COMPUTATIONAL DESIGN OF PROTEIN PATHWAY INHIBITORS

by

WAVERLY WILSON

A THESIS

Presented to the Department of Chemistry and Biochemistry
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2025

An Abstract of the Thesis of

Waverly Wilson for the degree of Bachelor of Science
in the Department of Chemistry and Biochemistry to be taken June 2025.

Title: Computational Design of Protein Pathway Inhibitors

Approved: *Dr. Parisa Hosseinzadeh*
Primary Thesis Advisor

Protein-Protein Interaction (PPI) networks are vital to how the cell functions. These networks are composed of individual proteins with specific binding partners that transmit signals in a cascading manner. When mutations occur to constituent proteins of the network, it creates a disruption of the network's function, which can cause disease states. This concept is extremely significant in the p53 interaction network. When disease states occur in the p53 network, it is a leading cause of cancer. A deeper understanding of how the PPI networks shift in the presence of disruptions is needed in order to understand how disease states can arise from disruptions. To explore the effects of these disruptions, protein inhibitors are computationally designed for known participant proteins in the p53 pathway. This thesis will explore the creation of a computationally designed inhibitor to S100B, a key protein in melanoma, along with the testing of the computationally designed inhibitors to three chosen proteins in the p53 pathway.

Acknowledgements

I wish to thank the PH lab for all their support during this thesis process. The lab has been super amazing and supportive, and I am grateful to be doing research in this lab. I want to thank Parisa Hosseinzadeh for guiding and supporting my research and being a great advisor. I want to thank Elizabeth Raisanen for all her support in the honors college and her help through the thesis process. I want to thank my mentor, Cassandra Gonzalez, for her insight and guidance through my research and her help with this project. I want to thank Brooklyn Cessna for being a great mentee and her energy in the lab. I want to thank my roommates, Alex Smith, Emma Lindquist, and Jessani Blakely for supporting me through this process and always hyping me up and showing interest in my work. I want to thank Marcus Polk for being so supportive and an amazing partner. I want to thank my family for their support and being amazing. I wish to thank my mom for being such an inspiration and supporting me so much. I wish to thank my dad for always pushing me and visiting me in Eugene. I wish to thank Whitney for always calling me and reminding me why I am doing this. I wish to thank Westerly for always being ready to call me out and also being very supportive and caring. I wish to thank Teagan Hunt for her support and celebratory boba runs. I wish to thank Luca Lippert for being the best lab buddy and giving such good advice and being an amazing friend. I wish to thank Henry Hutley for being a great study buddy and friend and going through the depths of chem with me. I wish to thank every single person in my life for supporting me and showing up. I wish to thank the Clark Honors College for supporting me through the Knight Campus Undergraduate Scholarship.

Table of Contents

Introduction 7
Protein-Protein Interactions 7
Background on Proteins 9
p53 9
S100B 10
CBP 11
SIRT1 11
Interaction of S100B and P53 12
Computational Protein Design 13
Methods 15
Initial Computational Methods 15
Initial Experimental Methods 17
Optimized Experimental Method 18
Testing Binders 19
Biolayer Interferometry 19
Fluorescence Polarization 20
S100B Computational Design Optimization 22
Results and Analysis 27
Discussion 42
Future Directions 43
Conclusion 44
Bibliography 55
Supplementary Information 45

List of Figures

Figure 1: P53 Pathway 10

Figure 2: Interaction of Target Proteins with CTD of P53 14

Figure 3: S100B with Highlights P53 Binding Hotspots 15

Figure 4: Protein Design Pipeline 19

Figure 5: Testing the new wet lab method in lab 20

Figure 6: Wet Lab Workflow 22

Figure 7: Fluorescence Polarization 24

Figure 8: Initial Design Metrics 26

Figure 9: Optimization of Potentials 27

Figure 10: Best Potential Combinations 28

Figure 11: Denoiser Runs 29

Figure 12: S100B plasmid insert compared against pet28a(+) and s100b sequence 30

Figure 13: SIRT-1 plasmid insert compared against pet28a(+) and SIRT-1 sequence. 31

Figure 14: CBP plasmid insert compared to pet28a(+) and CBP sequence 32

Figure 15: Purified S100B 33

Figure 16: SIRT-1 Purification 34

Figure 17: Purified CBP 35

Figure 18: CBP Loading Test 36

Figure 19: A1 Binder Tested with CBP 37

Figure 20: A2 and A3 binder tested with CBP 38

Figure 21: A4 and A5 binder tested with CBP 39

Figure 22: Fluorescent data for binder C3 with S100B 40

Figure 23: Fluorescent data for binder A1 with CBP 41

List of Tables

Table 1: Binder C3 39
Table 2: Binder C4 39
Table 3: Binder C5 40
Table 4: Binder C6 41

Introduction

Protein-Protein Interactions

Living things are composed of a multitude of cells. The human body contains trillions of cells, all working together to keep us alive and functioning.[1] Inside of a cell, proteins work together to keep the cell running and are involved in every cellular process.[2] Proteins are composed of amino acids and interact in a complex, dynamic manner.[3] The interactions between proteins are composed of individual proteins and their binding partners. Protein-Protein Interactions (PPIs) are vital to the normal function of the cell. These PPIs can chain together in a cascading manner, forming complex networks of interactions that manage diverse functions within the cell. The totality of these PPI networks is known as the interactome of a given organism. There are around 130000-650000 types of PPIs in the human interactome.[4] These networks have tolerances to some disruptions to their constituent PPIs, helping maintain homeostasis; however, PPIs can have mutations or be disrupted to the point that they do not function in the way they are supposed to. When this happens, the overall PPI network can be disrupted, leading to disease states in the cell. Characteristic disease states are associated with many different diseases, and important to the accurate modeling and treatment of a disease.[4]

Purposefully disrupting these PPIs can help understand the functions and kinetics within the networks. This can be done through the inhibition of individual proteins in the network. A typical small molecule drug discovery approach mainly focuses on protein-ligand interactions because the proteins contain well-defined ligand-binding sites that small molecules can interact with. A new science has been emerging that allows for the control of PPIs using drugs, which allows for the targeted treatment of diseases.[5] The human interactome has more than 645000 PPIs that are correlated to diseases.[6] By 2011, only 2% of the PPIs had therapeutic drugs to

target them. The rest have been considered undruggable because PPIs are difficult to modulate and remain poorly understood.[4] Despite these challenges, controlling PPIs is a powerful treatment option and some drugs that do so have made it to the market. Using small molecules to inhibit PPIs can possibly lead to creating therapies and treatments for multiple diseases, broadening their use case.

As new technologies are created, the ability to create PPI-based therapeutics becomes feasible.[5] Some of these new technologies focus on "hotspots" within proteins that are the main drivers of their interactions. Mimicking or co-opting these hotspots has made the designing of drugs for PPIs more efficient. Hotspots are found using structural biology, and they consist of amino acids that play a vital role in where binding interactions occur between proteins.[4] Alanine-scanning mutagenesis and X-ray crystallography are the common ways that hotspots can be found.[5] With the knowledge of hotspots, they can help in the computational design of proteins. This thesis will explore the computational design of proteins to inhibit three target proteins in the p53 pathway in the PPI network to gain a deeper understanding of the kinetics and the overall PPIs for potential therapeutic targets.

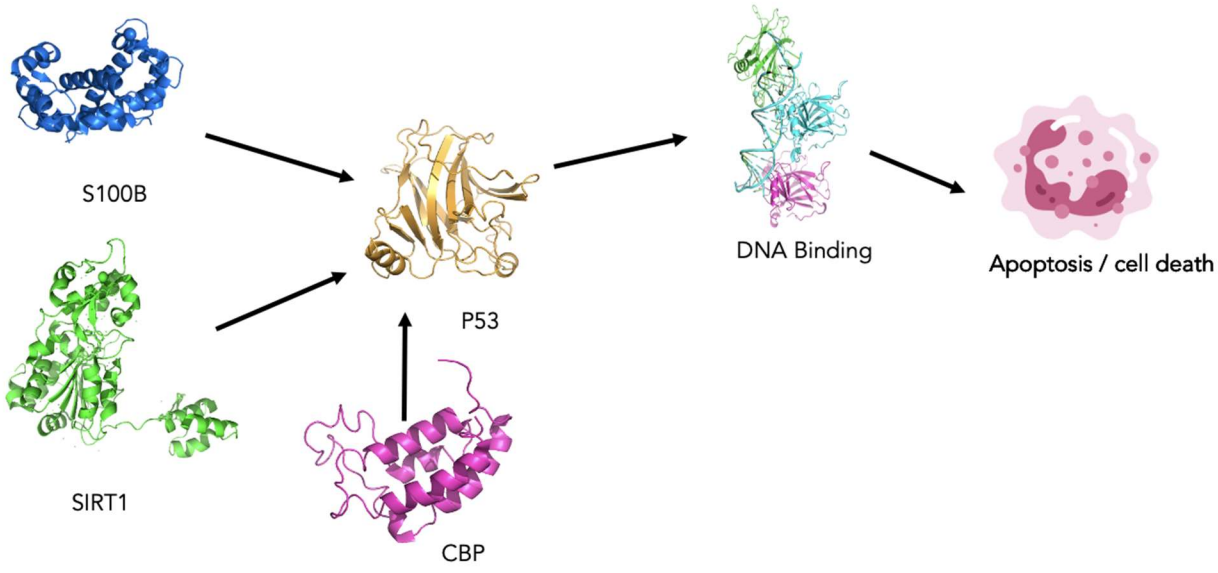


Figure 1: P53 Pathway

This is part of the P53 pathway that is the focus of this thesis. The pathway includes the three target proteins that bind to the c-terminal disordered region of P53. It also describes the main function of P53. When the cell is in a disease state, P53 becomes upregulated, leading to an increase in DNA binding, causing apoptosis or cell death. This is to prevent the spread of diseased cells.

Background on Proteins

p53

P53 is an important protein in the PPI network that plays a hub protein. Hub proteins are proteins that have many interacting partners in the network.[7] P53 is a widely studied protein due to its important role in preventing cancer and in fact, 50% of cancers have been correlated to mutations in the p53 pathway.[7,8] P53 is a tumor suppressor with numerous functions, and an important function is the induction of apoptosis, the term for programmed cell death. Apoptosis

is extremely important to the suppression of cancers because it kills the cancer cells before they can spread. Apoptosis is induced when DNA is damaged, or the cell is stressed.[9] P53's presence leads to the transcription of other proteins that promote apoptosis. The other functions of p53 are related to control of tumor angiogenesis and downstream cell cycle regulation.[7] P53 does this through the activation of P21, which stops the division of cells, halting the cell cycle.[8] P53 consists of three intrinsically disordered domains known as the transactivation domain, the linker region between the DNA binding domain and the tetramerization domain, and the c-terminal disordered domain.[10] C-terminal domain (CTD) is the part of p53 that changes conformation and is correlated to multiple gene products.[11] The CTD is one of the reasons behind the vast set of functions of p53.[7] Three proteins that we are exploring in this project are known to bind to different conformations of the CTD. These proteins are S100B, SIRT-1, and CBP and will be known as the target proteins for this project.[10] This thesis will explore the creation of computationally designed inhibitors for these proteins to understand the p53 network during disease states. The whole project includes the creation of inhibitors for all three of these proteins to gain a full understanding of the pathway. This thesis will focus on the computational design of S100B inhibitors and testing of designed inhibitors for all three proteins.

S100B

S100 calcium binding protein B (S100B) binds to the α-helical confirmation of the c-terminal disordered region of p53.[10] S100B is part of the S100 family that are calcium binding proteins and were named because they are soluble in 100% saturated ammonium sulfate. They are the largest family of EF-hand calcium binding proteins.[9] S100B is linked to malignant melanoma, which is a type of skin cancer that is very dangerous. When S100B binds to p53, it

inhibits apoptosis induced by p53.[9] The binding between p53 and S100B is a calcium dependent interaction.[12] When it binds, it inhibits p53 which stops p53 from suppressing tumors, leading to cancer, usually malignant melanoma. S100B has a calcium binding domain which changes the conformation of the protein and activates the binding structure. S100B is a calcium binding protein, so it naturally binds calcium weakly, but when S100B is bound to another protein, the calcium binding is stronger.[12] There are many members of the S100 protein family, which makes it difficult to create inhibitors to specifically target S100B.[13]

CBP

CREB binding protein (CBP) binds to β-turn confirmation of the c-terminal disordered region of p53.[10] CBP is a coactivator with P300 and mediates transcriptional activation. It was named due to its binding to the kinase inducible activation domain (KID) of CREB.[14] The KIX domain of CBP participates in a PPI with pCREB. P53 has also been shown to interact with the KIX domain of CBP.[15,14] CBP is recruited by the activation domain of p53.[14] It regulates the activity and stability of p53, and it increases the sequence specific DNA binding activity of p53.[16] CBP interacts with p53 after a lysine on the c-terminal disordered region has been acetylated. CBP is recruited to p53 target promoters and helps mitigate the transcriptional activity of p53, which is important in how p53 suppresses tumors.

SIRT1

SIRT1 binds to the β-strand confirmation of the c-terminal disordered region of p53.[10] SIRT-1 is part of the family Sirtuins that are NAD-positive-dependent-class III HDACs. They play an important role during chemotherapy for the tumor cells surviving and resisting drugs.

Mammalian Sirtuin 1 is a direct homolog of yeast Sir2.[17] SIRT-1 regulates insulin secretion and protects cells against oxidative stress.[18] It also modulates cell survival by regulating the transcriptional activities of p53. SIRT-1 activates stress defense and DNA repair mechanisms.[17]

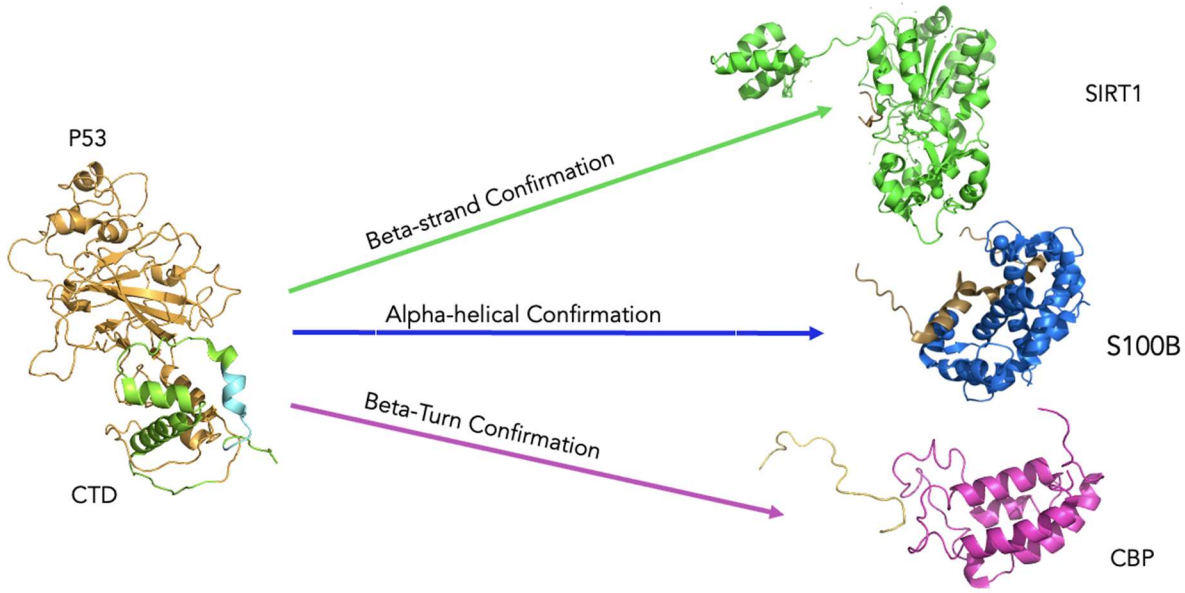


Figure 2: Interaction of Target Proteins with CTD of P53

The three target proteins interact with the C-Terminal Disordered region with different conformations that are shown.

