brain Tumors. *Vet. Pathol.* **2012**, *49* (5), 796–801. https://doi.org/10.1177/0300985811424734.

(39)Adapted from Human Internal Organs, 2023. https://app.biorender.com/illustrations/64bdffba9c8b64619d3272d0.

(40)Joerger, A. C.; Fersht, A. R. Structural Biology of the Tumor Suppressor P53. *Annu. Rev. Biochem.* **2008**, *77* (1), 557–582. https://doi.org/10.1146/annurev.biochem.77.060806.091238.

(41)Chillemi, G.; Kehrloesser, S.; Bernassola, F.; Desideri, A.; Dötsch, V.; Levine, A. J.; Melino, G. Structural Evolution and Dynamics of the P53 Proteins. *Cold Spring Harb. Perspect. Med.* **2017**, *7* (4), a028308. https://doi.org/10.1101/cshperspect.a028308.

(42)Bell, S.; Klein, C.; Müller, L.; Hansen, S.; Buchner, J. P53 Contains Large Unstructured Regions in Its Native State. *J. Mol. Biol.* **2002**, *322* (5), 917–927. https://doi.org/10.1016/S0022-2836(02)00848-3.

(43)*Cellular tumor antigen p53 (Q29537) - protein - InterPro*. https://www.ebi.ac.uk/interpro/protein/UniProt/Q29537/ (accessed 2023-07-04).

(44)Cho, Y.; Gorina, S.; Jeffrey, P. D.; Pavletich, N. P. Crystal Structure of a P53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations. *Science* **1994**, *265* (5170), 346–355. https://doi.org/10.1126/science.8023157.

(45)De Andrade, K. C.; Khincha, P. P.; Hatton, J. N.; Frone, M. N.; Wegman-Ostrosky, T.; Mai, P. L.; Best, A. F.; Savage, S. A. Cancer Incidence, Patterns, and Genotype–Phenotype Associations in Individuals with Pathogenic or Likely Pathogenic Germline TP53 Variants: An Observational Cohort Study. *Lancet Oncol.* **2021**, *22* (12), 1787–1798. https://doi.org/10.1016/S1470-2045(21)00580-5.

(46)Harvey, M.; Vogel, H.; Morris, D.; Bradley, A.; Bernstein, A.; Donehower, L. A. A Mutant P53 Transgene Accelerates Tumour Development in Heterozygous but Not Nullizygous P53–Deficient Mice. *Nat. Genet.* **1995**, *9* (3), 305–311. https://doi.org/10.1038/ng0395-305.

(47)Duffy, M. J.; Synnott, N. C.; O'Grady, S.; Crown, J. Targeting P53 for the Treatment of Cancer. *Semin. Cancer Biol.* **2022**, *79*, 58–67. https://doi.org/10.1016/j.semcancer.2020.07.005.

(48)Alsaihati, B. A.; Ho, K.-L.; Watson, J.; Feng, Y.; Wang, T.; Dobbin, K. K.; Zhao, S. Canine Tumor Mutational Burden Is Correlated with TP53 Mutation across Tumor Types and Breeds. *Nat. Commun.* **2021**, *12* (1), 4670. https://doi.org/10.1038/s41467-021-24836-9.

(49)Liou, S.-H.; Singh, S. K.; Singer, R. H.; Coleman, R. A.; Liu, W.-L. Structure of the P53/RNA Polymerase II Assembly. *Commun. Biol.* **2021**, *4* (1), 397. https://doi.org/10.1038/s42003-021-01934-4.

(50)Seo, M.; Lei, L.; Egli, M. Label-Free Electrophoretic Mobility Shift Assay (EMSA) for Measuring Dissociation Constants of Protein-RNA Complexes. *Curr. Protoc. Nucleic Acid Chem.* **2019**, *76* (1), e70. https://doi.org/10.1002/cpnc.70.

(51)Drescher, D. G.; Selvakumar, D.; Drescher, M. J. Analysis of Protein Interactions by Surface Plasmon Resonance. In *Advances in Protein Chemistry and Structural Biology*; Elsevier, 2018; Vol. 110, pp 1–30. https://doi.org/10.1016/bs.apcsb.2017.07.003.

(52)Kitayner, M.; Rozenberg, H.; Kessler, N.; Rabinovich, D.; Shaulov, L.; Haran, T. E.; Shakked, Z. Structural Basis of DNA Recognition by P53 Tetramers. *Mol. Cell* **2006**, *22* (6), 741–753. https://doi.org/10.1016/j.molcel.2006.05.015.

(53)Wang, Y.; Zhu, X.; Wu, M.; Xia, N.; Wang, J.; Zhou, F. Simultaneous and Label-Free Determination of Wild-Type and Mutant P53 at a Single Surface Plasmon Resonance Chip Preimmobilized with Consensus DNA and Monoclonal Antibody. *Anal. Chem.* **2009**, *81* (20), 8441–8446. https://doi.org/10.1021/ac9014269.

CHAPTER 2: CYAN THERMAL PROTEIN PROJECT

Introduction

The beginning of fluorescent protein development started with Osamu Shimomura's[1] isolation of the green fluorescent protein (GFP) from thousands of jellyfish, *Aequorea victoria*.[2,3] Martin Chalfie became the first to express the protein in prokaryotic *Escherichia coli* and eukaryotic *Caenorhabditis elegans* cells.[2,4] Following that, Roger Tsien developed different types of fluorescent proteins that could be used for various applications.[2,5] In 2008, all three of them were awarded a Nobel Prize for their work in the discovery and development of GFP.[2] As GFP could form a chromophore without needing external enzymes or cofactors,[3] it opened many new possibilities for its applications.[3]

Applications of Fluorescent Proteins. There are many different applications for fluorescent proteins as they are versatile tools for providing insights into physiological processes such as gene expression, signaling, protein transport, and other regulatory events.[6] Some of the major applications for fluorescent proteins include Fluorescence/ Förster Resonance Energy Transfer (FRET),[7] protein tracking and imaging,[8] super-resolution microscopy,[9] biosensors,[10] protein conformational studies,[11] cellular environment imaging,[12] and usage as tools for genetic expression studies.[13] Applications such as fluorescent labeling and tagging for flow cytometry can include up to 30 different tags simultaneously.[14]

An example of a biosensor involving a fluorescent protein is shown in Figure 22. As shown in the figure, a green fluorescent protein from a Japanese eel is adapted through molecular cloning to fluoresce after binding to bilirubin.[10] Many such advances have been made possible by developing and using different fluorescent proteins.[15–17]

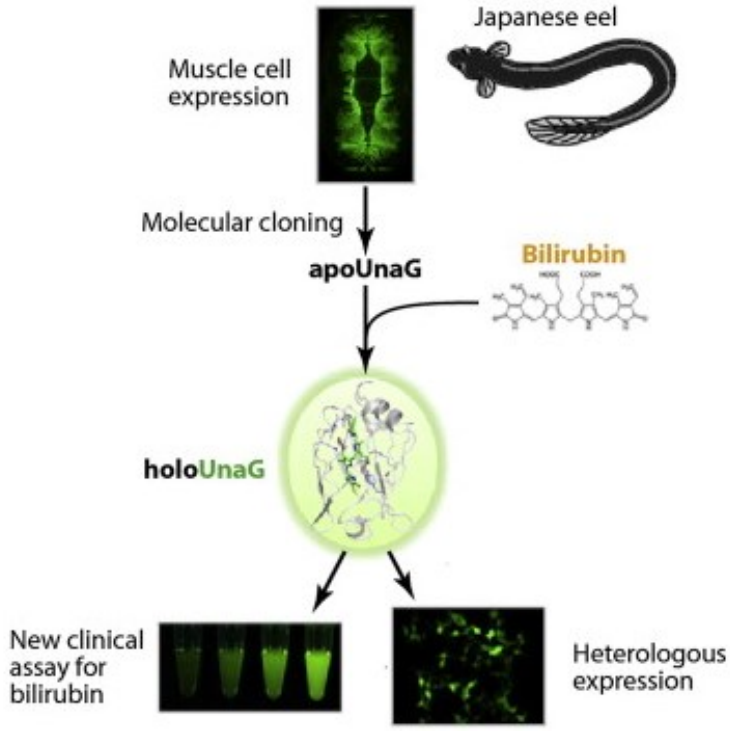

Figure 22: Example of a biosensor that tests liver function using bilirubin sensitivity. Reprinted from Cell *153* (7) Kumagai et al., A Bilirubin-Inducible Fluorescent Protein from Eel Muscle 1602–1611., (2013), with permission from Elsevier[10]

Development of Fluorescent Proteins. The original GFP had a fairly high quantum yield (QY = 0.6).[18] In the 1990s, researchers began to work on studying the different characteristics of GFP mutants in different pHs and to understand GFP's fluorescence properties better.[19] The mechanism of GFP's fluorescence involved the cyclization, dehydration, and oxidation of the Ser-Tyr-Gly residues forming a p-hydroxybenzylideneimidazolinone chromophore.[5] Researchers soon discovered that any alteration to the local environment around the chromophore, such as the positions of charged amino acid residues, hydrogen bonding networks, or hydrophobic interactions, resulted in red or blue shifts in the absorption and emission characteristics of the protein.[20] Bigger shifts in the spectral properties of fluorescent proteins are attributed to covalent structural changes as well as the π-orbital conjugation of the

chromophore.[20] For example, the presence of π-stacking interaction between Tyr66 and Tyr203 due to a T203Y mutation to GFP results in the formation of a yellow fluorescent protein (YFP). Similarly, as seen in Figure 23, a tyrosine to tryptophan (Y66W) mutation results in the conversion of EGFP (Enhanced Green Fluorescent Protein) to Cyan Fluorescent Protein (CFP).[20]

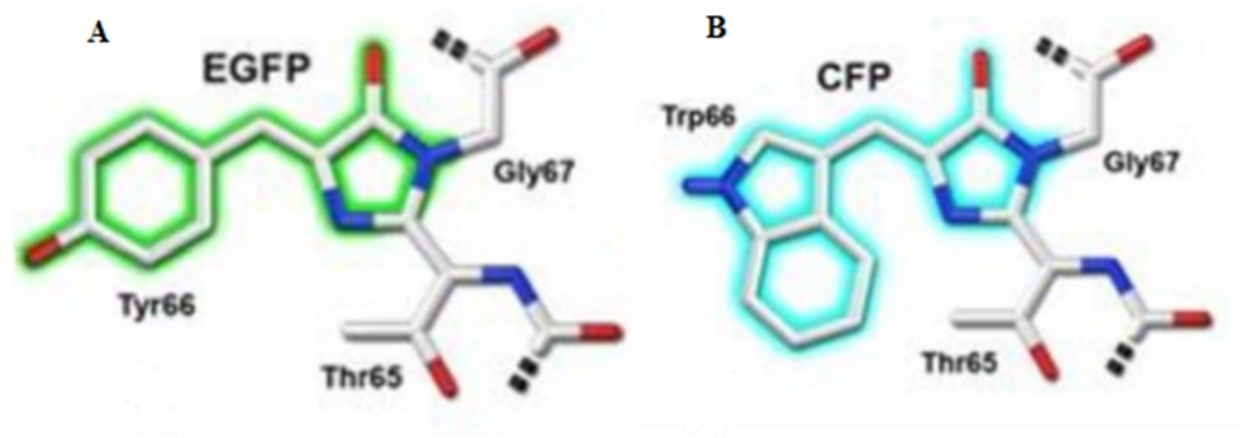

Figure 23: EGFP (A) to CFP (B) change based on Y66W mutation. Adapted with permission from Shaner et al., 2007[20]

Cyan Fluorescent Proteins (CFPs). CFPs were developed based on a tryptophan addition (Y66W) to the central residue of the chromophore.[17] The first major cyan fluorescent protein, enhanced cyan fluorescent protein (ECFP, QY= 0.36), was created using mutations that reduced the increased steric burden caused by the addition of tryptophan to the chromophore.[17,21] ECFP had alternate conformations of the seventh strand of the β-barrel,[22] which were improved by the creation of a newer variant, Cerulean (QY= 0.49), with Y145A and H148D mutations.[23] Using only the H158D mutation, another team created a better version known as Super Cyan Fluorescent Protein 3A (SCFP3A, QY= 0.56).[24] Although these proteins were significantly improved from ECFP, there were a few areas in need of improvement, such as pH stability and the loss of fluorescent capacity under illumination.[25] Some newer third-generation CFPs, such as

mCerulean2 and mCerulean3, offer an improvement in quantum yield using H148G and T65S mutations but do not completely address the flaws of the second-generation CFPs. These increases in QY show that directed evolution has ushered newer mutants with new applications.[15,26]