Interaction of S100B and P53

The interactions between two proteins are incredibly complex and dynamic. When two proteins interact, they usually have great chemical and geometric pairings. During the interactions, electrostatic, hydrophilic and hydrophobic forces play an important role in the strength of protein-protein interactions. There are certain amino acid residues that are key to the interactions between two proteins, and these are called hot spots.[19] The interaction between S100B and p53 is a calcium dependent interaction. In the presence of Ca2+, S100B undergoes a conformational change that is required for the interaction with p53. The binding occurs at the C-

terminal domain of p53 which has no native structure but is helical once bound to S100B. The binding of S100B and P53 occurs due to hydrophobic and electrostatic interactions.[20]

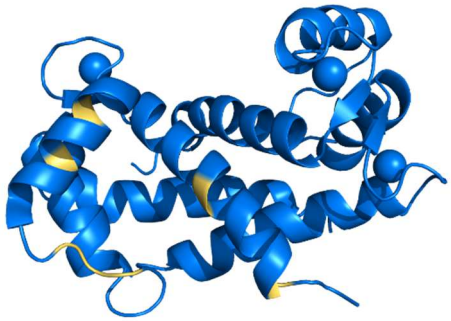


Figure 3: S100B with Highlights P53 Binding Hotspots

For the design of an inhibitor to S100B, hotspots are the amino acid sequence in which the interaction of P53 and S100B occurs. These hotspots are important for designing an inhibitor.

Computational Protein Design

In recent years, proteins have been designed using computation methods. The computational methods are based on the physics of the interactions between amino acids of proteins along with using knowledge of known structures and sequences. The first computationally designed protein was a zinc-finger protein designed by Dahiyat and Mayo in 1997.[21] The exciting step forward in the world of de novo design was in 2003 when Top7 was designed. Designed by David Baker and his lab, this was the first complex protein that had no structure or sequence similarity to any protein in the pdb bank, revolutionizing the protein engineering world.[22] This first official de novo designed protein was designed using the Rosetta Computer Program they developed for protein structure design and prediction.[23] This program has been used and improved and in 2023, RoseTTAFold Diffusion (RFDiffusion) was published.

RFDiffusion is a generative diffusion model that generates new protein structures using iterative denoising steps.[24] Another machine learning model used for computational protein design is ProteinMPNN. ProteinMPNN uses physics-based approaches and a message-passing neural network (MPNN) to predict protein sequences based on a given backbone.[25] An important part of protein design is to determine if the structures and sequences designed will actually fold into a conformation that will actually work in vivo. AlphaFold and ESM Fold are machine learning models that can predict a structure based on sequences and give metrics about the prediction of the structure.[26,27] To determine if the designed protein will have favorable interactions with the target used for the design, the protein is docked against the target protein and the energies of the conformations are scored to predict if binding will occur. AlphaFold-initiated Replica Exchange Docking (AlphaRED) is a docking machine learning model that has a success rate of 43%.[28] Another Machine Learning Model called EvoPro iterates between AlphaFold and ProteinMPNN and has shown success with protein binder design.[29] Overall, computational protein design is a new and promising field that incorporates a multitude of machine learning models.

Methods

Initial Computational Methods

To begin the computational design process, a deep understanding of the interaction that will be inhibited is needed. Through the use of the protein data bank (PDB), the structure of S100B binding to the C-terminal disordered region of P53 was found (PDB code: 1DT7). 1DT7 was then brought into protein modeling software called PyMOL. The structure was relaxed using pyrosetta, which is a process that samples movements of the protein structure in order to minimize it into a lower energy conformation, before being used as the target protein of interest.[30] A piece of the P53 CTD is used as a starting point for the creation of an inhibitor. This inhibitor will itself be a de-novo designed protein. Starting from a piece of a known interaction to create an inhibitor increases the chances of creating a successful inhibitor that actually binds to S100B when tested experimentally.[31] The amino acid residues that are important to the interaction of P53 CTD and S100B are noted and are used as the "hot spots" of the protein interaction. These hotspots are used to guide the design process and imputed into RFDiffusion. RFDiffusion is a machine learning model that is used to generate multiple plausible protein backbones with no particular amino acid sequence.[24] These backbones can be thought of as templates that we can assign amino acids sequences to, that will take on the backbone structure. After RFDiffusion creates multiple protein backbones, we assign an amino acid sequence to them using another machine learning model known as ProteinMPNN. ProteinMPNN uses the backbone created by RFDiffusion as input and assigns a sequence that is likely to take on the specific backbone structure.[25] The proteins with assigned sequences are then fed into AlphaFold2 multimer (AF2). AlphaFold2 multimer then predicts the structures of the protein inhibitor-S100B complexes given on the amino acid sequences. If AF2 predicts the structure to be similar to our

design with high confidence, then we keep those designs for experimental validation. The two most important scores that have been validated in literature are the pLDDT and the PAE. pLDDT is the predicted local difference distance test. The pLDDT is a measure of how the protein will fold. A pLDDT above 80 has a high likelihood of folding as designed. PAE is the predicted alignment error and is a measure of how much error there would be between the AF2 prediction and the true structure of a protein. The PAE can inform how well the binder will bind to the target protein. A PAE lower than 10 will have a higher chance of binding to the target protein. Once the protein designs are filtered through AF2, we are left with a small subset of candidate inhibitors with high pLDDT scores and low PAE.[32] Those binders are put into a modeling tool called Rosetta Docking to simulate their binding by putting the binders at different orientations and distances around the target and evaluating the energetics of the pair. After refining all the initial RFDiffusion designs to a set 20 binders pass all our filters, they are ordered as DNA oligonucleotides coding for the protein and the experimental validation portion begins.

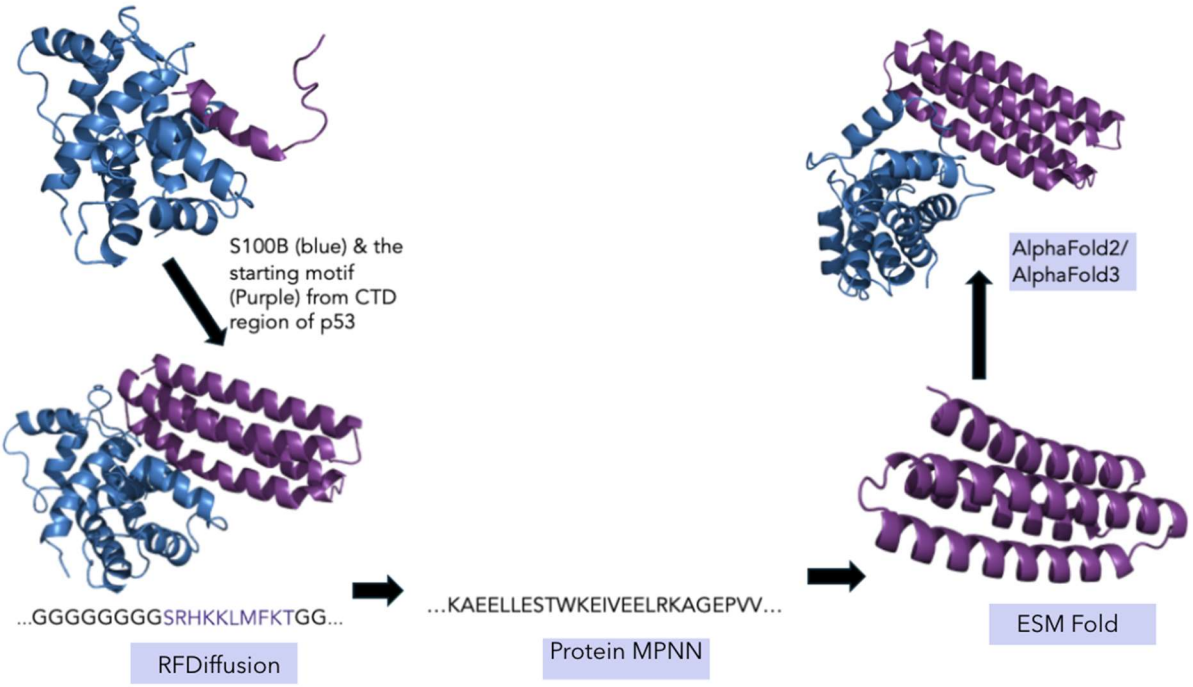


Figure 4: Protein Design Pipeline

The Design Pipeline outlined that was used to computationally design inhibitors to P53.

Initial Experimental Methods

Initially, to be able to use S100B, SIRT1 and CBP in the lab, they were ordered as DNA oligonucleotides. Once arrived, Gibson Assembly was performed which takes the DNA ordered and inputs it into a plasmid (pet28a+), which controls the expression of the proteins. The plasmid is then transformed through heat-shock transformation into NEB Turbo cells, a strain of E. coli that amplifies plasmids, to create multiple copies of the protein. After it is plasmid extracted to isolate the amplified plasmids. It is then sequenced to confirm the correct insertion of the gene encoding the protein of interest. If the sequence shows a successful transformation and integration of the gene, another heat shock transformation is performed of the resulting plasmid into BL21-DE3 e. coli cells which specialize in the expression of recombinant proteins. The transformed cells with the protein inserted in them are then grown up to a certain concentration

and spiked with IPTG to induce protein expression. The protein is then purified from all the components of the cell using immobilized metal affinity chromatography, employing Nickel, which binds a poly-histidine tag we intentionally add to the c-terminus of our proteins of interest. After purification, bio-layer interferometry was used to measure their binding affinity.

The method shown above was initially the process taken to express and purify the target proteins and the method that was going to be used for binders. The limitation of this method is that it is time consuming and would take a long time if used for all the binders, because each would have to be individually expressed.

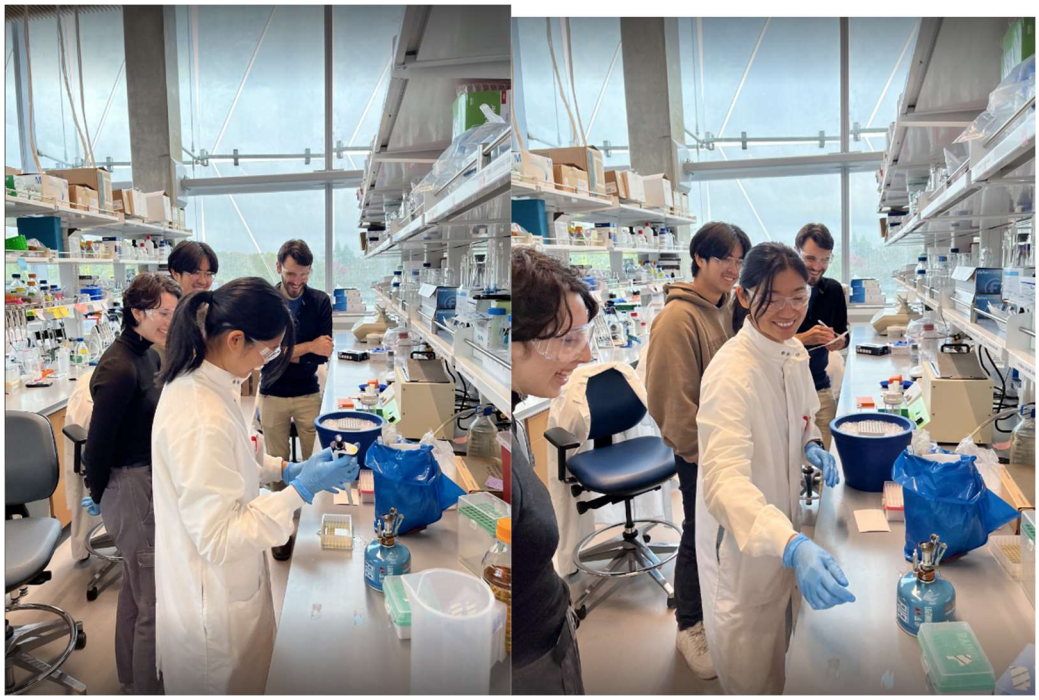


Figure 5: Testing the new wet lab method in lab

Graduate Student Karly Fear demonstrates the high throughput 96 well method to everyone.

Optimized Experimental Method

A new method was introduced to our lab by Karly Fear which encompassed High Throughput 96 well plate cloning. This would allow for the expression and purification of multiple proteins at once, allowing the expression and purification to be done in two days instead of two weeks like

the previous methods used. Through this protocol, the designed binders and target proteins were ordered in a 96 well block plate. Golden Gate assembly, another method to input the DNA ordered into the plasmid of choice, was performed for multiple constructs. The plasmid was transformed into BL21-DE3 cells and were grown up. The protein was expressed using auto induction media which induces protein expression using lac operons. The proteins were purified using immobilized metal affinity chromatography, and the results were tested using gel electrophoresis.

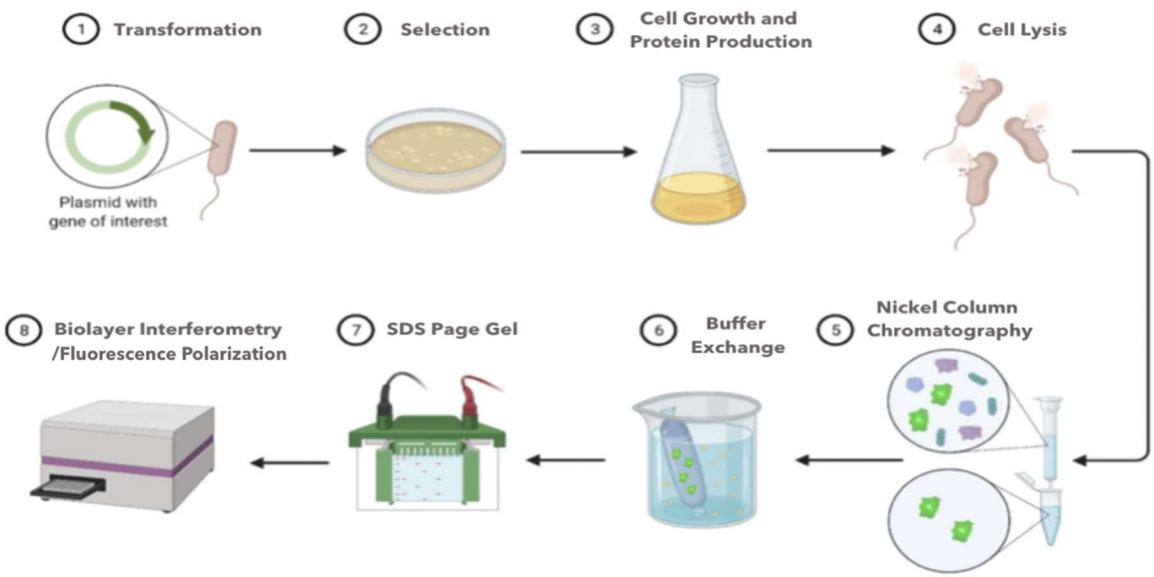


Figure 6: Wet Lab Workflow

The process of all the wet lab methods used for the expression, purification, and testing of the proteins.

Testing Binders

The binders were tested using two different methods:

Biolayer Interferometry

Biolayer Interferometry (BLI) is an optical method that is used to test the binding affinity of proteins. During testing, it requires no labelling of proteins and is able to measure the rate of

binding. It is able to measure the thickness of the bimolecular layer attached to the tip of the probe and uses interferometry to do so. It uses the light reflected off the probe and when a protein is attached, the length of the reflection increases and can be correlated to the mass. This is able to measure binding because the tip of the BLI probe has certain affinity tags that are initially used to pick up one of the proteins that have a correlated tag. The probe is then placed into a solution that has a protein that is a possible binder to the attached protein. If the protein binds to the other protein, the length of reflecting light increases and that change is able to be measured and is used to determine binding affinity.[33] Some limitations of BLI is that it takes a long time to prepare and is not useful in testing a multitude of proteins at once.

Fluorescence Polarization

Another setup that was used to test the binding affinity of the designed binders was fluorescent polarization. Fluorescent polarization uses polarized light to test if a protein is binding. One protein has a fluorophore attached, and the other protein is either much larger or is attached to a larger protein. When the protein with the fluorophore attached is in solution, it is rotating due to Brownian motion. If no binding occurs, it is by itself and has a fast rotation, meaning that the polarized light that excites the fluorophore becomes depolarized. If binding to the larger protein occurs, the rotation is slowed significantly, so the polarized light after exiting the fluorophore stays polarized. The amount of polarized or depolarized light present helps determine binding affinity of the proteins and gives an idea into if the designed inhibitors bind.[34]

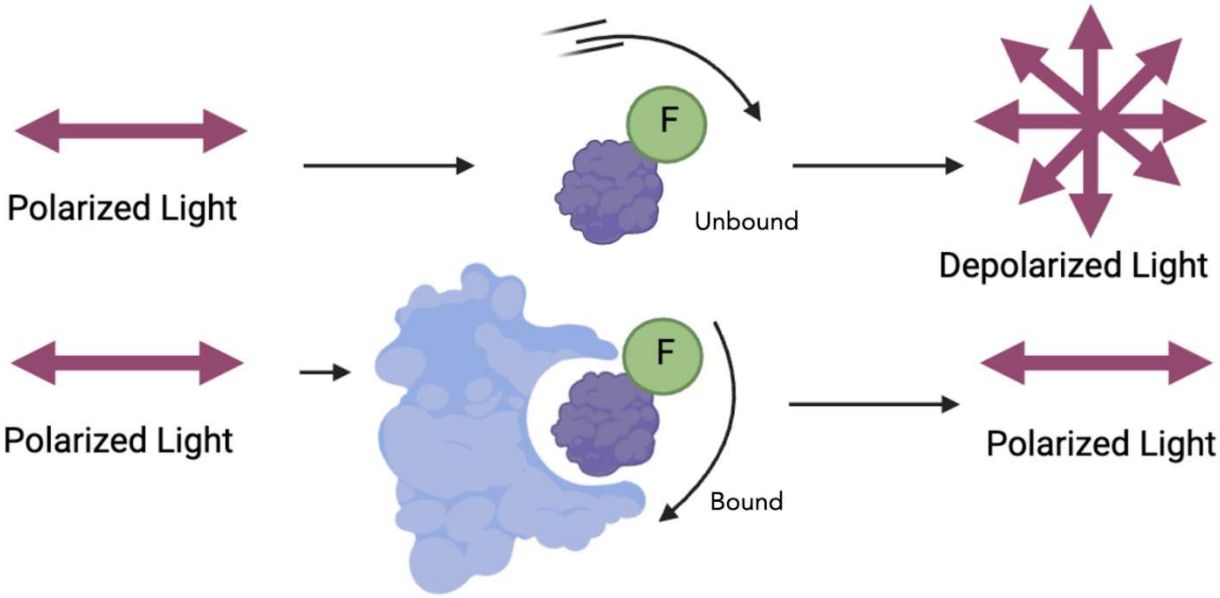


Figure 7: Fluorescence Polarization

Demonstration of how fluorescence polarization works. The bound protein spins slower and maintains polarized light due to Brownian motion. The unbound protein spins faster and depolarizes the light when excited.

S100B Computational Design Optimization

S100B is known to be a difficult target to design binders for due to the binding site having a mixture of multiple small helical sites, creating a binding pocket. To begin the design process, RFDiffusion was run, and the design pipeline was done on all the initial designs. When RFDiffusion was initially run, there was a mixture of designs that looked feasible and some that were just a single long strand of alpha helical loops. When the designs were examined using ESM fold and AF2, there were low pLDDT scores seen and no PAE scores were below the cutoff of 10. pLDDT is the predicted local distance difference test and a pLDDT score above 80 indicates that there is a high confidence in the predicted structure. To move forward with the designs, pLDDT scores above 80 and ideally above 90 are needed. These initial designs had very few above 80 and only a couple above 90. The PAE is the predicted aligned error and is important to have a low PAE, below 10, for a design to be feasible.[32] The initial designs had no PAE scores below 10, so optimization of the pipeline was needed.

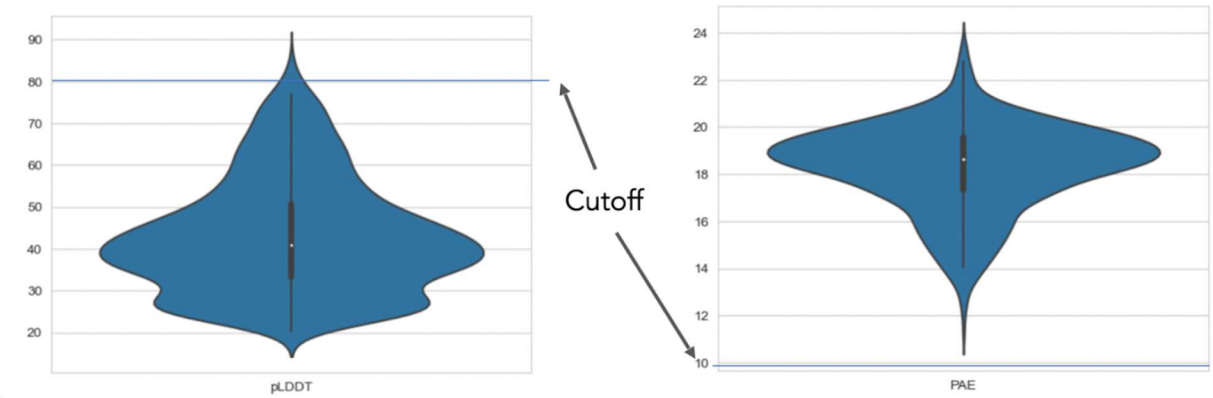


Figure 8: Initial Design Metrics

Initial design metrics before optimization. The results were poor and needed imporvement to meet the cutoff.

After reviewing different initial parameters in RFD to optimize, it was found that the radius of gyration (ROG) along with the interface contacts (interface n contacts) could have different inputs placed in which would change the parameters in which RFD designed binders by. The ROG potential controls how compact the model designs the proteins, and the calculations consist of the root mean square difference of the Calpha atoms from the center. A higher weight will give a more compact ROG. The Interface contacts shift the model to prioritize more contacts at the interface between the target protein (S100B) and the designed binder. A higher weight will gear more towards having more contacts at the interface. For both ROG and interface contacts, the recommended weight range is from 1-10.[35] Once these parameters were changed, outputs of RFD differed and the pLDDT and PAE of these outputs were compared.

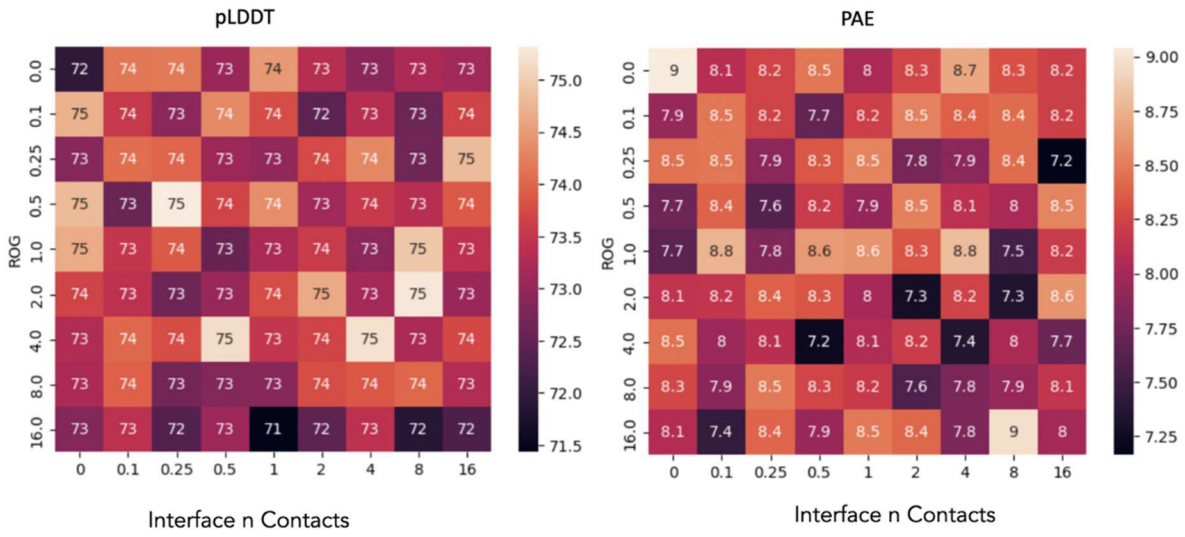


Figure 9: Optimization of Potentials

Different potentials for RFDiffusion were optimized and their combonitaions were displayed with the Alpha Fold Metrics.

It was seen that a ROG of 8 and an interface contact of 2 yielded a combination of the lowest PAE and highest pLDDT. The ROGs with the highest average pLDDTs were 0.5, 4, and 2. The interface contacts with the highest average pLDDTs were 0.25, 0.5, 4, and 8. The ROGs with the lowest average PAEs were 4, 2, and 0.25. The interface contacts with the lowest average PAEs were 0.5, 2, 8, and 16. The combinations of the best ROG and interface contacts were plotted for pLDDT and PAE distribution.

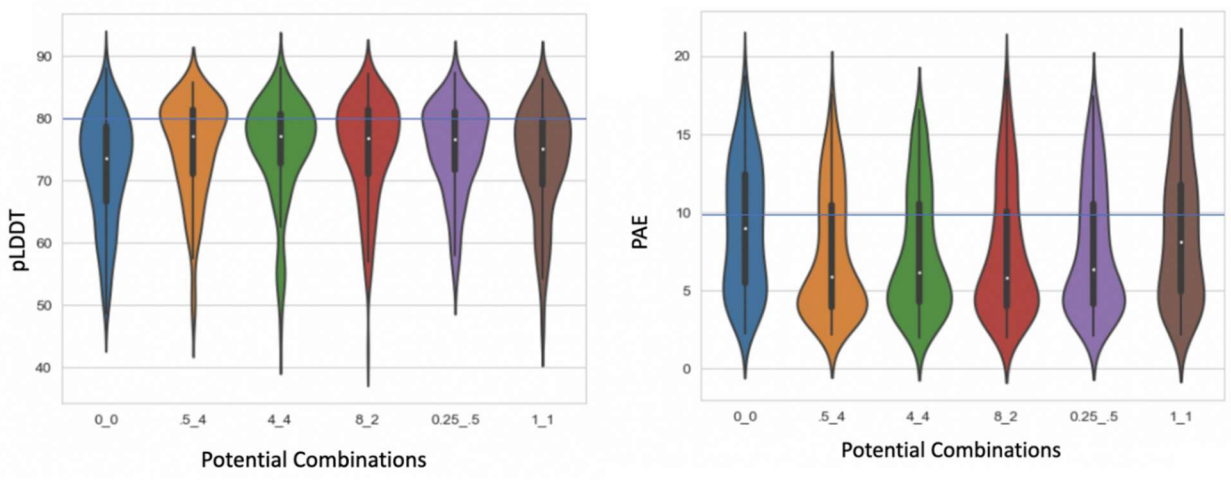


Figure 10: Best Potential Combinations

The potential combinations that had the best metrics were shown and analyzed by looking at the design distribution.

From these graphs, ROG of 0, 0.5, 4, 8, 0.25, and 1 were plotted with interface contact values of 0, 4, 4, 2, 0.5, and 1. The ROG and interface combination values chosen to move forward were 8 of ROG and 2 of interface contacts. Although the potentials were optimized, the designs overall did not have many low PAE scores or high pLDDT scores. To further optimize the designs, the denoiser scale was changed. The denoiser scales change the amount of noise added at the inference. Different denoiser amounts were tested, and the results are shown below.