One of the most widely used applications of CFPs is usage as a donor in combination with YFPs during FRET.[27] FRET includes the investigation of interactions between molecules in living cells by allowing the measurement of distances and changes in conformations via the detection of energy transfer between a donor and acceptor within a fluorophore.[7] In Figure 24 below, CFP is used as a sensor of the inactive state, is excited, and emits fluorescence at a specific wavelength.[26] Similarly, in the active state, FRET occurs due to the proximity of the CFP and the YFP, and specific wavelength fluorescence is emitted.[26] This helps provide better qualitative and quantitative insights from FRET imaging compared to existing measures.[26]

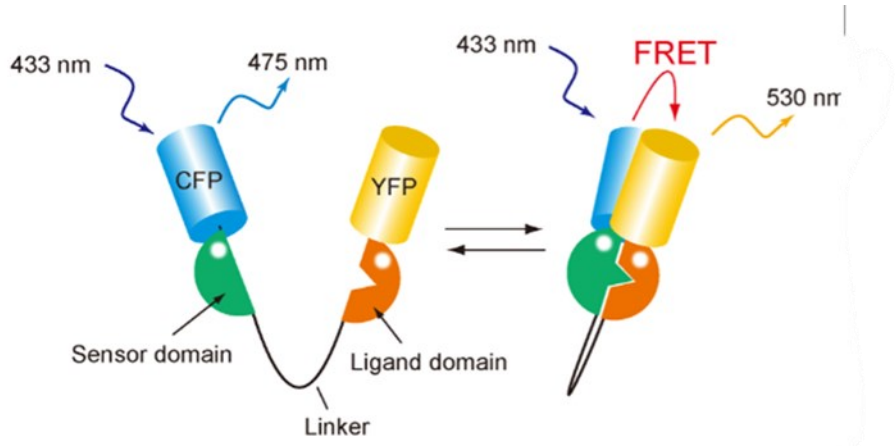

Figure 24: Dual usage of CFP and YFP as a FRET biosensor. Image adapted from Yamao et al., 2016,[26] using a creative commons attribution license.

Thermo Green Protein (TGP). Along with creating newer cyan fluorescent proteins, newer green fluorescent proteins were also being developed simultaneously. Compared to

synthetic fluorophores such as quantum dots and nanodots, most fluorescent proteins were unstable outside normal physiologic conditions, limiting their use in many applications.[16] Due to this, many efforts were made to develop variants with exceptional stability in thermophilic, denaturing conditions, such as those involved in many amyloid assays, lysosome functions, and assays involving thermophilic organisms.[16] The result was Thermo Green Protein (TGP) with exceptional thermostable and non-aggregation capabilities, derived from eCGP123, a type of consensus green protein, through directed evolution. The protein eCGP123 was derived from the synthetic consensus green protein (CGP).[28–30] TGP is an 11-stranded β-barrel protein with a QYG sequence, Gln62, Tyr63, and Gly64 residues forming the chromophore (Figure 25).[16,30] As mentioned previously, GFP has a SYG chromophore different from TGP's QYG chromophore. Additionally, TGP only shares 33% sequence identity with the original GFP, indicating they are completely different proteins.[16] The structure of TGP is shown in Figure 26 below.

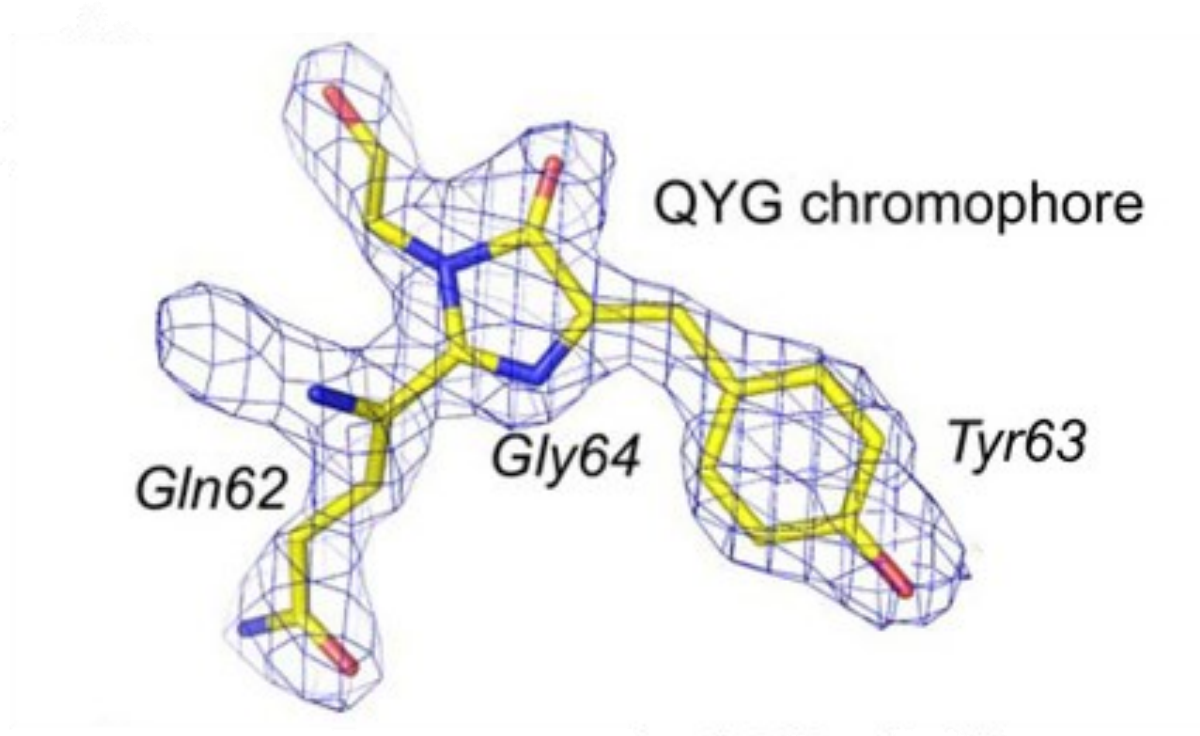

Figure 25: QYG chromophore of TGP. Adapted with permission from Close et al., 2015[16]

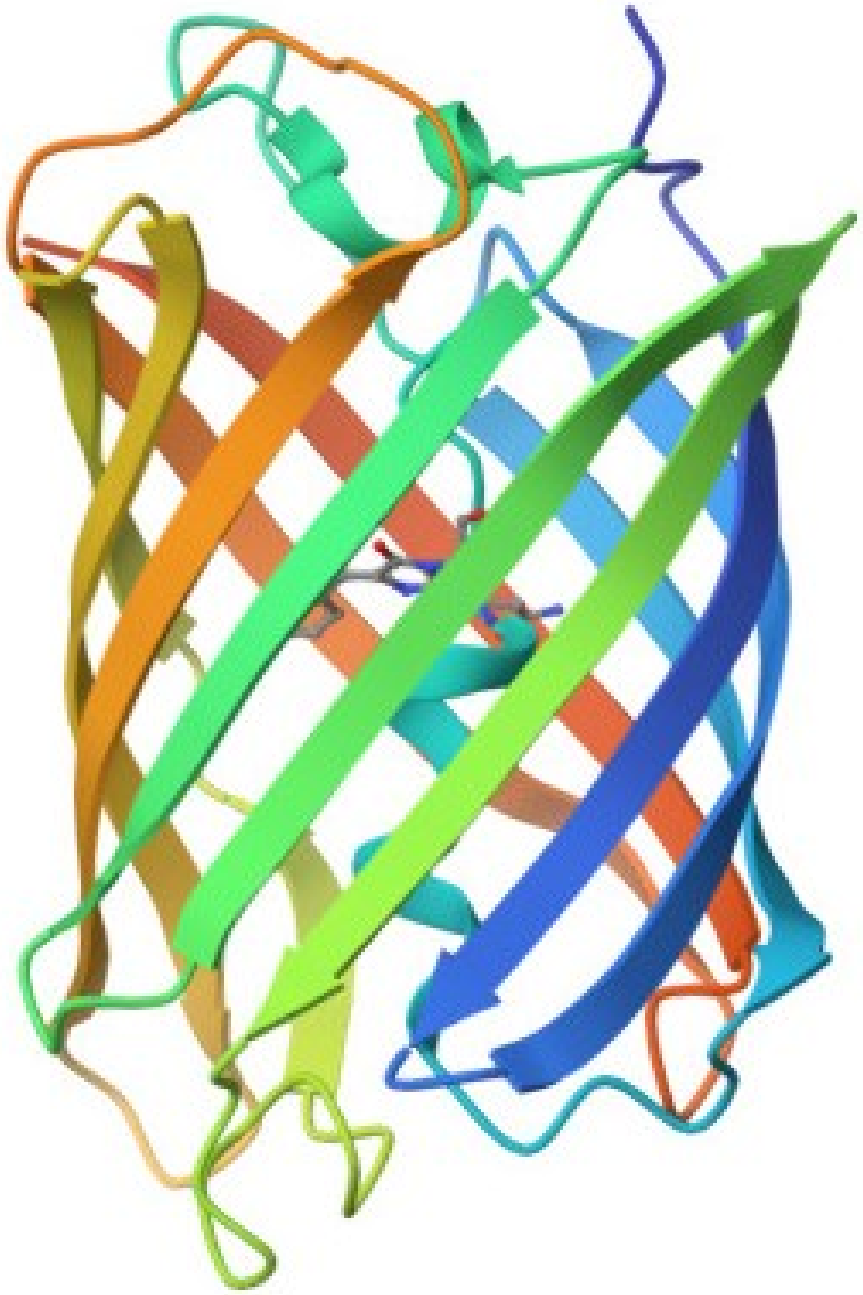

Figure 26: Structure of TGP (RCSB PDB 4TZA). Adapted from PDB, original structure by Close et al., 2015[16]

TGP is very unusual to other previously developed green fluorescent proteins because it has very high thermal stability and improvements in solubility due to negatively charged amino

acids like glutamate on the β-barrel surface.[16] Even with the improvements in these physical characteristics, it retains a very high QY, making it a beneficial fluorescent protein with wide-ranging applications from flow cytometry to amyloid assays.[16] The traditional drawbacks of CFPs are addressed by the engineering of TGP, which allows it to be used in many applications reserved previously for quantum dots and nanodots.[16]

However, there is a need for the development of more stable cyan fluorescent proteins, as many applications require the usage of a CFP, such as YFP-coupled FRET experiments and optical assays where red or green fluorescent proteins do not provide the right contrast.[17,27] Most CFPs currently share their lineage with the original ECFP and GFP. As TGP is a much more stable and efficient protein than GFP, an argument for creating CFPs with TGP as a starting point can be made. Due to the advantages provided by TGP over the original GFP, it can be reasoned that a CFP created from TGP might demonstrate a high quantum yield along with the retention of TGP's stability characteristics, which traditional CFPs lack.

Research Goals. As mentioned previously, there is a need to develop more robust versions of existing fluorescent proteins and newer fluorescent proteins that are more stable and efficient. The DeVore lab focuses on creating yellow thermal proteins (YTPs) and CTPs through rationally designed mutations to TGP. This research focused on creating a cyan fluorescent protein from TGP, named cyan thermal protein (CTP). A similar approach to creating CFP was used to create a CTP mutant by switching the tyrosine residue for a tryptophan residue (Y67W), resulting in a CTP mutant derived from TGP. CTP was fully characterized using assays to study its chemical, pH, temperature stability, and QY. More mutants were created to improve the QY of the proteins and impart better overall stability (CTP Q66E, I199T, and I199S). For GFP and some red fluorescent proteins such as DsRed, it was shown that altering the first residue of the

chromophore resulted in altered spectral characteristics and QY.[30] Altering the first residue of the chromophore of TGP, glutamine, has also shown to change the hydrogen bonding network of the protein due to the presence of an amide group in glutamine's side chain.[30] Due to these reasons, the glutamine residue of the chromophore was mutated to glutamate (Q66E), which does not contain an amide group and is a neutral amino acid to study its effect on the mutant. This mutation was also chosen as glutamate and glutamine have similar steric effects on neighboring residues due to their relatively similar size but may significantly change the surrounding hydrogen bonding network.[30] Residue 199 was mutated from isoleucine to either serine (I199S) or threonine (I199T) because, based on the structural alignment of TGP to mCerulean3, this residue should be within hydrogen bonding distance to the indole of the modified chromophore. A serine residue is found in the equivalent site of mCerulean3 and Aquamarine.[17] It was hypothesized that these mutations would provide more stability to the chromophore by providing a hydroxyl group to hydrogen bond directly to the chromophore.