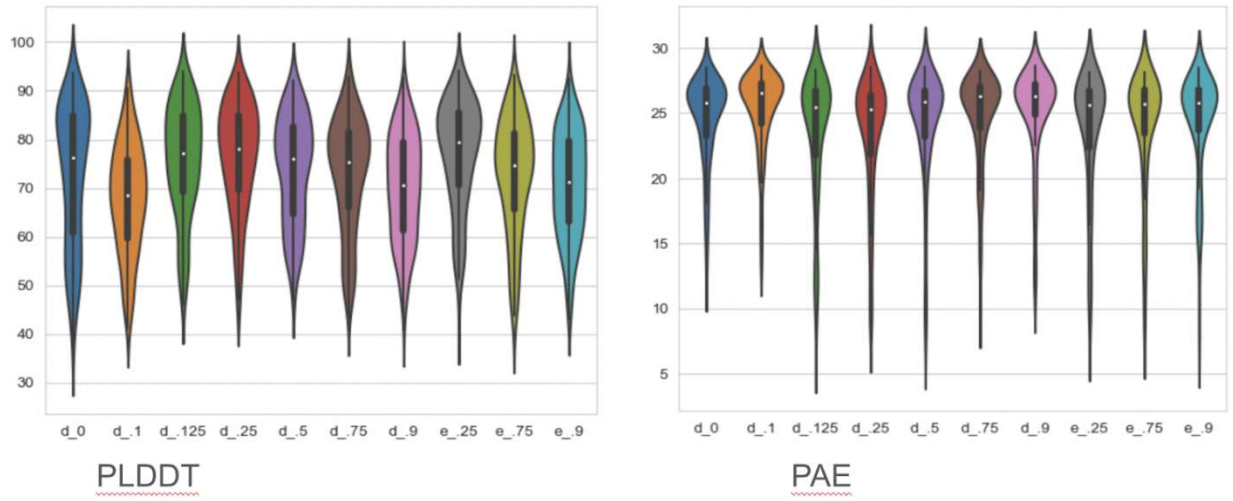


Figure 11: Denoiser Runs

RFDiffusion denoiser values were optimized, and their distributions were compared.

From changing the denoiser scale, there were no significant differences seen but it was determined that a denoiser scale of 0.125 gave the overall highest pLDDT combined with the overall lowest PAE. The optimization of RFDiffusion leads to overall higher pLDDT and lower PAE scores.

From the optimization steps, thousands of binders were generated using these parameters and five binders that passed the optimization steps with a high pLDDT and a low PAE were ordered for testing. Similar binder design pipelines were done, and similar optimization steps were performed for CBP and SIRT-1, and the best designs were ordered. These inhibitors were created by Cassandra Gonzalez (CBP) and Nguyen “Jenny” Nguyen (SIRT-1).

Results and Analysis

Once the sequences of the target proteins arrived after being ordered from IDT, they were cloned into the Pet28a(+) Plasmid using the Gibson Assembly method of joining DNA fragments. After the Gibson Assembly, the plasmids with the sequences inserted were transformed into NEB turbo cells. These cells are curated for great transformations and have rapid growth so they are great for newly cloned plasmids. To validate that the Gibson Assembly worked and that sequences were put into the plasmid, plasmid extraction was used, and the extracted plasmids were sent off to plasmidsaurous for sequencing. The three target proteins were all successfully inserted into the plasmids.

Figure 12: S100B plasmid insert compared against pet28a(+) and s100b sequence

The results show successful cloning.

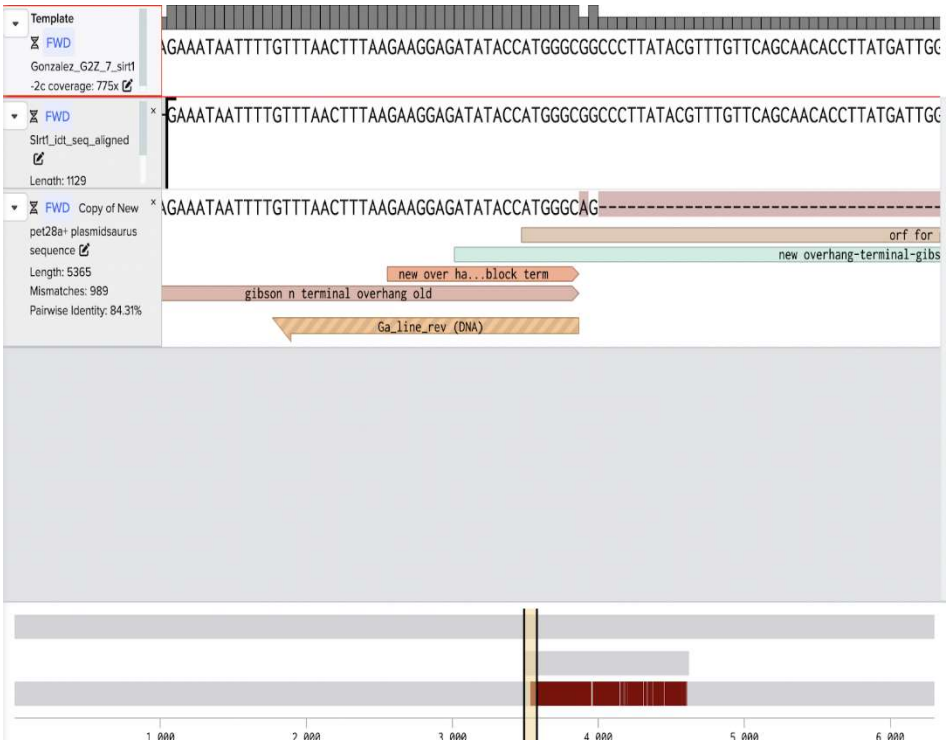


Figure 13: SIRT-1 plasmid insert compared against pet28a(+) and SIRT-1 sequence

The results show successful cloning.

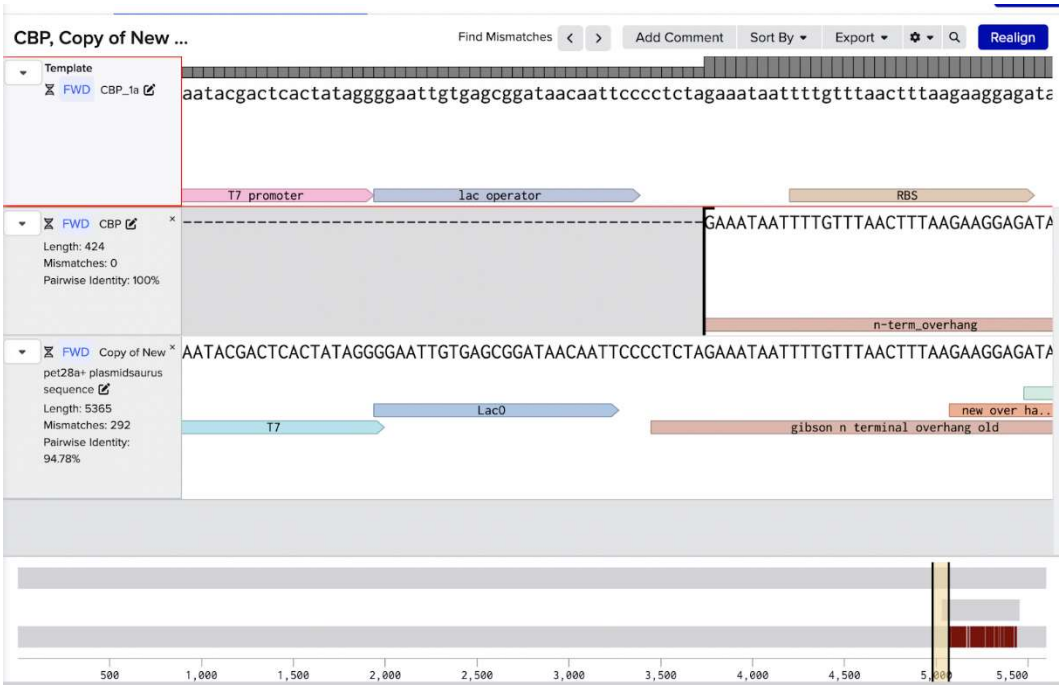


Figure 14: CBP plasmid insert compared to pet28a(+) and CBP sequence

The results show successful cloning.

The successfully cloned isolated plasmids were then transformed into BL21 DE3 cells and expressed using IPTG. The optimal expression conditions for S100B were induction at 18C overnight. The optimal expression conditions for SIRT-1 were induction at 37C for four hours. The optimal expression conditions for CBP were 30C for six hours. After expression, an abundance of proteins was created using BL21-DE3 cells and removed from the cells through cell lysis and sonication. Immobilized metal affinity chromatography using nickel chromatography is used to purify the protein. During the purification, the target proteins have a his tag that binds to the nickel beads, which allows them to be separated from the other components previously in the cell. Imidazole binds better to nickel and is then used at different concentrations. The low concentrations are used to wash away all the stuck debris and these are

called the wash steps. The higher concentrations are then used to replace the stuck proteins with imidazole and elute the proteins, and these are called the elution steps. The outputs from the nickel chromatography were run on a SDS page gel to determine if the protein was expressed and if the protein purified.

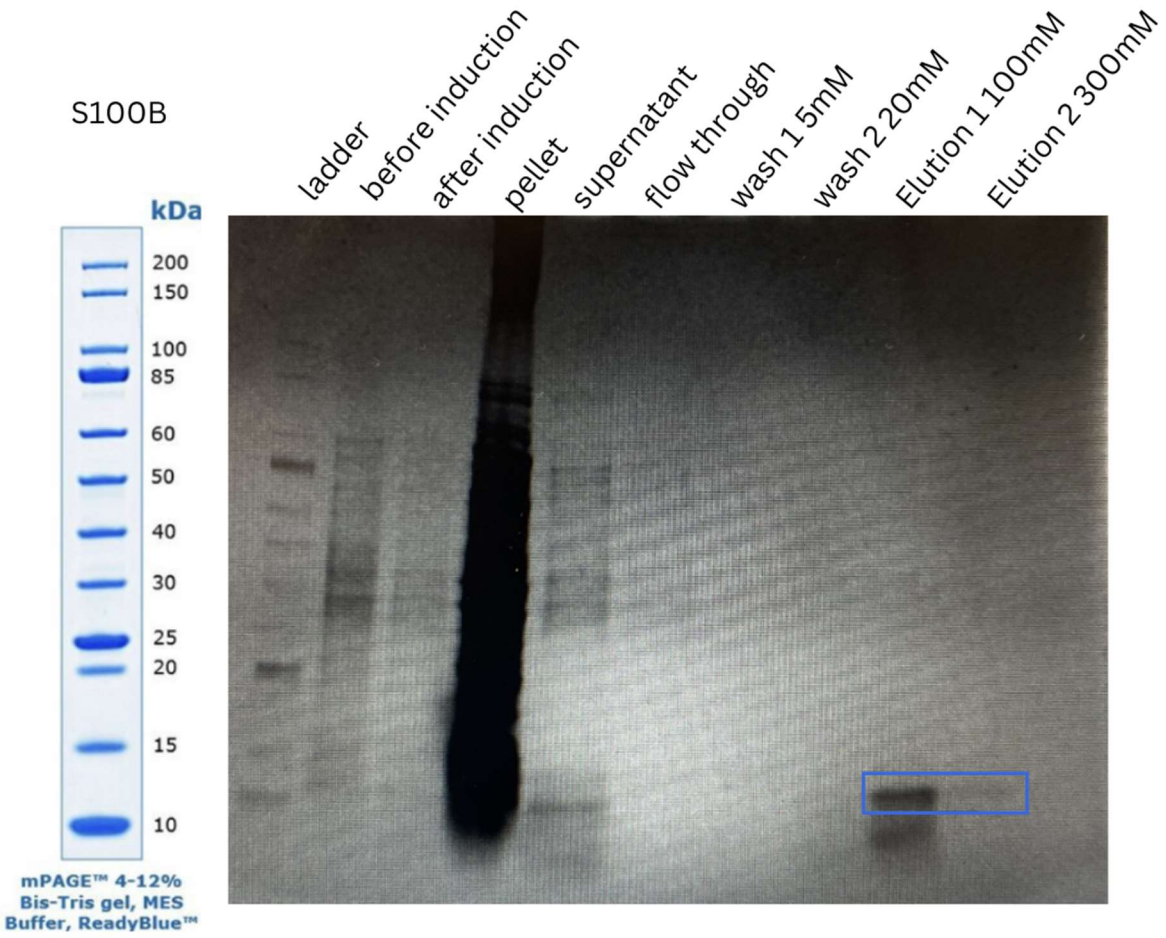


Figure 15: Purified S100B

The molar weight of S100B is 12 kDa. The gel of S100B purification shows that S100B was purified in Elution 1 and faintly in Elution 2.

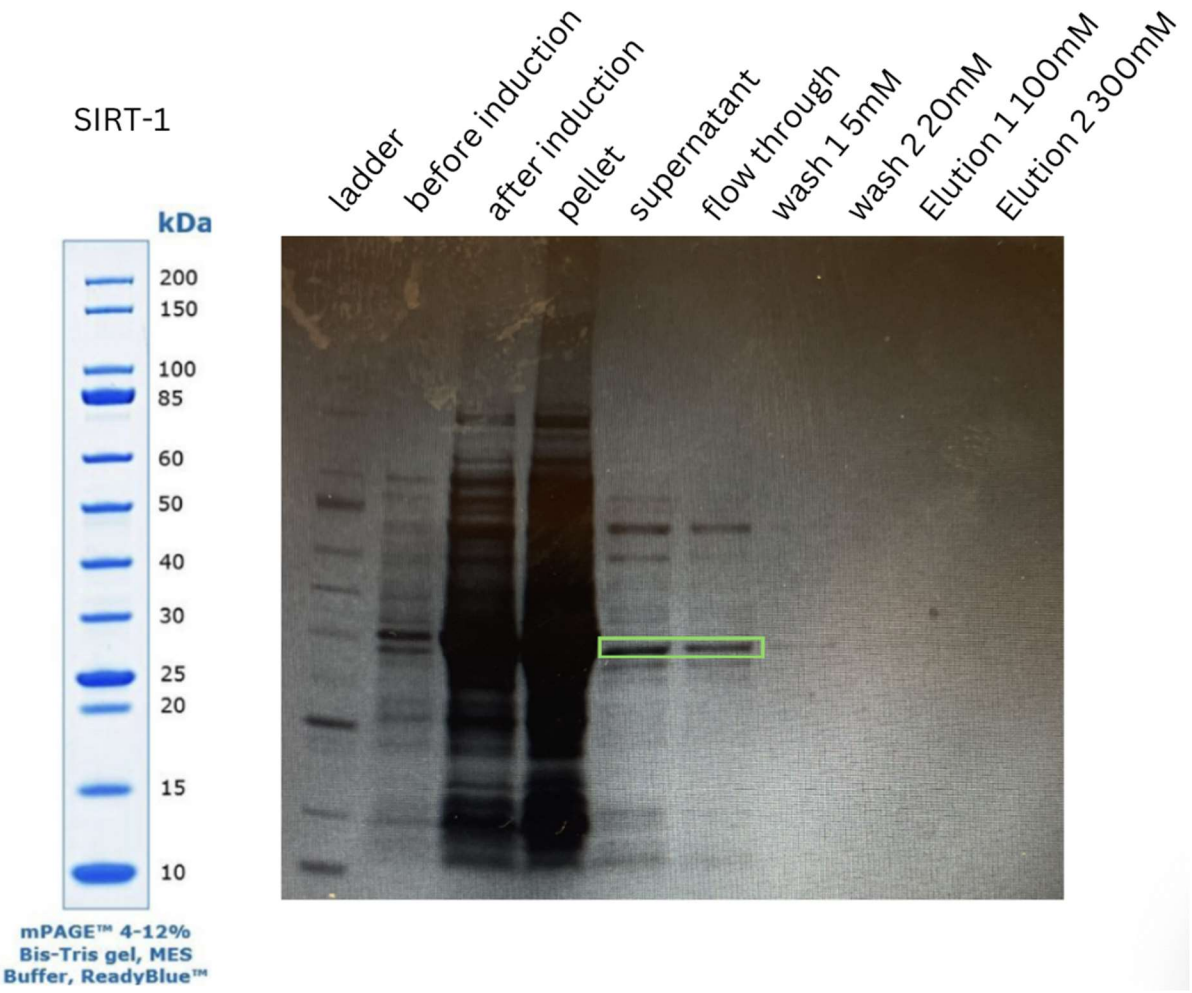


Figure 16: SIRT-1 Purification

The molar mass of SIRT-1 is 41 kDa. SIRT-1 was not purified and flowed through the column. This was experienced many times with SIRT-1 not purifying.