CTP was successfully created from TGP but showed a very significant drop in the QY of the protein. CTP-S led to a significant drop in the quantum yield from the native CTP, but CTP-T showed increased QY. CTP-E did not improve the QY of CTP as hypothesized. The full characterization of the mutants is currently ongoing. To fully understand the interactions occurring within the chromophore and between the protein, it is important to understand the proteins' three-dimensional structure. So, hanging drop vapor diffusion experiments were used to crystallize the proteins for X-ray diffraction studies. More studies are needed to fully characterize the different mutants and solve their three-dimensional structures to understand the exact complexities of interactions between the residues. In the future, more mutants will be screened to create a better version of CTP.

Methods

Site-Directed Mutagenesis. An *Escherichia coli* (*E. coli*) codon-optimized version of the TGP gene (Q-141324) was obtained from Los Alamos National Laboratory. Appropriate forward and reverse primers were purchased from Thermo Scientific to reflect the mutants shown in Figure 27. First, CTP was created from TGP by mutating residue 67 tyrosine to tryptophan (Y67W). Then, CTP was mutated to CTP-E by mutating residue 66 to glutamic acid (Q66E), to CTP-T by mutating the isoleucine residue at position 199 to threonine (I199T), and to CTP-S by mutating the isoleucine residue at position 199 to serine (I199S). The mutations were introduced using the Agilent QuickChange II kit. The polymerase chain reaction (PCR) was conducted using 18 cycles of denaturation (95˚C for 20 seconds), annealing (60˚C for 10 seconds), and extension (68 ˚C for 540 seconds). Colonies were obtained using the transformation of *E. coli* onto LB-kanamycin plates. The colonies were used to prepare plasmids using a GeneJet plasmid miniprep kit from Thermo Scientific and quantified using a nanodrop instrument. The identity of the plasmid and verification of correct mutations was done using a third-party service from ACGT. See Appendix D-G for the full profile of cyan proteins.

Expression of Proteins. A single colony was used to start an overnight starter culture containing 100 milliliters of LB media with 50 µg/mL kanamycin at 37 ˚C. The starter culture was divided and transferred the next day to two one-liter flasks containing Lennox Broth (LB) media and grown for a few hours until the optical density at 600 nanometers (OD_{600}) reached an absorbance value of 0.8, following which protein expression was induced using 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG). The temperature was reduced to 30 ˚C and grown for one day. Following that, the culture was centrifuged at 5000 rpm for 20 minutes, and the pellets were resuspended in a resuspension buffer (100 mM Tris hydrochloride, pH 7.4, 10% glycerol, 300

mM NaCl). The resuspended cells were then lysed using 4- 30 second sonication cycles and centrifuged at 20,000 rpm for 20 minutes to separate soluble protein from cell debris. The pellets were discarded, and the supernatant (lysate) was collected for purification using different chromatography steps.

```
               10         20         30         40         50         60
TGP   MGAHASVIKP EMKIKLRMEG AVNGHKFVIE GEGIGKPYEG TQTLDLTVEE GAPLPFSYDI
CTP   .......... .......... .......... .......... .......... ..........
CTP-E .......... .......... .......... .......... .......... ..........
CTP-T .......... .......... .......... .......... .......... ..........
CTP-S .......... .......... .......... .......... .......... ..........

               70         80         90        100        110        120
TGP   LTPAFQYGNR AFTKYPEDIP DYFKQAFPEG YSWERSMTYE DQGICIATSD ITMEGDCFFY
CTP   ......W... .......... .......... .......... .......... ..........
CTP-E .....EW... .......... .......... .......... .......... ..........
CTP-T ......W... .......... .......... .......... .......... ..........
CTP-S ......W... .......... .......... .......... .......... ..........

              130        140        150        160        170        180
TGP   EIRFDGINFP PNGPVMQKKT LKWEPSTEKM YVEDGVLKGD VEMALLLEGG GHYRCDFKTT
CTP   .......... .......... .......... .......... .......... ..........
CTP-E .......... .......... .......... .......... .......... ..........
CTP-T .......... .......... .......... .......... .......... ..........
CTP-S .......... .......... .......... .......... .......... ..........

              190        200        210        220        230        240
TGP   YKAKKDVRLP DAHEVDHRIE ILSHDKDYNK VRLYEHAEAR YSGGGSGGGA SGKPIPNPLL
CTP   .......... .......... .......... .......... .......... ..........
CTP-E .......... .......... .......... .......... .......... ..........
CTP-T .......... ........T. .......... .......... .......... ..........
CTP-S .......... ........S. .......... .......... .......... ..........

              250
TGP   GLDSTHHHHH H
CTP   .......... .
CTP-E .......... .
CTP-T .......... .
CTP-S .......... .
```

Figure 27: Sequence alignment of CTP mutants to TGP. CTP-T and CTP-E were originally expressed and purified by Trey Norman and Andrew Yates.

Purification of Proteins. The lysate was loaded onto a Nickel-nitriloacetic acid (Ni-NTA, Gold biotechnology) affinity column for primary purification. The protein was washed

with a wash buffer (100 mM Tris hydrochloride, pH 7.4, 10% glycerol, 100 mM NaCl, 10 mM imidazole) and eluted using an elution buffer (100 mM Tris hydrochloride, pH 7.4, 10% glycerol, 200 mM imidazole). The elution fraction from primary purification underwent a secondary purification step using ion-exchange chromatography. A diethylaminoethyl cellulose (DEAE-C) column was used to bind the primary purified protein, which was diluted five times with ion-exchange wash buffer (50 mM Tris hydrochloride, pH 7.4, 10% glycerol) to reduce the concentration of NaCl. The protein was then eluted using an elution buffer (50 mM Tris hydrochloride, pH 7.4, 10% glycerol, and 500 mM NaCl). A list of buffers used is shown in Table 3. The purified protein was then quantified using an absorbance measurement (Shimadzu UV-2101 PC spectrophotometer) from a 700 to 250 nm range. The calculations for yield were done using Beer's law (280 nm peak), and the protein was further concentrated using Amicon ultra-4 centrifugal filter units until the final volume was reduced to less than 500 µL.

Table 3: List of buffers used for CTP purification.

Buffer (pH 7.4)	Tris HCl (mM)	Glycerol (%)	Sodium Chloride (mM)	Imidazole (mM)
Ni-NTA Resuspension	100	10	300	-
Ni-NTA Wash	100	10	100	10
Ni-NTA Elution	100	10	-	200
CM Wash	50	10	-	-
CM Elution	50	10	500	-

pH Sensitivity Assay. The pH sensitivity of the purified protein was tested using a pH sensitivity assay with five replicates. The purified protein (1 mg/mL) was diluted 15 times using a glycine/phosphate/citrate sample buffer of 0.1 M sodium chloride at different pHs, from pH 3.0 to pH 10.0. The excitation wavelength used was 430 nanometers, and the emission wavelength

used was 480 nanometers, corresponding to cyan fluorescent proteins. The fluorescence was measured after a one-hour incubation period using a SpectraMax M5 plate reader. The fluorescence data was normalized to the pH with the highest fluorescence value.

Chemical Stability Assay. A chemical stability assay was used to study the stability of the protein in the presence of a strong chemical denaturant, guanidinium hydrochloride. The purified protein was treated with increasing concentrations of guanidinium hydrochloride (0-8 M) in an assay buffer (0.1M Tris HCl pH 7.4, 20 mM $MgCl_2$). The plate was incubated at room temperature, and the fluorescence was measured at one hour, five days, and ten days with an excitation wavelength of 430 nm and an emission wavelength of 480 nm. The fluorescence was normalized against the control (0 M guanidinium HCl), and the assay was conducted in quintuplicates.

Thermal Stability Assay. The purified protein samples were diluted in the same manner as the previous assays, and the effect of temperature was studied using a thermal stability assay. The samples were heated to 60 ˚C for two hours to study unfolding. Following the unfolding experiment, the temperature was reduced to 20 ˚C to study refolding of the protein over two hours. The measurements were taken in 30-second intervals with the excitation wavelength set to 430 nm and the emission wavelength set to 480 nm.

Quantum Yield. The quantum yields of the proteins were studied using a plot of the area under fluorescence vs. absorbance of proteins with respect to a known standard, fluorescein. The proteins were diluted in an assay buffer (0.5 M Tris, pH 8.0, and 2 mM $MgCl_2$). Sequential amounts of protein were added, and an absorbance scan (500-350 nm, slow scan speed, 0.5 nm scan interval, 2.0 nm slit width, Shimadzu UV – 2101 PC), and a fluorescence scan (425- 600 nm, excitation 410 nm, Perkin Elmer LS 55 fluorescence spectrometer) were taken at each

addition. The quantum yield of the proteins was calculated using the equation represented below.[31]

$$Q_X = Q_{ST}\left(\frac{m_x}{m_{ST}}\right)\left(\frac{n_x^2}{n_{ST}^2}\right)$$

The subscripts ST and X represent the standard and the sample, respectively. M represents the slope of the area under fluorescence vs. absorbance plot, n represents the refractive indexes (1.33 literature value for solvents used),[31] and Q is the quantum yield.

Crystallography. The purified protein was then used to set up plates of crystals via the hanging drop vapor diffusion method. One microliter of protein, generally around 20 mg/mL, was combined with the same amount of well solution on a coverslip. The coverslip is then used to seal the well using either immersion oil or grease. Over 100 crystal screen conditions were set up using commercially available third-party screens such as Hampton Research Crytal Screen I and II. The plates were allowed to grow in an undisturbed environment for weeks/months and periodically checked for crystal growth. Once the crystals were large enough, they were looped and affixed to a Rigaku Xtalab Synergy-S pedestal and flash-frozen at 100 K. Preliminary data was collected to check for diffracting crystals. Unfortunately, none of the crystals diffracted sufficiently to collect a full dataset.

Results

Mutants. A total of four different cyan mutants were studied during this research. First, TGP was converted to CTP using a Y67W mutation. Following the successful creation of this protein, three additional mutations were induced in CTP to create CTP- S (I199S), CTP-E (Q66E), and CTP-T (I199T). CTP was fully characterized and compared with TGP, and quantum yield data was collected on the other mutants to understand the improvements or degradation in

the efficiency of the chromophore. The CTP-T and CTP-E mutants were expressed and purified by Trey Norman and Andrew Yates. The overall goal was to develop a protein with a significantly greater quantum yield than the native CTP protein.

Expression and Purification. A total of ten different rounds of expression and purification were conducted for CTP and CTP-S mutants. The initial few rounds of expression were done using protocols that resulted in successful expression and purification of other fluorescent proteins. It was discovered that CTP and CTP-S favored a two liter-one day growth and a growth temperature of 26-28 ˚C for better yield. The best expression yielded roughly 22 mg/L of protein, and the worst resulted in less than 10 mg/L. The Ni-NTA and DEAE purification column for the best CTP prep are shown in Figures 28 and 29. For the Ni-NTA purification, the red box represents the wash fraction, and the green box represents the elution fraction. For the DEAE column, the elution fraction is denoted by the green box, which shows a very sharp peak, denoting a high amount of eluted protein from the column.

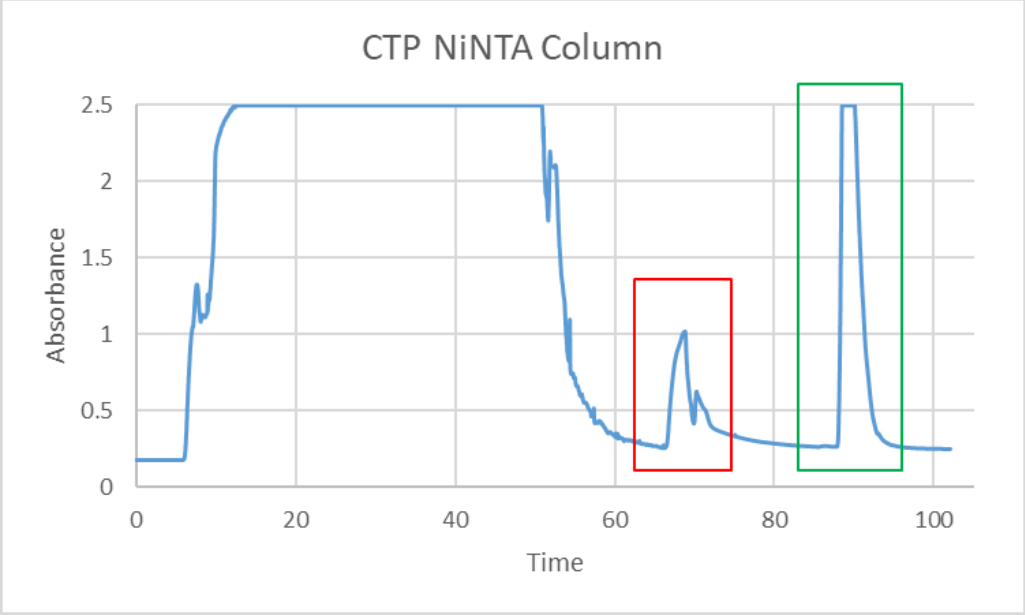

Figure 28: CTP purification using a Ni-NTA column. The red and green boxes denote the wash and elution fractions, respectively.

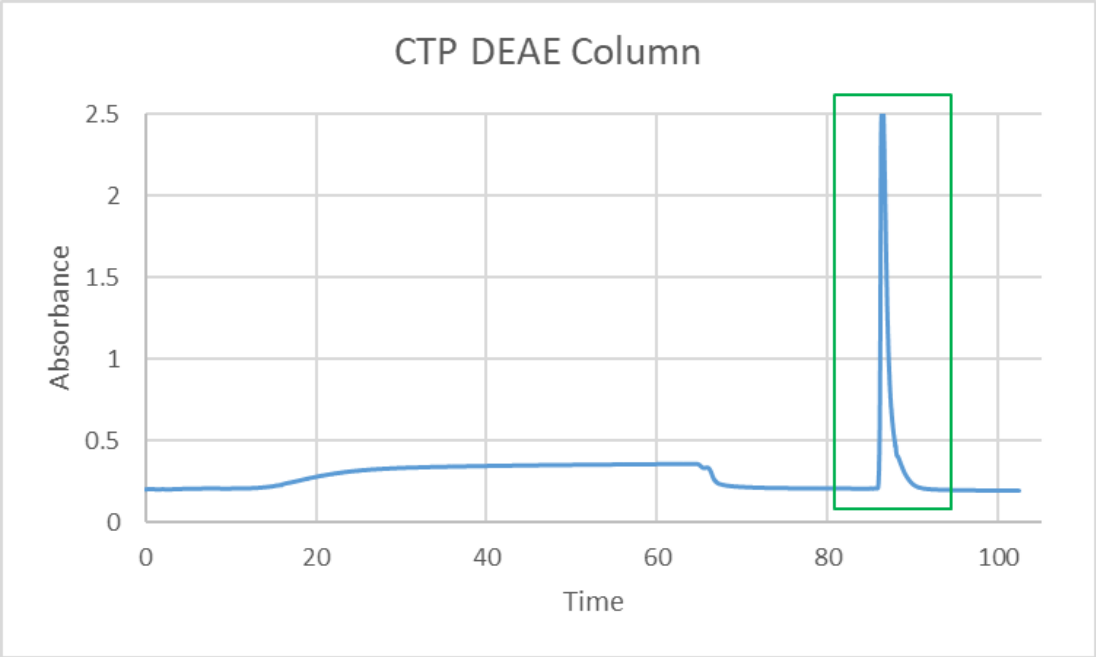

Figure 29: CTP purification using a DEAE-C column. The green box denotes the elution fraction.

Full Characterization of CTP. The complete characterization of CTP was conducted using different assays to test its robustness in different environments. The data collected were used to compare different proteins using literature.

Excitation and Emission Scans. The excitation and emission scans of CTP are shown in Figure 30 below. The λ_{max} values for absorption and emission scans of CTP were compared to the data previously obtained by our lab for TGP.[30] The λ_{max} of absorbance and emission for TGP was 494 nm and 509 nm, respectively.[30] The λ_{max} values of absorbance and emission for the newly developed CTP were 423 nm and 476 nm, respectively, reflecting a successful mutation and blue shift of TGP to cyan. The absorption and emission values for CTP are consistent with other cyan fluorescent proteins.

pH Stability Assay. A pH study assay was conducted to study the stability of CTP at different pH values ranging from pH 3 to pH 9 (Figure 31). Based on the normalized

fluorescence data below, it can be seen that CTP is most stable at pH 8.0. Below pH 8.0, there is a gradual decline to 80% at pH 6.0, followed by a sharper decline to 30% at pH 3.0. The fluorescence at pH 9.0 and 10.0 (not shown) declines to 50% showing that CTP is unsuitable for basic environments.

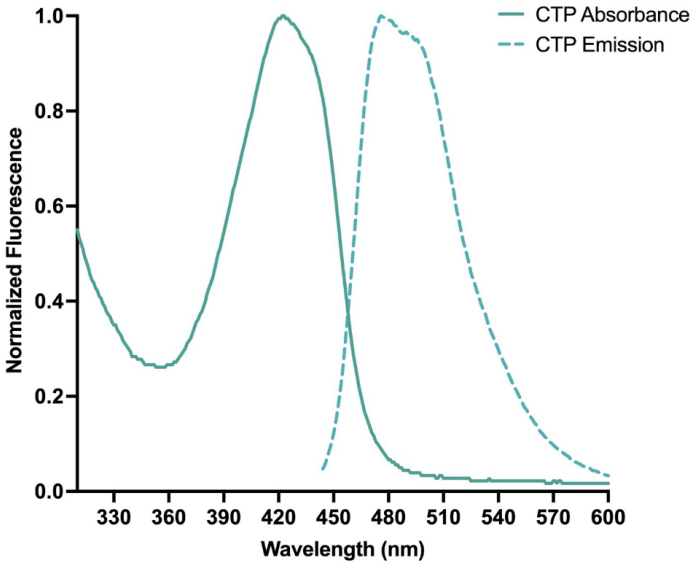

Figure 30: Absorbance and emission scans for CTP.

Chemical Stability Assay. A chemical denaturation study was conducted using Guanidinium Hydrochloride (Gdn HCl) to test the effect of a strong chemical denaturant on protein stability. Gdn HCl concentrations from 0 M to 8 M were used to conduct this experiment, and the fluorescence values were measured after one hour, five days, and ten days. The fluorescence values were normalized to the control samples with the highest fluorescence at the one-hour time point. From Figure 32, it can be noted that there is a roughly 20% drop in the fluorescence values of the 0M samples on day 5 and day 10. Both day 5 and day 10 samples show almost identical trends, meaning that there is not a substantial loss of fluorescence from day 5 to day 10 between samples at the same concentrations of Gdn HCl.

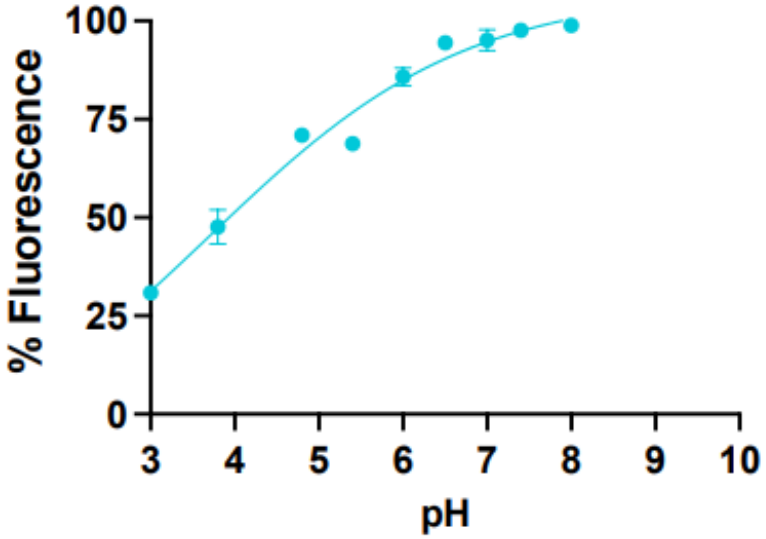

Figure 31: pH stability assay for CTP.

Interestingly, there is no meaningful drop in fluorescence values at Gdn HCl concentrations greater than 4M. All samples greater than 4M Gdn HCl show a normalized fluorescence value of 20% on day 5 and day 10. As fluorescence values of proteins are directly correlated to their folding, it could mean that CTP retains some level of folding even after 10 days in 8 M Gdn HCl.

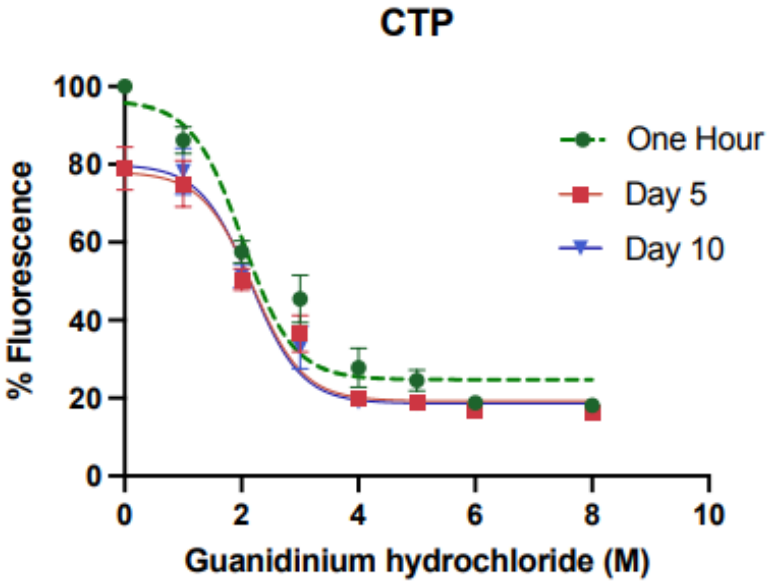

Figure 32: Chemical stability of CTP.

Thermal Stability Assay. A temperature stability assay was devised to understand the effect of temperature on the stability of CTP (Figure 33). As the temperature ramped up to 60 ˚C, the fluorescence of CTP decreased and stabilized to roughly 28% of the original fluorescence. Due to instrumental limitations, it was not possible to study the thermal degradation of CTP at higher temperatures; the normalized fluorescence of CTP would be expected to decrease further at temperatures higher than 60 ˚C.

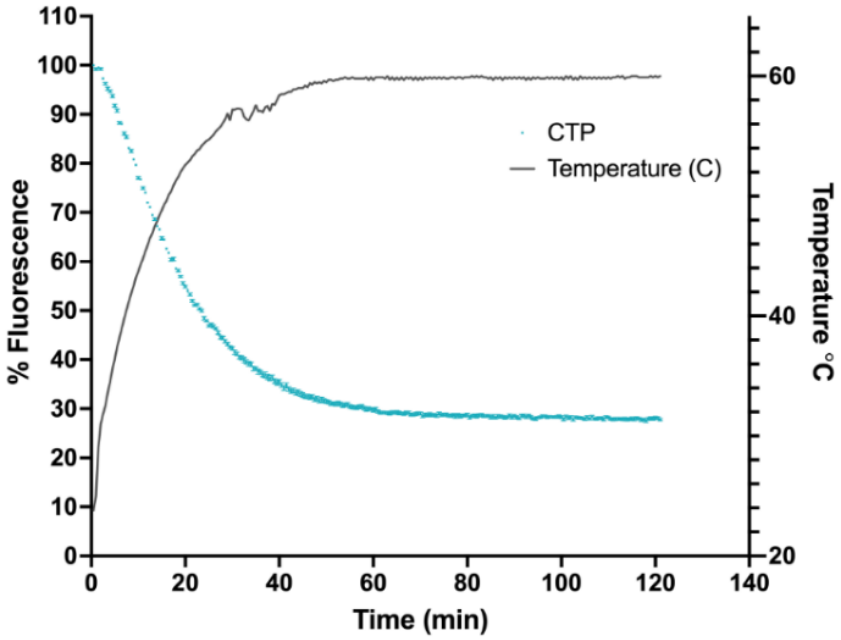

Figure 33: Temperature study assay for CTP.