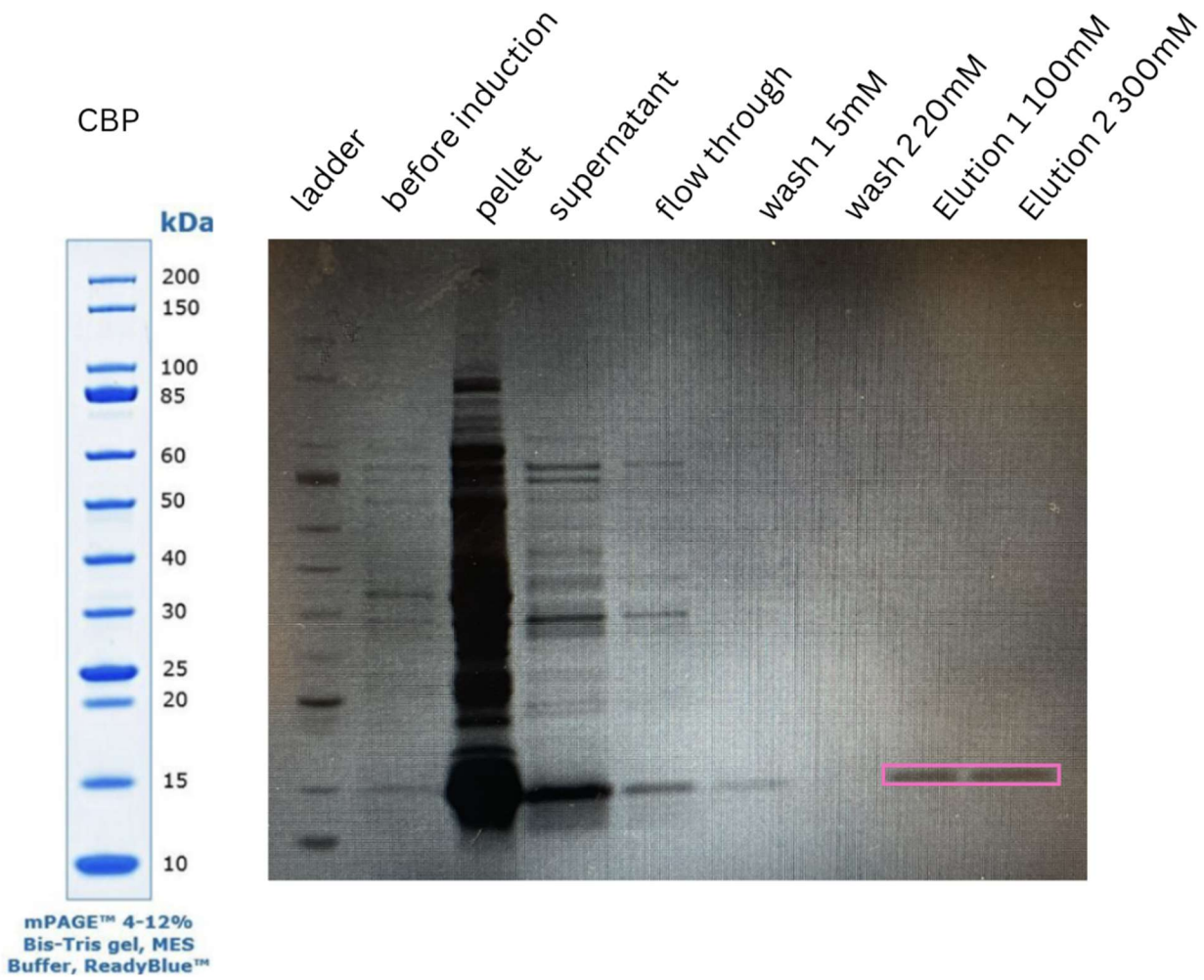


Figure 17: Purified CBP

The molar mass of CBP is 41 kDa. CBP was purified and was seen in wash 1, elution 1, and elution 2.

After the expression and purification of the three target proteins, the computationally designed binders were ordered in a 96 well plate. The binders are named after their placement in their individual wells on the 90 well plate. They were cloned into plasmids and transformed into BL21-DE3 cells. The binders were expressed using auto induction media and purified using nickel chromatography. Most of the binders successfully purified and only had one band in the gel, which meant they were usable for testing them against a target protein. To begin the testing

of binding affinity, BLI was used. CBP and the binders designed for CBP were used because CBP successfully purified and only had one band on the gel and many of the binders also purified.

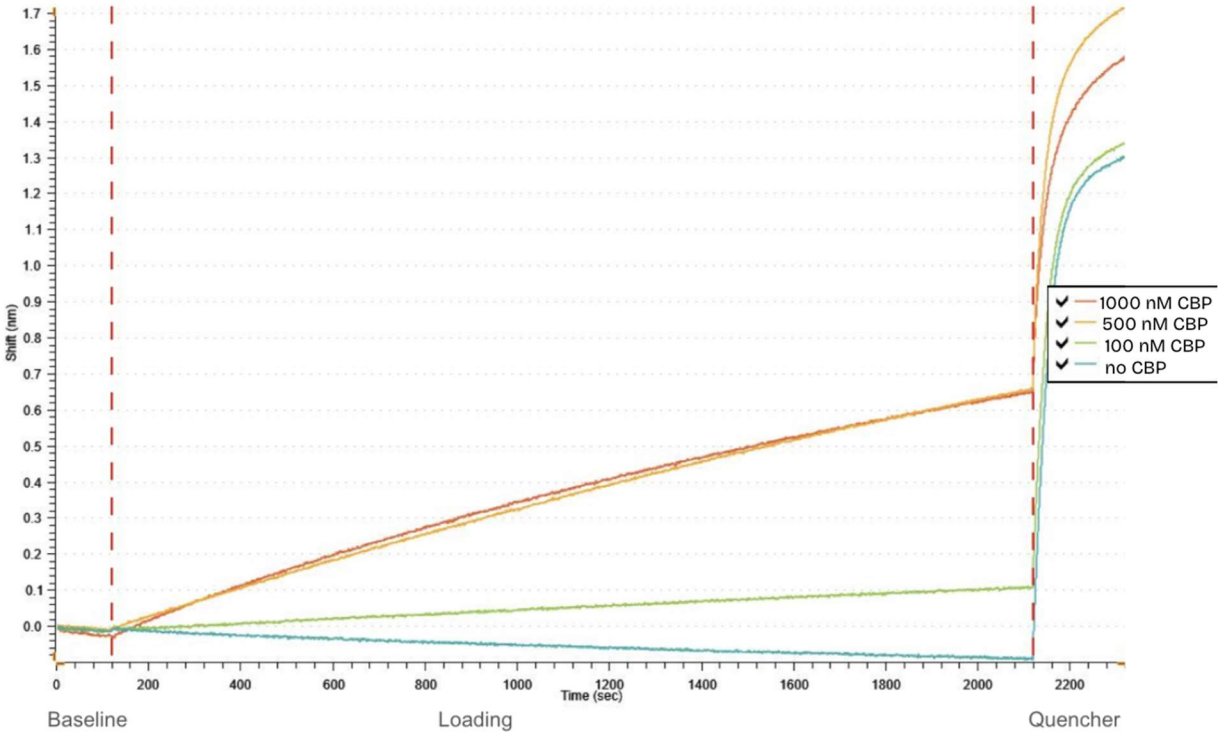


Figure 18: CBP Loading Test

The loading of CBP onto the probes was tested. It was shown that CBP did load onto the probes because the higher concentrations of CBP had a higher shift that increased with time. The probe with no CBP had a lower shift as time increased.

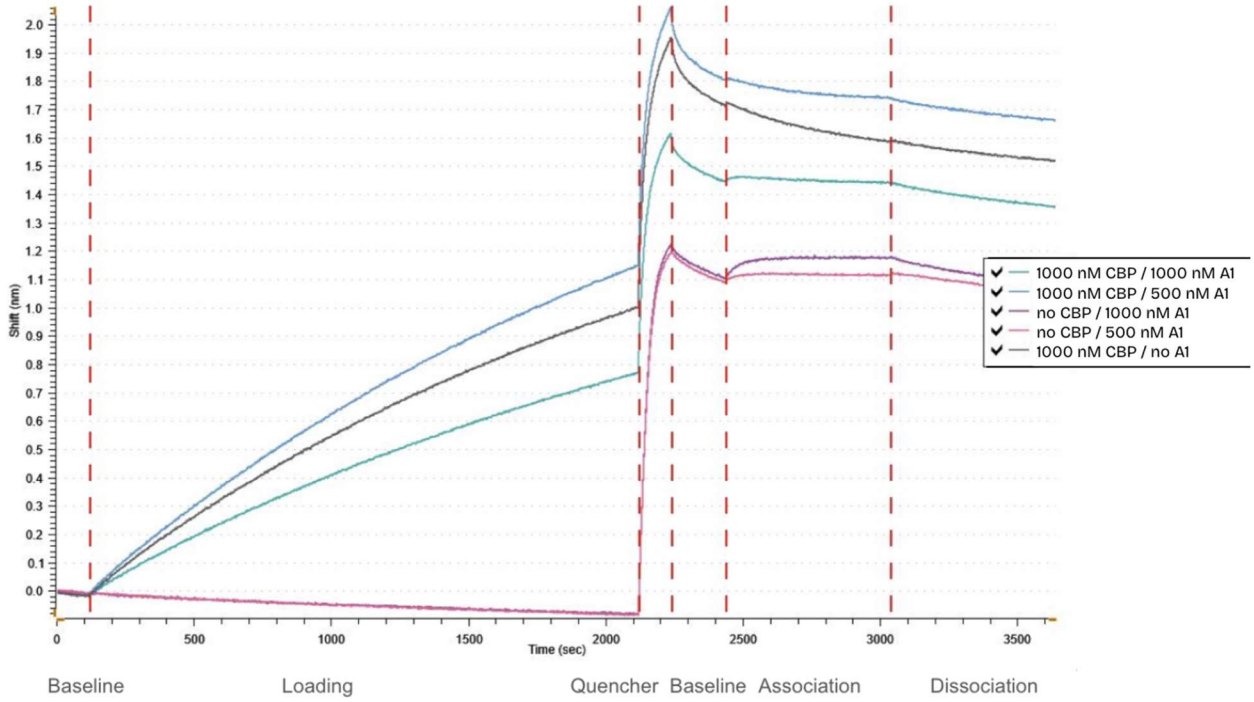


Figure 19: A1 binder tested with CBP

The binder named A1 was tested at different concentrations against different concentrations of CBP. The association step is when the binder would be loaded onto the probe and there was no increase in shift, so no binding occurred with A1 and CBP.

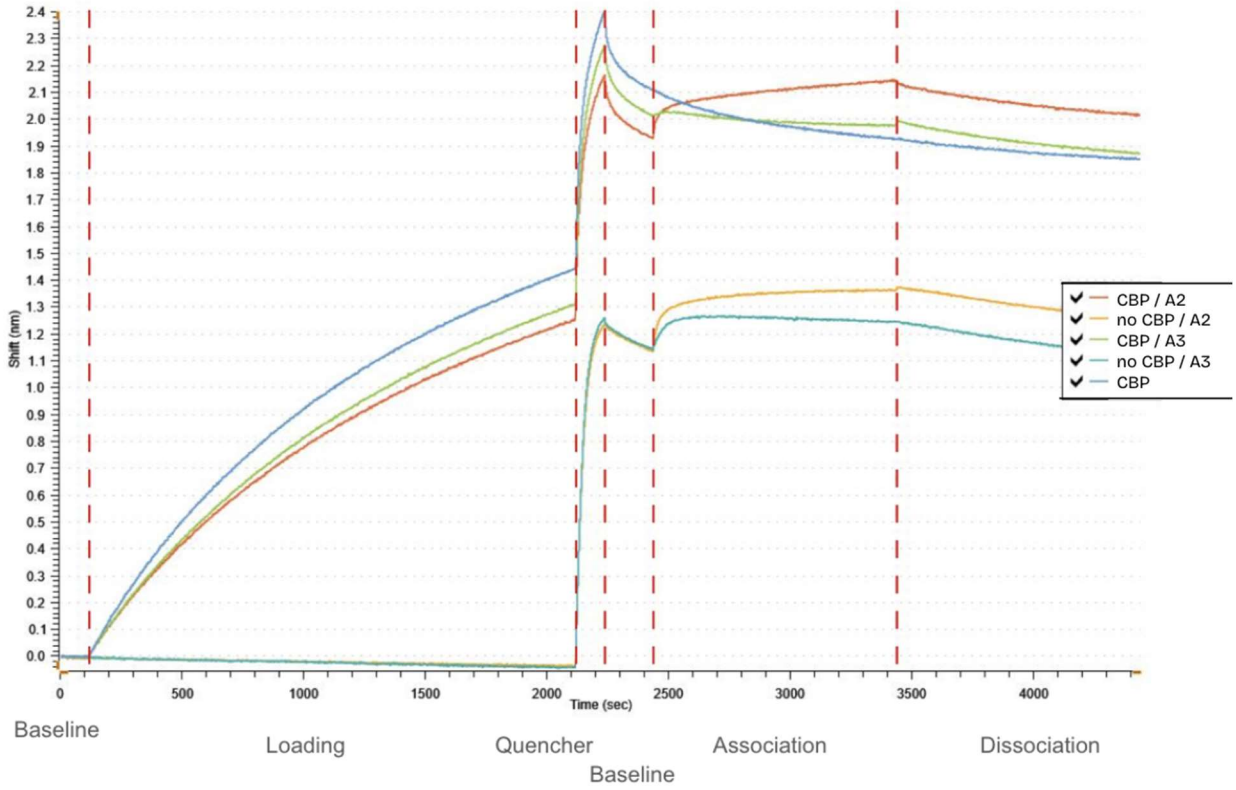


Figure 20: A2 and A3 binder tested with CBP

The binders A2 and A3 were tested against CBP. The association step is when the binder would be loaded onto the probe and there was some increase in binding seen, but it was also seen when no CBP was present, which meant there was some binding to the probe.