After a heat cycle, a refolding experiment was conducted to see the effect of cooling to the original temperature. The temperature was brought down to room temperature for two hours, and fluorescence measurements were taken every 30 seconds. From Figure 34 below, it can be noted that CTP refolded to 90% of its original value after a refolding cycle. TGP only refolds back to 30% of its original value after one cycle of unfolding.[30] This shows that CTP can refold better to its original state than TGP.

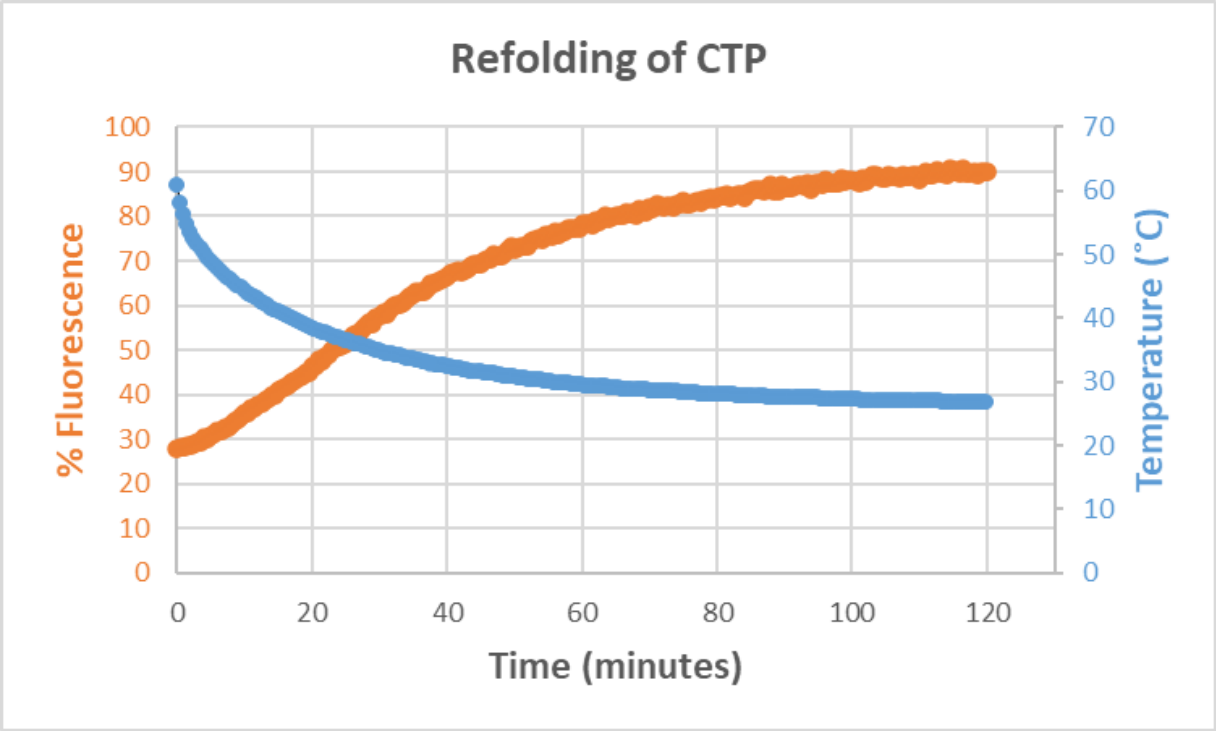

Figure 34: Refolding of CTP.

Fluorescence Quantum Yields. The quantum yield of proteins was obtained using the slope of the area under the fluorescence vs. absorbance plot. The quantum yield values of CTP mutants were calculated using quantum yield values of a known standard, fluorescein. The quantum yield of CTP was determined to be 0.0581 (data not shown, the experiment was conducted by Caitlin Padgett). Although CTP shows a significant decrease in the QY from TGP (QY= 0.83)[30], the QY of YTP, another protein created by the DeVore lab, was 0.02,[30] much worse than CTP. All quantum yields were calculated using the equation below.[31]

$$Q_X = Q_{ST}\left(\frac{m_x}{m_{ST}}\right)\left(\frac{n_x^2}{n_{ST}^2}\right)$$

Quantum Yield of CTP-S. The quantum yield value for CTP-S was the worst among all CTP mutants. The quantum yield value for CTP-S was 0.0185 (Figure 35), which is only 31.6% of the quantum yield value for CTP.

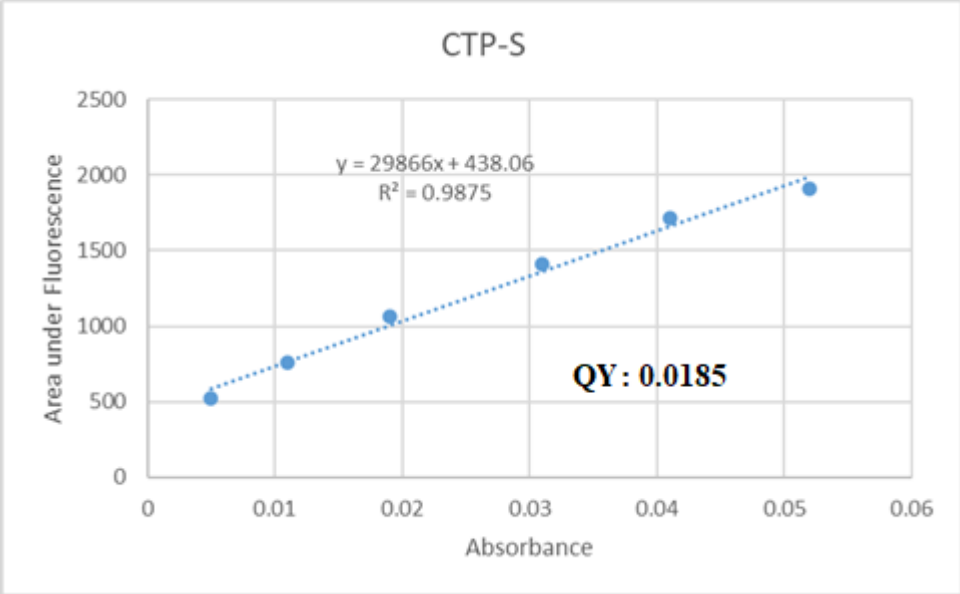

Figure 35: Area under fluorescence vs. absorbance plot for CTP-S. The quantum yield was 0.0184.

Quantum Yield of CTP-E. The quantum yield of CTP-E at pH 8.0 was 0.0303 (Figure 36-left), and it was 0.0406 at pH 7.4 (Figure 36 -right). Initially, all mutants were studied at pH 8.0 as native CTP is most stable at pH 8.0, but preliminary pH characterization revealed that CTP-E was most stable at pH 7.4. Going from pH 8.0 to pH 7.4, the QY increased by roughly 34%.

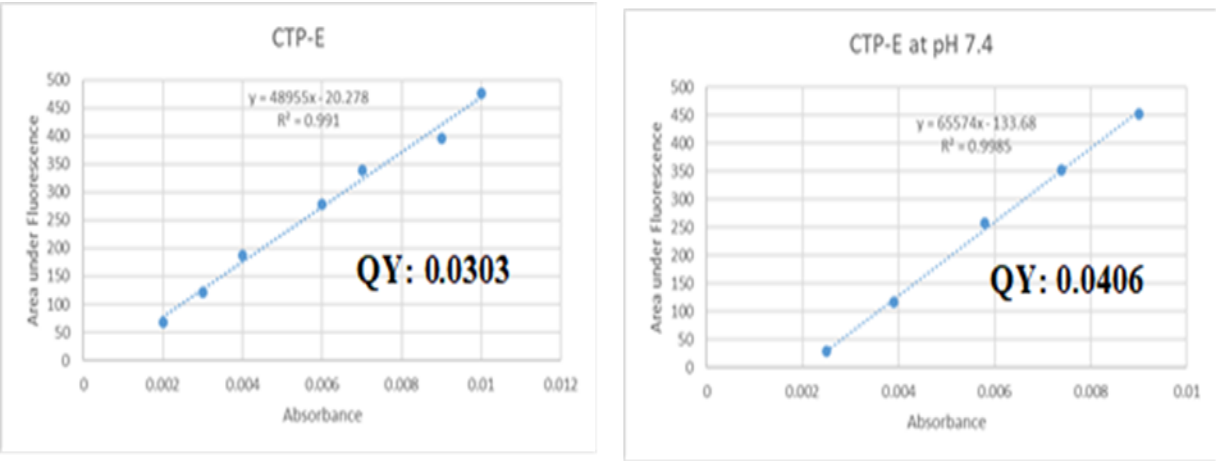

Figure 36: Area under fluorescence vs. absorbance plot for CTP-E. The quantum yield was 0.0303 at pH 8.0 (left) and 0.0406 at pH 7.4 (right). Reduction in pH shows an increase in the QY value of CTP-E

Quantum Yield of CTP-T. The quantum yield of CTP-T at pH 8.0 was 0.0394 (Figure 35-left), and it was 0.0737 at pH 6.5 (Figure 37 -right). Initially, all mutants were studied at pH 8.0 as native CTP is most stable at pH 8.0, but preliminary pH characterization revealed that CTP-T was most stable at pH 6.5. Going from pH 8.0 to pH 6.5, the QY increased by roughly 87%.

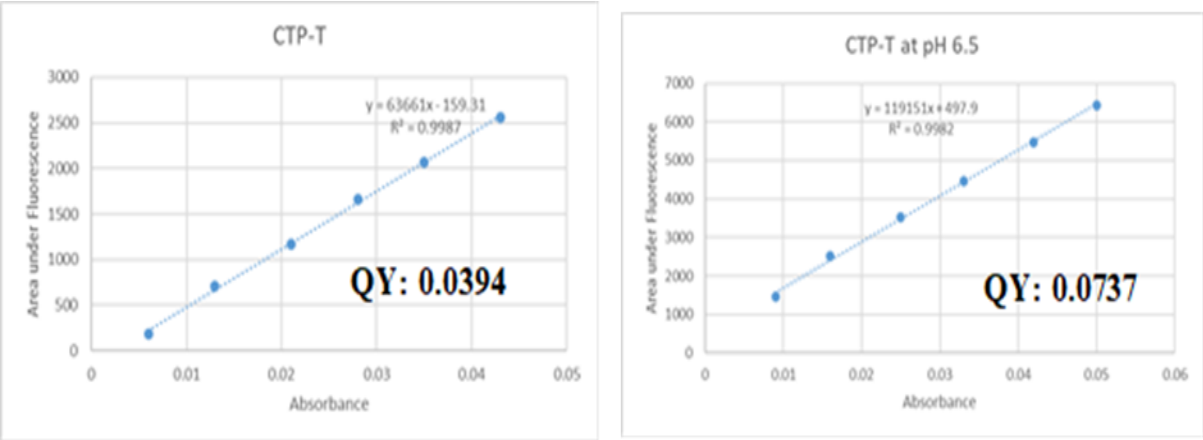

Figure 37: Area under fluorescence vs. absorbance plot for CTP-T. The quantum yield was 0.0394 at pH 8.0 (left) and 0.0737 at pH 6.5 (right). Reduction in pH shows an increase in the QY value of CTP-E

Crystallography. After purification and concentration, the proteins were screened using Hampton Research Crystal Screen I and II to check for crystal hits (Appendix B and C). The conditions that led to crystals were optimized to create better quality crystals. CTP and CTP-S yielded crystals at a concentration of 20 mg/mL. Over 200 unique conditions were used to promote crystal growth, and the best crystals were used for X-ray diffraction (XRD) experiments. Some of the conditions that yielded data sets for other fluorescent proteins were also used to increase the odds of collecting a good dataset. The best conditions for crystal growth included 0.35 M Magnesium chloride hexahydrate, 0.1 M Tris hydrochloride pH 8.5, and 32.5% polyethylene glycol 8,000 in the well solution. Some of the best crystals for CTP and CTP-S

took nearly six months to grow, leading to difficulty in optimizing. Over fifteen crystals were screened and placed on the XRD instrument for data collection.

Unfortunately, none of the crystals diffracted well enough to collect a full dataset and solve the structure of the proteins. The efforts to crystallize and optimize more conditions are ongoing. An example of a looped cyan crystal is shown in Figure 38 below.

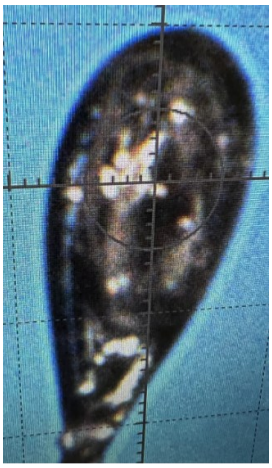

Figure 38: Looped CTP-S crystal. Well condition: 0.35 M Magnesium chloride hexahydrate, 0.1 M Tris hydrochloride pH 8.5, and 32.5% polyethylene glycol 8,000.

Discussion

Four different cyan mutants were studied during this research. CTP-T and CTP-E were expressed and purified by Trey Norman and Andrew Yates. TGP was first converted into CTP using a Y67W mutation. CTP was then converted to CTP-S, CTP-E, and CTP-T using I199S, Q66E, and I199T mutations, respectively. The proteins were expressed and purified using a BL-21(DE3) *E. coli* cell line.

The native CTP showed a spectral shift to 423 nm for absorption maxima, and 476 nm for emission maxima, showing a successful shift to cyan. The pH stability of CTP was

significantly greater than the reported stability of TGP[30] by Anderson et al. at all pHs, except for pH 9.0 and pH 10.0, indicating that CTP may be more suitable for acidic environments. In terms of chemical stability and denaturation by guanidinium hydrochloride, CTP showed poor stability at all concentrations compared to the chemical stability data reported for TGP by Anderson et al.[30] TGP maintained nearly 90% of its fluorescence even at 3M guanidinium hydrochloride concentrations, indicating the retention of most of its fluorescent properties even in harsh environments.[30] Additionally, TGP also showed more thermal stability at all temperatures compared to CTP; however, CTP showed better % recovery after the first refolding cycle as it regained around 90% of its fluorescence compared to only 30% recovery for TGP.[30]

The QY of other CFPs is significantly greater than all CTP mutants showing the need for more optimization. Regarding stability, ECFP (QY = 0.36)[17] showed complete denaturation at a temperature of 75.7 ˚C.[32] Due to instrumental limitations, the CTP thermal stability experiments were only conducted at a temperature of 60 ˚C, where CTP retained 30% of its fluorescence, leading to difficulty in determining the point of complete denaturation. Regarding pH stability, CTP shows greater % fluorescence in acidic conditions than ECFP, which showed a complete loss of fluorescence intensity around pH 4.0 or lower.[33] CTP retained roughly 50% of its normalized fluorescence intensity at pH 4.0.

Unfortunately, CTP also showed a significant loss of QY compared to TGP, which shows that although the Y67W mutation leads to the creation of CTP, it may also cause an increased steric burden due to the bulkier side chain of tryptophan. Similar reductions in QY were observed during the creation of ECFP (QY = 0.36),[17] indicating the need to stabilize the increased steric burden of tryptophan to improve the QY. Another reason behind the reduction in QY could be the alteration of π-stacking interactions with histidine 193 due to the loss of the

tyrosine residue within the chromophore. Although tryptophan can also participate in π-stacking interactions with histidine 193, the bulkier side chain of tryptophan may disturb the chromophore enough to cause a decrease in QY.

The addition of Q66E mutation to both TGP and YTP has increased the thermal stability of both proteins.[30] The full characterization of the cyan mutants is still ongoing; therefore, it is difficult to predict if the same is true for CTP. The structural determination of TGP-E showed that the stability increase was due to an additional hydrogen bond with the backbone, leading to a shorter bond length of 2.8 Å, compared to 3.0 Å in native TGP.[34] If the same is true for CTP-E, there may be an increase in thermal stability compared to CTP. Due to the CTP-E mutation leading to no improvement in QY, even with increased thermal stability, such a protein would need to be optimized further.

The addition of I199S and I199T mutations led to very different results. Although the idea behind adding each mutation was to directly hydrogen bond to the chromophore, the I199T mutation increased the QY, whereas the I199S led to a major decrease. As threonine is more sterically similar to isoleucine, this might explain why serine, which is also able to form hydrogen bonds, is unable to stabilize the networks, resulting in a lower quantum yield, whereas the threonine mutation results in an overall better quantum yield. Table 4 includes a summary of QYs for CTP mutants.

Table 4: Summary of Quantum Yields.

Protein	Quantum Yield
CTP (pH 8.0) (Caitlin Padgett)	0.0581
CTP-S (pH 8.0)	0.0185
CTP-E (pH 8.0)	0.0303
CTP-E (pH 7.4)	0.0406
CTP-T (pH 8.0)	0.0394
CTP-T (pH 6.5)	0.0737

The overall quantum yield for all cyan mutants is still quantitatively low compared to other CFPs, and more optimization is needed to improve the QY of CTP mutants. Solving the three-dimensional structures of the mutants would help us understand how each mutation disrupts or improves the hydrogen bonding and π-stacking networks. Without solved structures of the CTP mutants, it is challenging to accurately determine the ongoing interactions within the chromophore, as a cyan version of TGP has never been created, and the sequences of other CFPs are too different for direct comparison of interactions. A better understanding of the underlying networks would help create better mutants, and more optimization is needed to increase the chances of getting a diffracting crystal. For hanging drop experiments, the conditions that have led to crystal formation need to be repeated with small changes in well solution concentrations to improve the quality of crystals for XRD experiments. The full characterization of other cyan mutants is ongoing.

Lastly, CTP-T showed the best QY out of all the mutants, and full characterization of all mutants, especially CTP-T, is needed to understand why it has a higher QY and what can be done to improve the QY of the next generation of cyan mutants. Despite many failures and the inability to get enough diffraction data for structure solution, this work laid the foundation for future experiments on different mutants.

References

(1) Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* **1962**, *59* (3), 223–239. https://doi.org/10.1002/jcp.1030590302.

(2) Zimmer, M. GFP: From Jellyfish to the Nobel Prize and Beyond. *Chem. Soc. Rev.* **2009**, *38* (10), 2823–2832. https://doi.org/10.1039/b904023d.

(3) Bartkiewicz, M.; Kazazić, S.; Krasowska, J.; Clark, P. L.; Wielgus-Kutrowska, B.; Bzowska, A. Non–Fluorescent Mutant of Green Fluorescent Protein Sheds Light on the Mechanism of Chromophore Formation. *FEBS Lett.* **2018**, *592* (9), 1516–1523. https://doi.org/10.1002/1873-3468.13051.

(4) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green Fluorescent Protein as a Marker for Gene Expression. *Science* **1994**, *263* (5148), 802–805. https://doi.org/10.1126/science.8303295.

(5) Tsien, R. Y. The Green Fluorescent Protein. *Annu. Rev. Biochem.* **1998**, *67*, 509–544. https://doi.org/10.1146/annurev.biochem.67.1.509.

(6) Wu, C.; Bull, B.; Szymanski, C.; Christensen, K.; McNeill, J. Multicolor Conjugated Polymer Dots for Biological Fluorescence Imaging. *ACS Nano* **2008**, *2* (11), 2415–2423. https://doi.org/10.1021/nn800590n.

(7) Xia, Z.; Liu, Y. Reliable and Global Measurement of Fluorescence Resonance Energy Transfer Using Fluorescence Microscopes. *Biophys. J.* **2001**, *81* (4), 2395–2402. https://doi.org/10.1016/S0006-3495(01)75886-9.

(8) Kubala, M. H.; Kovtun, O.; Alexandrov, K.; Collins, B. M. Structural and Thermodynamic Analysis of the GFP:GFP-Nanobody Complex. *Protein Sci.* **2010**, *19* (12), 2389–2401. https://doi.org/10.1002/pro.519.

(9) Brakemann, T.; Stiel, A. C.; Weber, G.; Andresen, M.; Testa, I.; Grotjohann, T.; Leutenegger, M.; Plessmann, U.; Urlaub, H.; Eggeling, C.; Wahl, M. C.; Hell, S. W.; Jakobs, S. A Reversibly Photoswitchable GFP-like Protein with Fluorescence Excitation Decoupled from Switching. *Nat. Biotechnol.* **2011**, *29* (10), 942–947. https://doi.org/10.1038/nbt.1952.

(10) Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. A Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* **2013**, *153* (7), 1602–1611. https://doi.org/10.1016/j.cell.2013.05.038.

(11) Ghisaidoobe, A.; Chung, S. Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* **2014**, *15* (12), 22518–22538. https://doi.org/10.3390/ijms151222518.

(12) Suhling, K.; Siegel, J.; Phillips, D.; French, P. M. W.; Lévêque-Fort, S.; Webb, S. E. D.; Davis, D. M. Imaging the Environment of Green Fluorescent Protein. *Biophys. J.* **2002**, *83* (6), 3589–3595. https://doi.org/10.1016/S0006-3495(02)75359-9.

(13) Bose, J. L.; Fey, P. D.; Bayles, K. W. Genetic Tools To Enhance the Study of Gene Function and Regulation in Staphylococcus Aureus. *Appl. Environ. Microbiol.* **2013**, *79* (7), 2218–2224. https://doi.org/10.1128/AEM.00136-13.

(14) McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **2018**, *120*, 5.1.1-5.1.11. https://doi.org/10.1002/cpim.40.

(15) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from Discosoma Sp. Red Fluorescent Protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567–1572. https://doi.org/10.1038/nbt1037.

(16) Close, D. W.; Paul, C. D.; Langan, P. S.; Wilce, M. C. J.; Traore, D. A. K.; Halfmann, R.; Rocha, R. C.; Waldo, G. S.; Payne, R. J.; Rucker, J. B.; Prescott, M.; Bradbury, A. R. M. Thermal Green Protein, an Extremely Stable, Nonaggregating Fluorescent Protein Created by Structure-Guided Surface Engineering. *Proteins Struct. Funct. Bioinforma.* **2015**, *83* (7), 1225–1237. https://doi.org/10.1002/prot.24699.

(17) Gotthard, G.; Von Stetten, D.; Clavel, D.; Noirclerc-Savoye, M.; Royant, A. Chromophore Isomer Stabilization Is Critical to the Efficient Fluorescence of Cyan Fluorescent Proteins. *Biochemistry* **2017**, *56* (49), 6418–6422. https://doi.org/10.1021/acs.biochem.7b01088.

(18) Pirovano, G.; Roberts, S.; Kossatz, S.; Reiner, T. Optical Imaging Modalities: Principles and Applications in Preclinical Research and Clinical Settings. *J. Nucl. Med.* **2020**, *61* (10), 1419–1427. https://doi.org/10.2967/jnumed.119.238279.

(19) Haupts, U.; Maiti, S.; Schwille, P.; Webb, W. W. Dynamics of Fluorescence Fluctuations in Green Fluorescent Protein Observed by Fluorescence Correlation Spectroscopy. *Proc. Natl. Acad. Sci.* **1998**, *95* (23), 13573–13578. https://doi.org/10.1073/pnas.95.23.13573.

(20) Shaner, N. C.; Patterson, G. H.; Davidson, M. W. Advances in Fluorescent Protein Technology. *J. Cell Sci.* **2007**, *120* (24), 4247–4260. https://doi.org/10.1242/jcs.005801.

(21) Cubitt, A. B.; Woollenweber, L. A.; Heim, R. Chapter 2: Understanding Structure—Function Relationships in the Aequorea Victoria Green Fluorescent Protein. In *Methods in Cell Biology*; Elsevier, 1998; Vol. 58, pp 19–30. https://doi.org/10.1016/S0091-679X(08)61946-9.

(22) Hyun Bae, J.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H. J.; Azim, M. K.; Kim, J.-S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N. Expansion of the Genetic Code Enables Design of a Novel "Gold" Class of Green Fluorescent Proteins. *J. Mol. Biol.* **2003**, *328* (5), 1071–1081. https://doi.org/10.1016/S0022-2836(03)00364-4.

(23) Rizzo, M. A.; Springer, G. H.; Granada, B.; Piston, D. W. An Improved Cyan Fluorescent Protein Variant Useful for FRET. *Nat. Biotechnol.* **2004**, *22* (4), 445–449. https://doi.org/10.1038/nbt945.