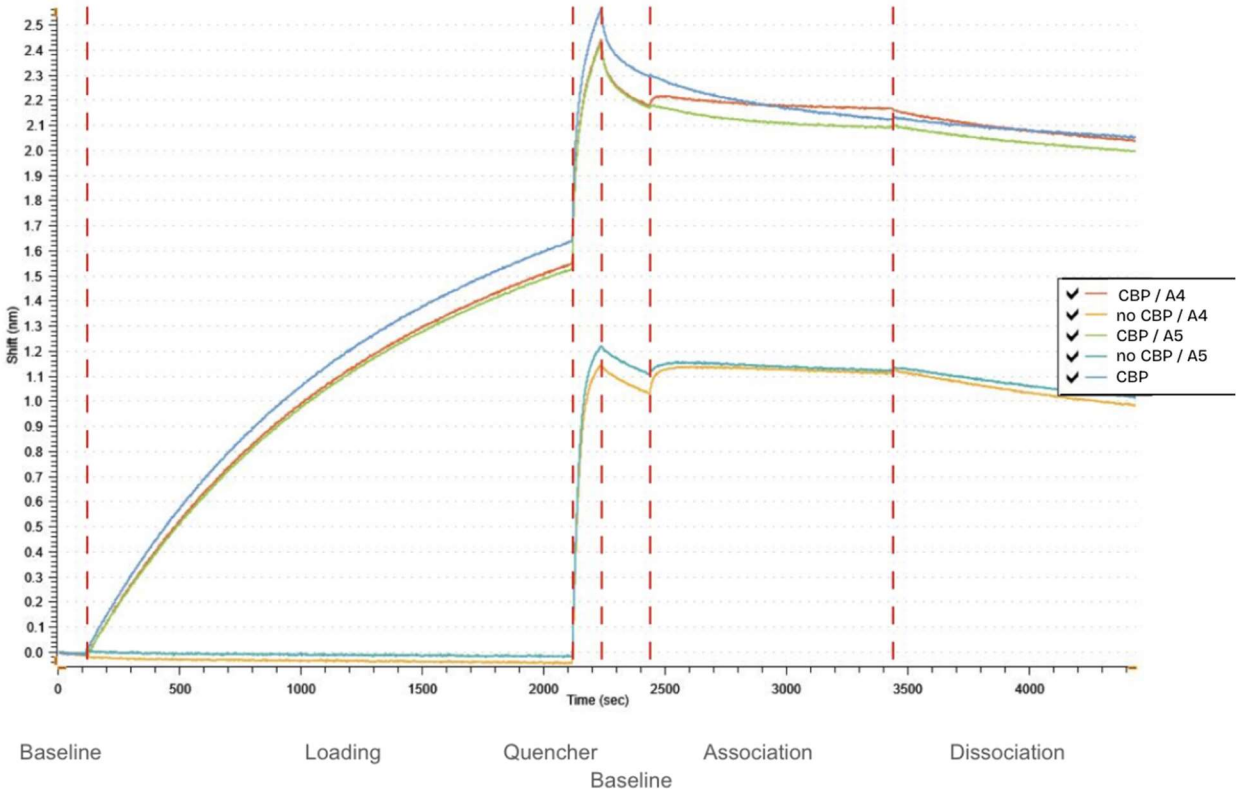


Figure 21: A4 and A5 binder tested with CBP

The binders A4 and A5 were tested against CBP. The association step is when the binder would be loaded onto the probe and there was some increase in binding seen, but it was also seen when no CBP was present, which meant there was some binding to the probe.

Overall, there were a couple of problems that arose when doing BLI. The first is that it is difficult to test many binders at once because many controls are needed to confirm results. The main problem that arose is that both the target protein and the binders had his tags. His probes were used for the BLI, which meant that the binders could also bind to the probe, so it was difficult to confirm accurate readings.

For more efficient binder testing, Fluorescence Polarization was used. With FP, the target proteins had the fluorophore, and the binders had a large protein attached.

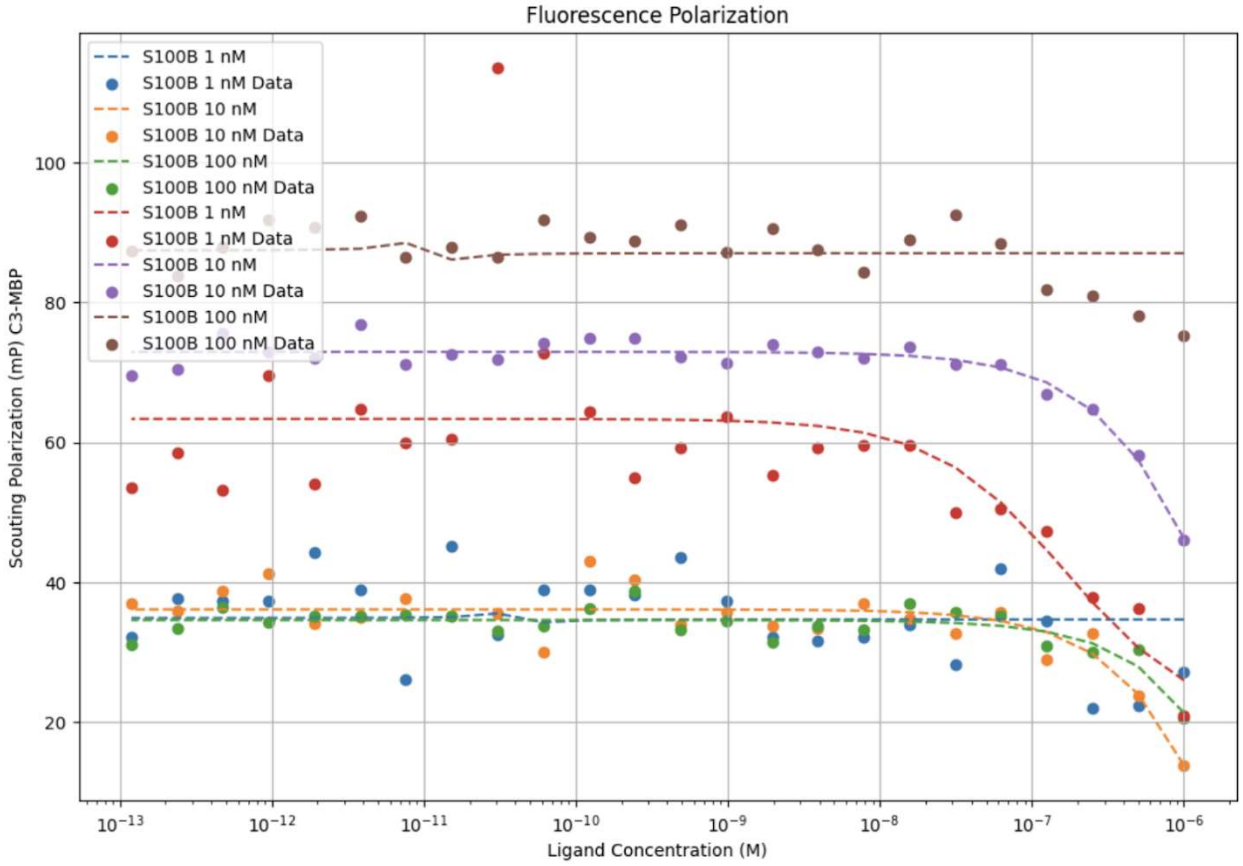


Figure 22: Fluorescent data for binder C3 with S100B

Different concentrations of S100B with the fluorophore attached. The two sets of S100B concentrations represent the fluorophore attached to the c-terminus and the fluorophore attached to the n-terminus to make sure the placement of the fluorophore does not interfere with binding. If there was binding occurring, there would be an upward trend with the data as concentration increased. No binding was seen and the other graphs for the other binders designed showed similar results, meaning that there were no successful S100B binders. The other graphs are in the supplementary information.

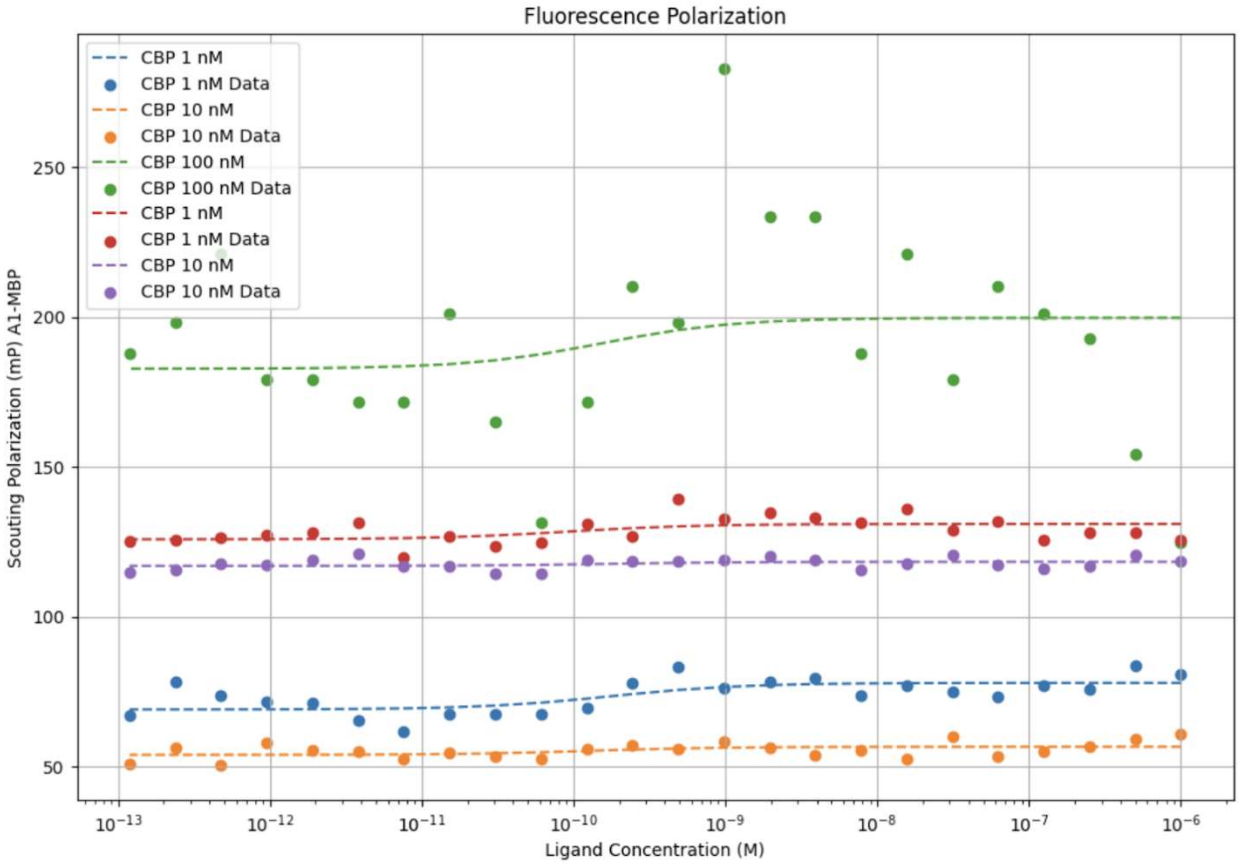


Figure 23: Fluorescent data for binder A1 with CBP

Different concentrations of CBP with the fluorophore attached. The two sets of CBP concentrations represent the fluorophore attached to the c-terminus and the fluorophore attached to the n-terminus to make sure the placement of the fluorophore does not interfere with binding. If binding was occurring, there would be an upward trend with the data as concentration increased. No binding was seen and the other graphs for the other binders designed showed similar results, meaning that there were no successful CBP binders. The other graphs are in the supplementary information.

When looking closer into the construction of the S100B binders and modelling them with AlphaFold 3, it was seen why they did not bind.

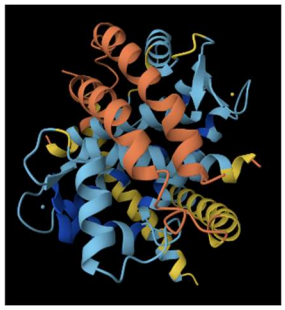
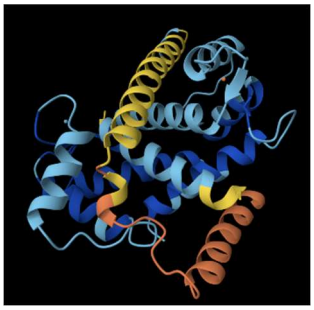
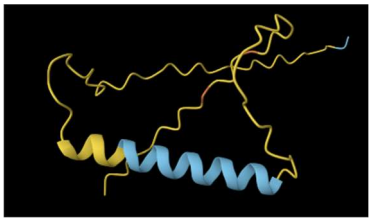
Two Binders	One Binder	Binder Itself
		

Table 1: Binder C3

The binder is modelled by itself, and it is shown how it would bind to S100B.

Binder C3 was modelled with one binder, two binders, and the binder itself. Alpha Fold 3 had low prediction scores for this binder. The binder by itself looks unstable and probably won't be able to fold. The model with one binder was reaching over to both binding sites and not folding into an actual feasible strcuture. The model with two binders was both wrapped around S100B and would not be able to bind.

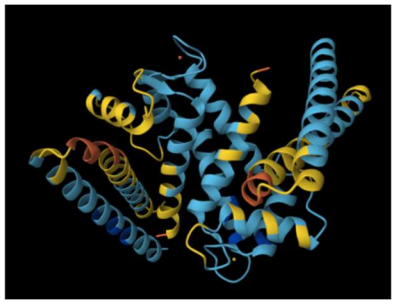
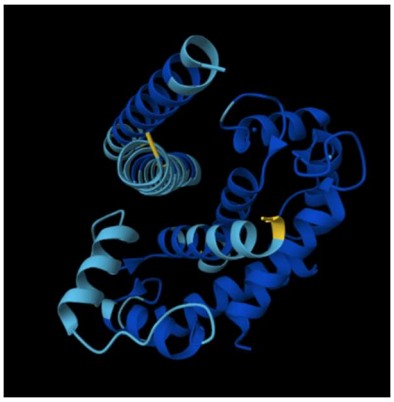
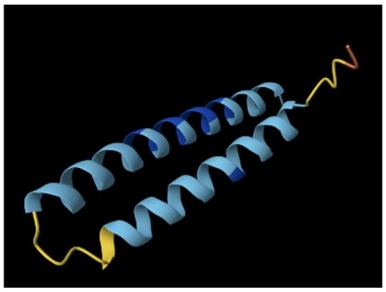
Two Binders	One Binder	Binder Itself
		

Table 2: Binder C4

The binder is modelled by itself, and it is shown how it would bind to S100B.