(24) Kremers, G.-J.; Goedhart, J.; Van Munster, E. B.; Gadella, T. W. J. Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius ˚. *Biochemistry* **2006**, *45* (21), 6570–6580. https://doi.org/10.1021/bi0516273.

(25) Sinnecker, D.; Voigt, P.; Hellwig, N.; Schaefer, M. Reversible Photobleaching of Enhanced Green Fluorescent Proteins. *Biochemistry* **2005**, *44* (18), 7085–7094. https://doi.org/10.1021/bi047881x.

(26) Yamao, M.; Aoki, K.; Yukinawa, N.; Ishii, S.; Matsuda, M.; Naoki, H. Two New FRET Imaging Measures: Linearly Proportional to and Highly Contrasting the Fraction of Active Molecules. *PLOS ONE* **2016**, *11* (10), e0164254. https://doi.org/10.1371/journal.pone.0164254.

(27) Goedhart, J.; Von Stetten, D.; Noirclerc-Savoye, M.; Lelimousin, M.; Joosen, L.; Hink, M. A.; Van Weeren, L.; Gadella, T. W. J.; Royant, A. Structure-Guided Evolution of Cyan Fluorescent Proteins towards a Quantum Yield of 93%. *Nat. Commun.* **2012**, *3* (1), 751. https://doi.org/10.1038/ncomms1738.

(28) Dai, M.; Fisher, H. E.; Temirov, J.; Kiss, C.; Phipps, M. E.; Pavlik, P.; Werner, J. H.; Bradbury, A. R. M. The Creation of a Novel Fluorescent Protein by Guided Consensus Engineering. *Protein Eng. Des. Sel.* **2007**, *20* (2), 69–79. https://doi.org/10.1093/protein/gzl056.

(29) Kiss, C.; Temirov, J.; Chasteen, L.; Waldo, G. S.; Bradbury, A. R. M. Directed Evolution of an Extremely Stable Fluorescent Protein. *Protein Eng. Des. Sel.* **2009**, *22* (5), 313–323. https://doi.org/10.1093/protein/gzp006.

(30) Anderson, M. R.; Padgett, C. M.; Dargatz, C. J.; Nichols, C. R.; Vittalam, K. R.; DeVore, N. M. Engineering a Yellow Thermostable Fluorescent Protein by Rational Design. *ACS Omega* **2023**, *8* (1), 436–443. https://doi.org/10.1021/acsomega.2c05005.

(31) Yvon, Jobin. A Guide to Recording Fluorescence Quantum Yields. https://www.chem.uci.edu/~dmitryf/manuals/Fundamentals/Quantum%20yield%20determination.pdf.

(32) Lelimousin, M.; Noirclerc-Savoye, M.; Lazareno-Saez, C.; Paetzold, B.; Le Vot, S.; Chazal, R.; Macheboeuf, P.; Field, M. J.; Bourgeois, D.; Royant, A. Intrinsic Dynamics in ECFP and Cerulean Control Fluorescence Quantum Yield. *Biochemistry* **2009**, *48* (42), 10038–10046. https://doi.org/10.1021/bi901093w.

(33) Fredj, A.; Pasquier, H.; Demachy, I.; Jonasson, G.; Levy, B.; Derrien, V.; Bousmah, Y.; Manoussaris, G.; Wien, F.; Ridard, J.; Erard, M.; Merola, F. The Single T65S Mutation Generates Brighter Cyan Fluorescent Proteins with Increased Photostability and PH Insensitivity. *PLoS ONE* **2012**, *7* (11), e49149. https://doi.org/10.1371/journal.pone.0049149.

(34) Anderson, M. R. Structural Engineering of Thermostable Fluorescent Proteins TGP-E and YTP-E and Crystal Structure of TGP-E. *MSU Grad. Theses* **2023**, *3828*.

SUMMARY

This section will attempt to summarize both projects. Between the two projects, eighteen rounds of purification and expression were conducted (10 for CTP, 8 for canine p53). In both projects, proteins were successfully expressed and purified for further analysis.

The native CTP was fully characterized using quantum yield (QY) data, temperature stability assay, pH stability assay, and chemical stability assay. The QY data for all cyan mutants were collected and compared. For the p53 protein project, nine EMSA runs and nine SPR runs were conducted to test the binding affinity of canine protein to human DNA. The EMSA showed a K_d value of 13.876 μM for hairpin loop #1 and 12.54 μM for hairpin loop #2, whereas the SPR runs could not yield fruitful results.

For the crystallization of proteins between both projects, a total of over 200 unique conditions per project were tested for crystal growth. Between the two projects, there were over 1000 small protein drops in hanging drop vapor diffusion wells for crystal growth. Hundreds of crystals grew between the projects, but most were too small to screen for X-ray diffraction. In total, roughly 40 crystals (around 25 for p53 and 15 for the CTP project) were screened using X-ray diffraction experiments, but unfortunately, none of them diffracted sufficiently for a structure solution.

Several advancements were made between both projects that will be beneficial going forward. Future experiments can build on the progress made by these projects, and the hope is that careful optimization could lead to solved protein structures for both projects.

APPENDICES

Appendix A: p53 Protein Profile

Profile created using ProtParam feature by Expasy.

```
Number of amino acids: 207

Molecular weight: 2:   1.67

Theoretical pI: 8.99

Amino acid composition:
Ala (A)   7       3.4%
Arg (R)  16       7.7%
Asn (N)  10       4.8%
Asp (D)   7       3.4%
Cys (C)   9       4.3%
Gln (Q)   3       1.4%
Glu (E)  11       5.3%
Gly (G)  12       5.8%
His (H)  13       6.3%
Ile (I)   6       2.9%
Leu (L)  15       7.2%
Lys (K)   9       4.3%
Met (M)   4       1.9%
Phe (F)   7       3.4%
Pro (P)  17       8.2%
Ser (S)  20       9.7%
Thr (T)  14       6.8%
Trp (W)   2       1.0%
Tyr (Y)   9       4.3%
Val (V)  16       7.7%
Pyl (O)   0       0.0%
Sec (U)   0       0.0%

 (B)   0          0.0%
 (Z)   0          0.0%
 (X)   0          0.0%


Total number of negatively charged residues (Asp + Glu): 18
Total number of positively charged residues (Arg + Lys): 25
```

Atomic composition:

Carbon	C	1031
Hydrogen	H	1597
Nitrogen	N	305
Oxygen	O	300
Sulfur	S	13

Formula: $C_{1031}H_{1597}N_{305}O_{300}S_{13}$
Total number of atoms: 3246

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient	24910
Abs 0.1% (=1 g/l)	1.061, assuming all pairs of Cys residues form cystines

Ext. coefficient	24410
Abs 0.1% (=1 g/l)	1.040, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is L (Leu).

The estimated half-life is: 5.5 hours (mammalian reticulocytes, in vitro).
3 min (yeast, in vivo).
2 min (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 65.25
This classifies the protein as unstable.

Aliphatic index: 65.36

Grand average of hydropathicity (GRAVY): -0.556

Appendix B: Hampton Research Crystal Screen I

Tube #	Salt	Buffer ◊	Precipitant
1.	0.02 M Calcium chloride dihydrate	0.1 M Sodium acetate trihydrate pH 4.6	30% v/v (+/-)-2-Methyl-2,4-pentanediol
2.	None	None	0.4 M Potassium sodium tartrate tetrahydrate
3.	None	None	0.4 M Ammonium phosphate monobasic
4.	None	0.1 M TRIS hydrochloride pH 8.5	2.0 M Ammonium sulfate
5.	0.2 M Sodium citrate tribasic dihydrate	0.1 M HEPES sodium pH 7.5	30% v/v (+/-)-2-Methyl-2,4-pentanediol
6.	0.2 M Magnesium chloride hexahydrate	0.1 M TRIS hydrochloride pH 8.5	30% w/v Polyethylene glycol 4,000
7.	None	0.1 M Sodium cacodylate trihydrate pH 6.5	1.4 M Sodium acetate trihydrate
8.	0.2 M Sodium citrate tribasic dihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% v/v 2-Propanol
9.	0.2 M Ammonium acetate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	30% w/v Polyethylene glycol 4,000
10.	0.2 M Ammonium acetate	0.1 M Sodium acetate trihydrate pH 4.6	30% w/v Polyethylene glycol 4,000
11.	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	1.0 M Ammonium phosphate monobasic
12.	0.2 M Magnesium chloride hexahydrate	0.1 M HEPES sodium pH 7.5	30% v/v 2-Propanol
13.	0.2 M Sodium citrate tribasic dihydrate	0.1 M TRIS hydrochloride pH 8.5	30% v/v Polyethylene glycol 400
14.	0.2 M Calcium chloride dihydrate	0.1 M HEPES sodium pH 7.5	28% v/v Polyethylene glycol 400
15.	0.2 M Ammonium sulfate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% w/v Polyethylene glycol 8,000
16.	None	0.1 M HEPES sodium pH 7.5	1.5 M Lithium sulfate monohydrate
17.	0.2 M Lithium sulfate monohydrate	0.1 M TRIS hydrochloride pH 8.5	30% w/v Polyethylene glycol 4,000
18.	0.2 M Magnesium acetate tetrahydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	20% w/v Polyethylene glycol 8,000
19.	0.2 M Ammonium acetate	0.1 M TRIS hydrochloride pH 8.5	30% v/v 2-Propanol
20.	0.2 M Ammonium sulfate	0.1 M Sodium acetate trihydrate pH 4.6	25% w/v Polyethylene glycol 4,000
21.	0.2 M Magnesium acetate tetrahydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% v/v (+/-)-2-Methyl-2,4-pentanediol
22.	0.2 M Sodium acetate trihydrate	0.1 M TRIS hydrochloride pH 8.5	30% w/v Polyethylene glycol 4,000
23.	0.2 M Magnesium chloride hexahydrate	0.1 M HEPES sodium pH 7.5	30% v/v Polyethylene glycol 400
24.	0.2 M Calcium chloride dihydrate	0.1 M Sodium acetate trihydrate pH 4.6	20% v/v 2-Propanol
25.	None	0.1 M Imidazole pH 6.5	1.0 M Sodium acetate trihydrate
26.	0.2 M Ammonium acetate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	30% v/v (+/-)-2-Methyl-2,4-pentanediol
27.	0.2 M Sodium citrate tribasic dihydrate	0.1 M HEPES sodium pH 7.5	20% v/v 2-Propanol
28.	0.2 M Sodium acetate trihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% w/v Polyethylene glycol 8,000
29.	None	0.1 M HEPES sodium pH 7.5	0.8 M Potassium sodium tartrate tetrahydrate
30.	0.2 M Ammonium sulfate	None	30% w/v Polyethylene glycol 8,000
31.	0.2 M Ammonium sulfate	None	30% w/v Polyethylene glycol 4,000
32.	None	None	2.0 M Ammonium sulfate
33.	None	None	4.0 M Sodium formate
34.	None	0.1 M Sodium acetate trihydrate pH 4.6	2.0 M Sodium formate
35.	None	0.1 M HEPES sodium pH 7.5	0.8 M Sodium phosphate monobasic monohydrate 0.8 M Potassium phosphate monobasic
36.	None	0.1 M TRIS hydrochloride pH 8.5	8% w/v Polyethylene glycol 8,000
37.	None	0.1 M Sodium acetate trihydrate pH 4.6	8% w/v Polyethylene glycol 4,000
38.	None	0.1 M HEPES sodium pH 7.5	1.4 M Sodium citrate tribasic dihydrate
39.	None	0.1 M HEPES sodium pH 7.5	2% v/v Polyethylene glycol 400 2.0 M Ammonium sulfate
40.	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	20% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
41.	None	0.1 M HEPES sodium pH 7.5	10% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
42.	0.05 M Potassium phosphate monobasic	None	20% w/v Polyethylene glycol 8,000
43.	None	None	30% w/v Polyethylene glycol 1,500
44.	None	None	0.2 M Magnesium formate dihydrate
45.	0.2 M Zinc acetate dihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	18% w/v Polyethylene glycol 8,000
46.	0.2 M Calcium acetate hydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	18% w/v Polyethylene glycol 8,000
47.	None	0.1 M Sodium acetate trihydrate pH 4.6	2.0 M Ammonium sulfate
48.	None	0.1 M TRIS hydrochloride pH 8.5	2.0 M Ammonium phosphate monobasic
49.	1.0 M Lithium sulfate monohydrate	None	2% w/v Polyethylene glycol 8,000
50.	0.5 M Lithium sulfate monohydrate	None	15% w/v Polyethylene glycol 8,000

Appendix C: Hampton Research Crystal Screen II

Tube #	Salt	Tube #	Buffer ◊	Tube #	Precipitant
1.	2.0 M Sodium chloride	1.	None	1.	10% w/v Polyethylene glycol 6,000
2.	0.5 M Sodium chloride 0.01 M Magnesium chloride hexahydrate	2.	None	2.	0.01 M Hexadecyltrimethylammonium bromide
3.	None	3.	None	3.	25% v/v Ethylene glycol
4.	None	4.	None	4.	35% v/v 1,4-Dioxane
5.	2.0 M Ammonium sulfate	5.	None	5.	5% v/v 2-Propanol
6.	None	6.	None	6.	1.0 M Imidazole pH 7.0
7.	None	7.	None	7.	10% w/v Polyethylene glycol 1,000 10% w/v Polyethylene glycol 8,000
8.	1.5 M Sodium chloride	8.	None	8.	10% v/v Ethanol
9.	None	9.	0.1 M Sodium acetate trihydrate pH 4.6	9.	2.0 M Sodium chloride
10.	0.2 M Sodium chloride	10.	0.1 M Sodium acetate trihydrate pH 4.6	10.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
11.	0.01 M Cobalt(II) chloride hexahydrate	11.	0.1 M Sodium acetate trihydrate pH 4.6	11.	1.0 M 1,6-Hexanediol
12.	0.1 M Cadmium chloride hydrate	12.	0.1 M Sodium acetate trihydrate pH 4.6	12.	30% v/v Polyethylene glycol 400
13.	0.2 M Ammonium sulfate	13.	0.1 M Sodium acetate trihydrate pH 4.6	13.	30% w/v Polyethylene glycol monomethyl ether 2,000
14.	0.2 M Potassium sodium tartrate tetrahydrate	14.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	14.	2.0 M Ammonium sulfate
15.	0.5 M Ammonium sulfate	15.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	15.	1.0 M Lithium sulfate monohydrate
16.	0.5 M Sodium chloride	16.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	16.	2% v/v Ethylene imine polymer
17.	None	17.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	17.	35% v/v tert-Butanol
18.	0.01 M Iron(III) chloride hexahydrate	18.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	18.	10% v/v Jeffamine® M-600®
19.	None	19.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	19.	2.5 M 1,6-Hexanediol
20.	None	20.	0.1 M MES monohydrate pH 6.5	20.	1.6 M Magnesium sulfate heptahydrate
21.	0.1 M Sodium phosphate monobasic monohydrate 0.1 M Potassium phosphate monobasic	21.	0.1 M MES monohydrate pH 6.5	21.	2.0 M Sodium chloride
22.	None	22.	0.1 M MES monohydrate pH 6.5	22.	12% w/v Polyethylene glycol 20,000
23.	1.6 M Ammonium sulfate	23.	0.1 M MES monohydrate pH 6.5	23.	10% v/v 1,4-Dioxane
24.	0.05 M Cesium chloride	24.	0.1 M MES monohydrate pH 6.5	24.	30% v/v Jeffamine® M-600®
25.	0.01 M Cobalt(II) chloride hexahydrate	25.	0.1 M MES monohydrate pH 6.5	25.	1.8 M Ammonium sulfate
26.	0.2 M Ammonium sulfate	26.	0.1 M MES monohydrate pH 6.5	26.	30% w/v Polyethylene glycol monomethyl ether 5,000
27.	0.01 M Zinc sulfate heptahydrate	27.	0.1 M MES monohydrate pH 6.5	27.	25% v/v Polyethylene glycol monomethyl ether 550
28.	None	28.	None	28.	1.6 M Sodium citrate tribasic dihydrate pH 6.5
29.	0.5 M Ammonium sulfate	29.	0.1 M HEPES pH 7.5	29.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
30.	None	30.	0.1 M HEPES pH 7.5	30.	10% w/v Polyethylene glycol 6,000 5% v/v (+/-)-2-Methyl-2,4-pentanediol
31.	None	31.	0.1 M HEPES pH 7.5	31.	20% v/v Jeffamine® M-600®
32.	0.1 M Sodium chloride	32.	0.1 M HEPES pH 7.5	32.	1.6 M Ammonium sulfate
33.	None	33.	0.1 M HEPES pH 7.5	33.	2.0 M Ammonium formate
34.	0.05 M Cadmium sulfate hydrate	34.	0.1 M HEPES pH 7.5	34.	1.0 M Sodium acetate trihydrate
35.	None	35.	0.1 M HEPES pH 7.5	35.	70% v/v (+/-)-2-Methyl-2,4-pentanediol
36.	None	36.	0.1 M HEPES pH 7.5	36.	4.3 M Sodium chloride
37.	None	37.	0.1 M HEPES pH 7.5	37.	10% w/v Polyethylene glycol 8,000 8% v/v Ethylene glycol
38.	None	38.	0.1 M HEPES pH 7.5	38.	20% w/v Polyethylene glycol 10,000
39.	0.2 M Magnesium chloride hexahydrate	39.	0.1 M Tris pH 8.5	39.	3.4 M 1,6-Hexanediol
40.	None	40.	0.1 M Tris pH 8.5	40.	25% v/v tert-Butanol
41.	0.01 M Nickel(II) chloride hexahydrate	41.	0.1 M Tris pH 8.5	41.	1.0 M Lithium sulfate monohydrate
42.	1.5 M Ammonium sulfate	42.	0.1 M Tris pH 8.5	42.	12% v/v Glycerol
43.	0.2 M Ammonium phosphate monobasic	43.	0.1 M Tris pH 8.5	43.	50% v/v (+/-)-2-Methyl-2,4-pentanediol
44.	None	44.	0.1 M Tris pH 8.5	44.	20% v/v Ethanol
45.	0.01 M Nickel(II) chloride hexahydrate	45.	0.1 M Tris pH 8.5	45.	20% w/v Polyethylene glycol monomethyl ether 2,000
46.	0.1 M Sodium chloride	46.	0.1 M BICINE pH 9.0	46.	20% v/v Polyethylene glycol monomethyl ether 550
47.	None	47.	0.1 M BICINE pH 9.0	47.	2.0 M Magnesium chloride hexahydrate
48.	None	48.	0.1 M BICINE pH 9.0	48.	2% v/v 1,4-Dioxane 10% w/v Polyethylene glycol 20,000

Appendix D: CTP Profile

Profile created using ProtParam feature by Expasy

Number of amino acids: 251

Molecular weight: 28256.78

Theoretical pI: 5.61

Amino acid composition:

Ala (A)	14	5.6%
Arg (R)	9	3.6%
Asn (N)	6	2.4%
Asp (D)	17	6.8%
Cys (C)	3	1.2%
Gln (Q)	5	2.0%
Glu (E)	22	8.8%
Gly (G)	27	10.8%
His (H)	13	5.2%
Ile (I)	13	5.2%
Leu (L)	16	6.4%
Lys (K)	19	7.6%
Met (M)	8	3.2%
Phe (F)	11	4.4%
Pro (P)	16	6.4%
Ser (S)	11	4.4%
Thr (T)	14	5.6%
Trp (W)	3	1.2%
Tyr (Y)	13	5.2%
Val (V)	11	4.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 39
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1265
Hydrogen	H	1911
Nitrogen	N	337
Oxygen	O	379
Sulfur	S	11

Formula: $C_{1265}H_{1911}N_{337}O_{379}S_{11}$
Total number of atoms: 3903

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 35995
Abs 0.1% (=1 g/l) 1.274, assuming all pairs of Cys residues form cystines

Ext. coefficient 35870
Abs 0.1% (=1 g/l) 1.269, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 30.94
This classifies the protein as stable.

Aliphatic index: 63.35

Grand average of hydropathicity (GRAVY): -0.644

Appendix E: CTP-E Profile

Profile created using ProtParam feature by Expasy

Number of amino acids: 251

Molecular weight: 28257.77

Theoretical pI: 5.52

Amino acid composition:

Ala (A)	14	5.6%
Arg (R)	9	3.6%
Asn (N)	6	2.4%
Asp (D)	17	6.8%
Cys (C)	3	1.2%
Gln (Q)	4	1.6%
Glu (E)	23	9.2%
Gly (G)	27	10.8%
His (H)	13	5.2%
Ile (I)	13	5.2%
Leu (L)	16	6.4%
Lys (K)	19	7.6%
Met (M)	8	3.2%
Phe (F)	11	4.4%
Pro (P)	16	6.4%
Ser (S)	11	4.4%
Thr (T)	14	5.6%
Trp (W)	3	1.2%
Tyr (Y)	13	5.2%
Val (V)	11	4.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 40
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1265
Hydrogen	H	1910
Nitrogen	N	336
Oxygen	O	380
Sulfur	S	11

Formula: $C_{1265}H_{1910}N_{336}O_{380}S_{11}$
Total number of atoms: 3902

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 35995
Abs 0.1% (=1 g/l) 1.274, assuming all pairs of Cys residues form cystines

Ext. coefficient 35870
Abs 0.1% (=1 g/l) 1.269, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 30.34
This classifies the protein as stable.

Aliphatic index: 63.35

Grand average of hydropathicity (GRAVY): -0.644

Appendix F: CTP-T Profile

Profile created using ProtParam feature by Expasy

Number of amino acids: 251

Molecular weight: 28244.73

Theoretical pI: 5.61

Amino acid composition:

Ala (A)	14	5.6%
Arg (R)	9	3.6%
Asn (N)	6	2.4%
Asp (D)	17	6.8%
Cys (C)	3	1.2%
Gln (Q)	5	2.0%
Glu (E)	22	8.8%
Gly (G)	27	10.8%
His (H)	13	5.2%
Ile (I)	12	4.8%
Leu (L)	16	6.4%
Lys (K)	19	7.6%
Met (M)	8	3.2%
Phe (F)	11	4.4%
Pro (P)	16	6.4%
Ser (S)	11	4.4%
Thr (T)	15	6.0%
Trp (W)	3	1.2%
Tyr (Y)	13	5.2%
Val (V)	11	4.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 39
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1263
Hydrogen	H	1907
Nitrogen	N	337
Oxygen	O	380
Sulfur	S	11

Formula: $C_{1263}H_{1907}N_{337}O_{380}S_{11}$
Total number of atoms: 3898

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 35995
Abs 0.1% (=1 g/l) 1.274, assuming all pairs of Cys residues form cystines

Ext. coefficient 35870
Abs 0.1% (=1 g/l) 1.270, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 29.95
This classifies the protein as stable.

Aliphatic index: 61.79

Grand average of hydropathicity (GRAVY): -0.665

Appendix G: CTP-S Profile

Profile created using ProtParam feature by Expasy

Number of amino acids: 251

Molecular weight: 28230.70

Theoretical pI: 5.61

Amino acid composition:

Ala (A)	14	5.6%
Arg (R)	9	3.6%
Asn (N)	6	2.4%
Asp (D)	17	6.8%
Cys (C)	3	1.2%
Gln (Q)	5	2.0%
Glu (E)	22	8.8%
Gly (G)	27	10.8%
His (H)	13	5.2%
Ile (I)	12	4.8%
Leu (L)	16	6.4%
Lys (K)	19	7.6%
Met (M)	8	3.2%
Phe (F)	11	4.4%
Pro (P)	16	6.4%
Ser (S)	12	4.8%
Thr (T)	14	5.6%
Trp (W)	3	1.2%
Tyr (Y)	13	5.2%
Val (V)	11	4.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 39
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1262
Hydrogen	H	1905
Nitrogen	N	337
Oxygen	O	380
Sulfur	S	11

Formula: $C_{1262}H_{1905}N_{337}O_{380}S_{11}$
Total number of atoms: 3895

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 35995
Abs 0.1% (=1 g/l) 1.275, assuming all pairs of Cys residues form cystines

Ext. coefficient 35870
Abs 0.1% (=1 g/l) 1.271, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 31.70
This classifies the protein as stable.

Aliphatic index: 61.79

Grand average of hydropathicity (GRAVY): -0.665