Binder C4 was modelled in three ways in Alpha Fold 3. It was found that there was some confidence for the fold of the binder itself, but it only has two helices and is long, which is not desirable for a binder. The prediction with one binder looks like it maintains the structure and could possibly bind to the binding site. The prediction with two binders has lower confidence and the binders lose their fold in this prediction.

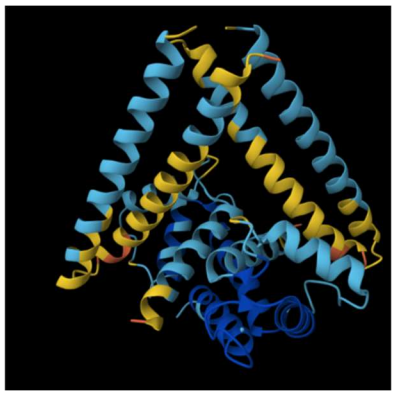
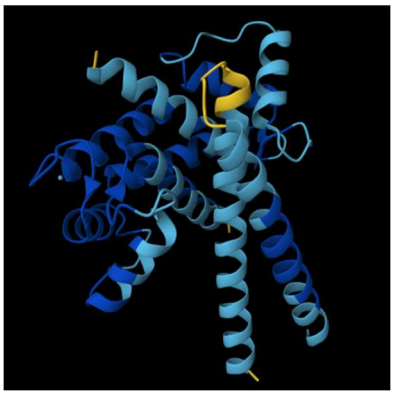
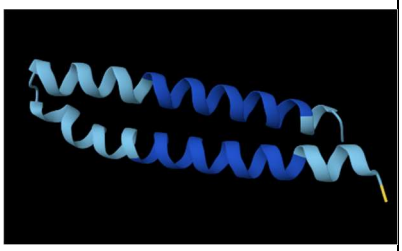
Two Binders	One Binder	Binder Itself
		

Table 3: Binder C5

The binder is modelled by itself, and it is shown how it would bind to S100B.

Binder C5 was predicted by itself to be two helices. The prediction of one binder and target protein in the complex showed that the binder would bind close to the binding site but still be in the two-helix formation. The prediction of two binders and the S100B dimer showed that the binders might come into contact because they are so long which would not allow good binding. Additionally, the two-helix conformation is not desirable for binding.

Two Binders	One Binder	Binder Itself

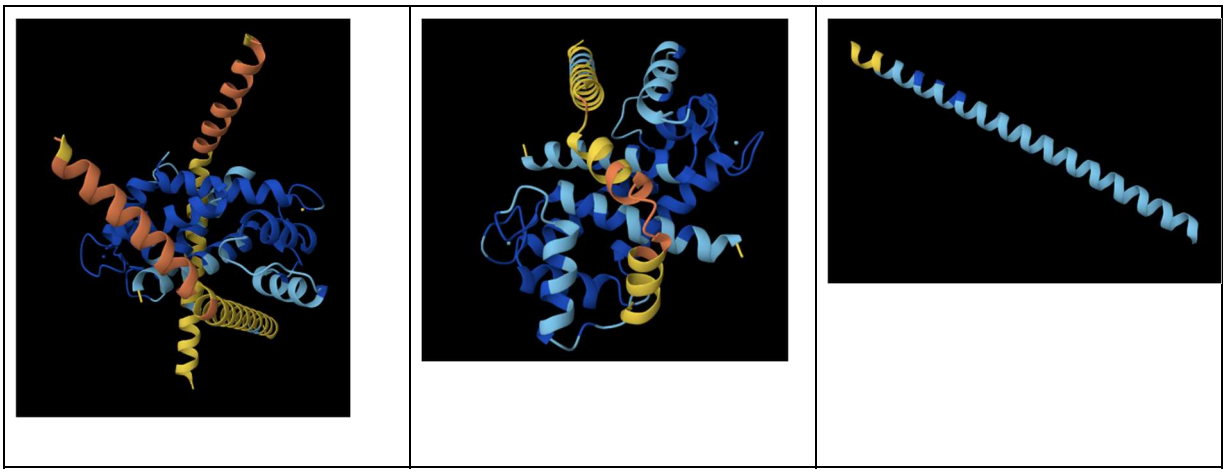


Table 4: Binder C6

The binder is modelled by itself, and it is shown how it would bind to S100B.

Binder C6 was predicted to be a helix strand which is not an actual protein. When modelled bound to the protein with one and two binders, it maintains its form. C6 was predicted to wrap around S100B which would not be a stable structure of a binder. It was predicted to bind to the desired binding site of S100B and wrap around to both binding sites of the dimer. The prediction also has low confidence of the fold. This is probably why it did not work.

By observing the binders designed for CBP, the designed binders had similar reasons for most likely not working based on the modelling. SIRT-1 has yet to be successfully purified so it will be experimented on later. SIRT-1 has around 20 binders designed that have been ordered and are ready for testing.

Discussion

To create inhibitors for S100B, CBP, and SIRT-1 to understand how the P53 pathway changes during a disease state, a multitude of proteins were computationally designed. Through the computational design process, the steps were optimized to create a binder for S100B, and a couple of successful binders were ordered. The optimization consisted of changing the weights for the radius of gyration, interface contacts, and denoiser scales in RFDiffusion. The confidence of the binders was determined by multiple metrics through AlphaFold 2 and ESM Fold. The binders that passed the metrics were ordered for testing in wetlab. This was done for all three target proteins. These binders were expressed and purified along with the target proteins. The testing consisted of BLI and FP. From the testing, it was seen that none of the binders for S100B and CBP were successful. Through further analysis of the binders, it was seen that they did not work due to unstable folds, low confidence binding, and interaction with each other after binding to the binding domains. Further work will be needed to create binders that will be successful and bind to their target proteins.

Future Directions

Due to the previously designed binders not working, the computational design process will be continued and optimized further. Using the optimized potentials of interface contacts, radius of gyration, and denoiser scales, more binders will be designed. The design process will take three approaches. The first approach uses the hot spots found from the PDB bank and paper and runs RFDiffusion using the optimized potentials. This way, thousands of designs will be created, given a sequence using Protein MPNN, and given metrics using ESM Fold, AlphaFold 2, AlphaFold 3. The designs will be docked using AlphaRED. The second approach uses the motif of the C-terminal Disordered region of P53 to start designing off of. This was the approach used for the previous binder design and is called motif grafting. Thousands of binders will be generated and tested using this method and modelled using the same process of the first method. The third method will be partial diffusion. Partial diffusion will take already generated designs that look promising and noise and denoise the design to further optimize them. They will then be tested in the same way as the other methods. Those are the main three methods that will be the focus of the future directions of the S100B and CBP binder design process. The fourth possible method will be generating small helices to different sections of the binding region of the target protein and connecting them by beta loops, creating a protein.

For SIRT-1, the designs have been ordered and will be tested using BLI. There is a BLI method that uses the supernatant directly after expression and does not need to undergo the purification step. That will be used for testing the SIRT-1 binders. If the designs aren't successful, more computational optimization will occur.

Once three inhibitors have been successfully designed and tested, they will be used to see how removing one of the three target proteins from the pathway changes the disease state.

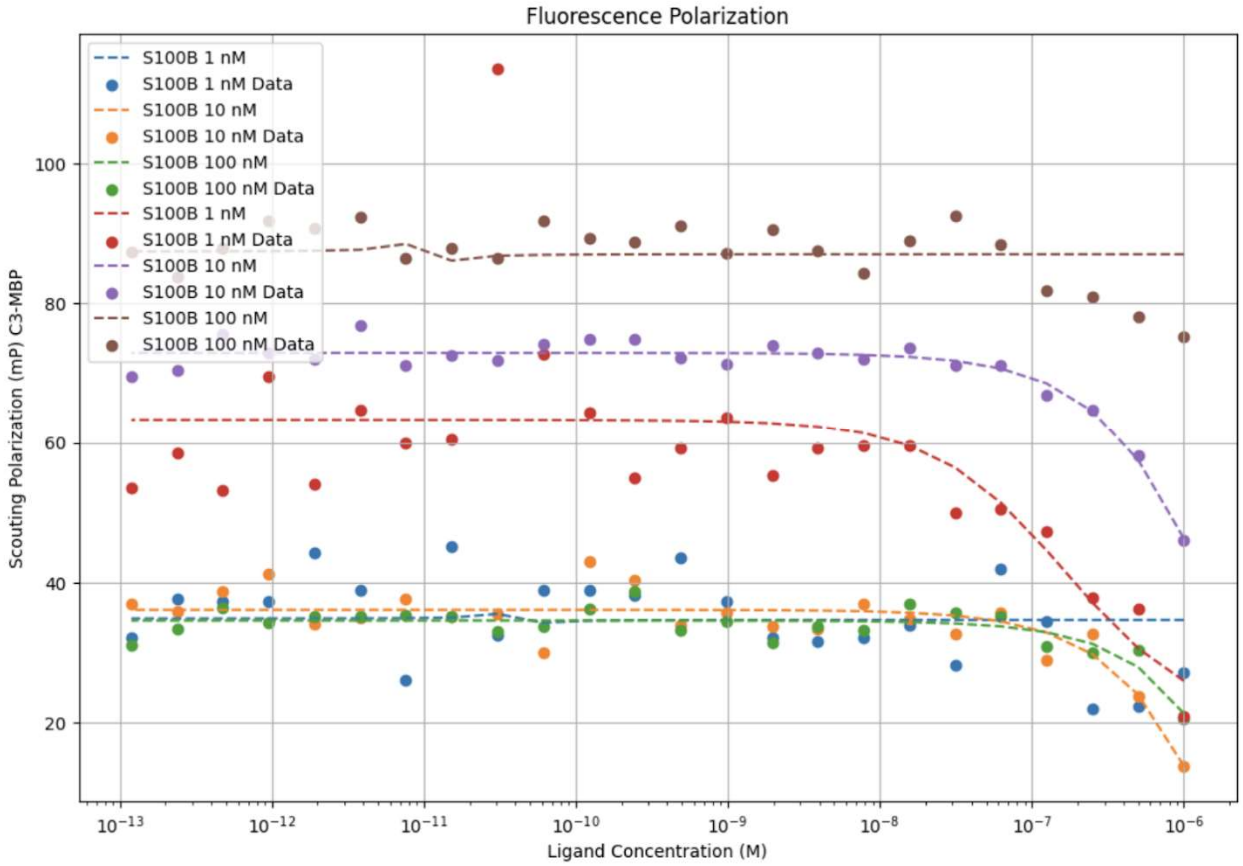
Conclusion

This thesis explored the creation of computationally designed inhibitors to understand protein-protein interaction pathways during disease states. Inhibitors were designed for three target proteins that are a part of the p53 pathway. These three proteins all bind to the c-terminal disordered region of p53. The computational design of S100B inhibitors was explored. Designed inhibitors to S100B and CBP were tested in wet lab to understand if any of the inhibitors would bind. It was found that none of the inhibitors bound, and further computational optimization will be explored. The goal is to design an inhibitor to all three of the proteins to test the kinetics of the p53 pathway.

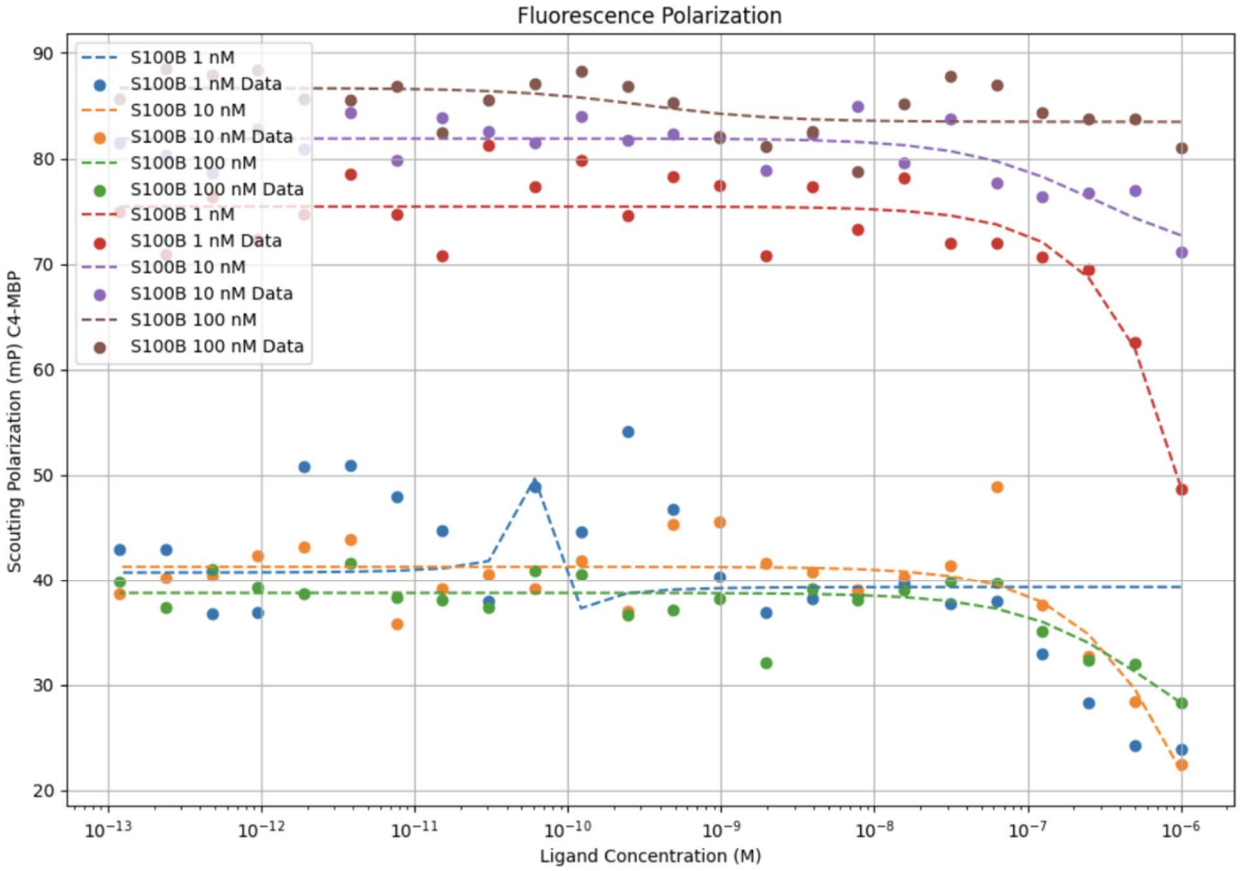
Supplementary Information

All the FP data:

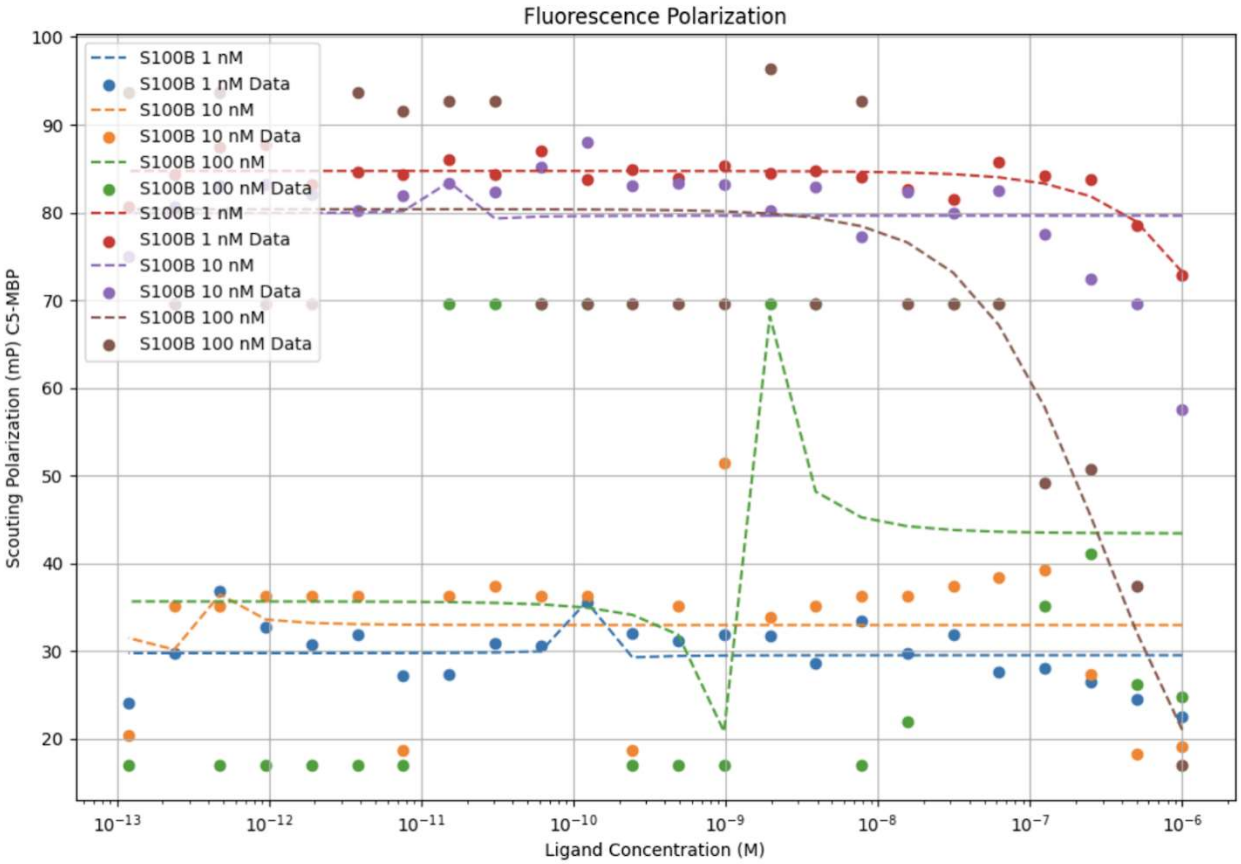
C3:



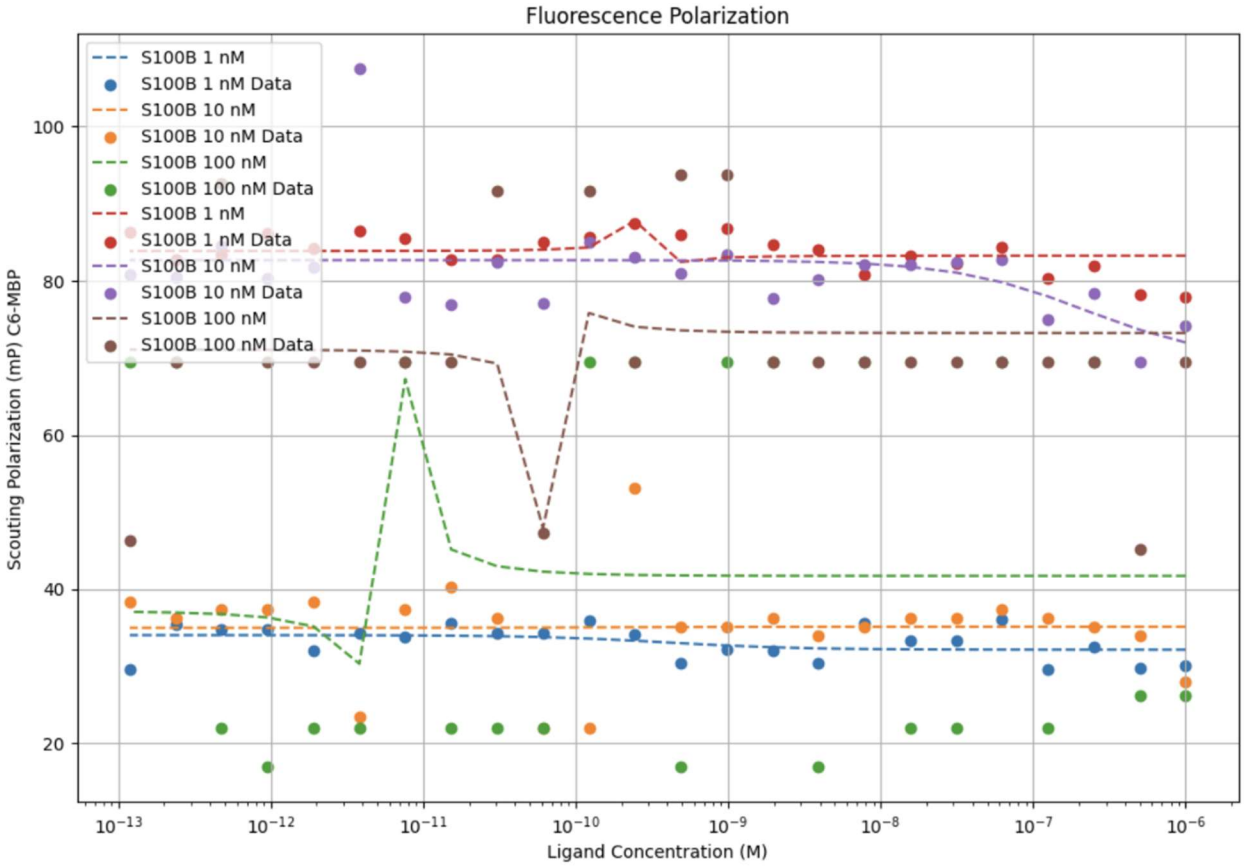
C4:


Fluorescence Polarization
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
Scouting Polarization (mP) C4-MBP
Ligand Concentration (M)
90
80
70
60
50
40
30
20
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}

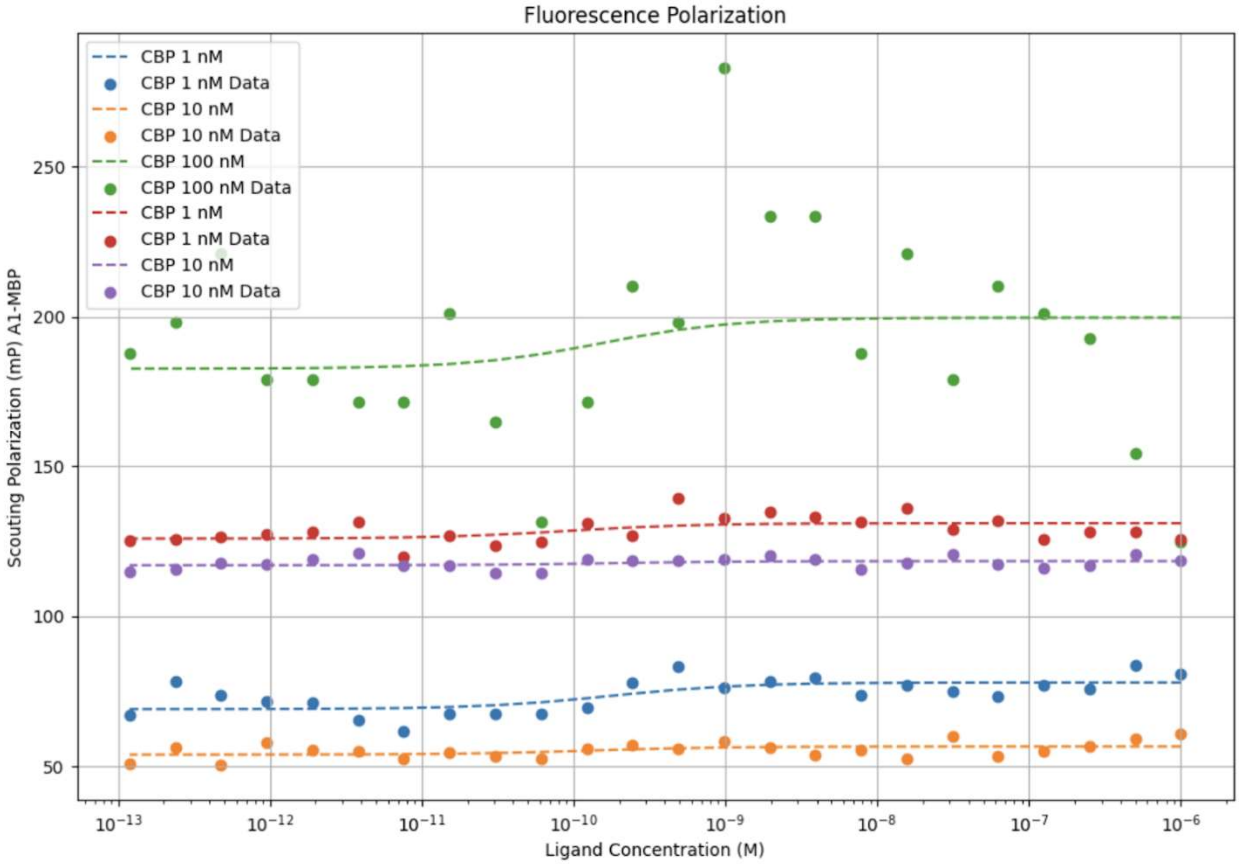
C5:


Fluorescence Polarization
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
Scouting Polarization (mP) C5-MBP
Ligand Concentration (M)

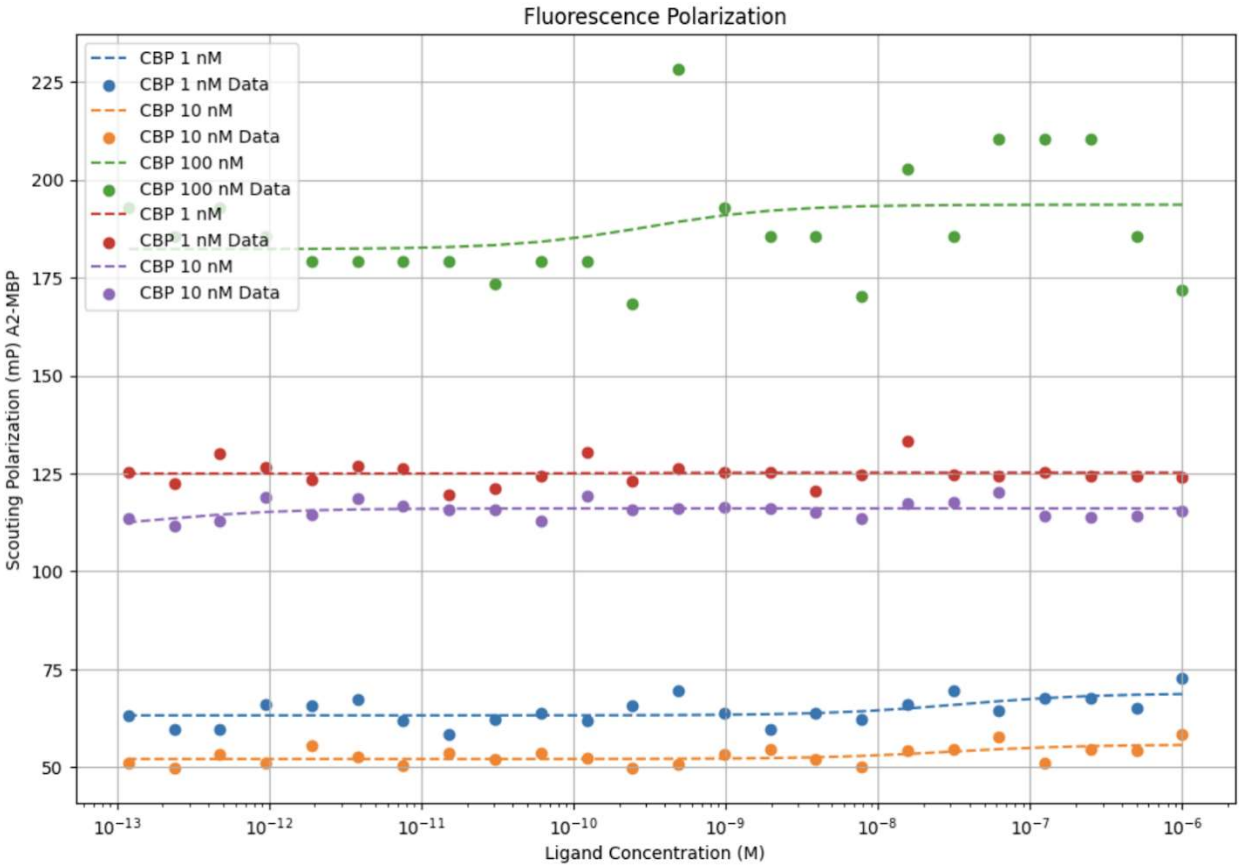
C6:


Fluorescence Polarization
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
S100B 1 nM
S100B 1 nM Data
S100B 10 nM
S100B 10 nM Data
S100B 100 nM
S100B 100 nM Data
Scouting Polarization (mP) C6-MBP
Ligand Concentration (M)

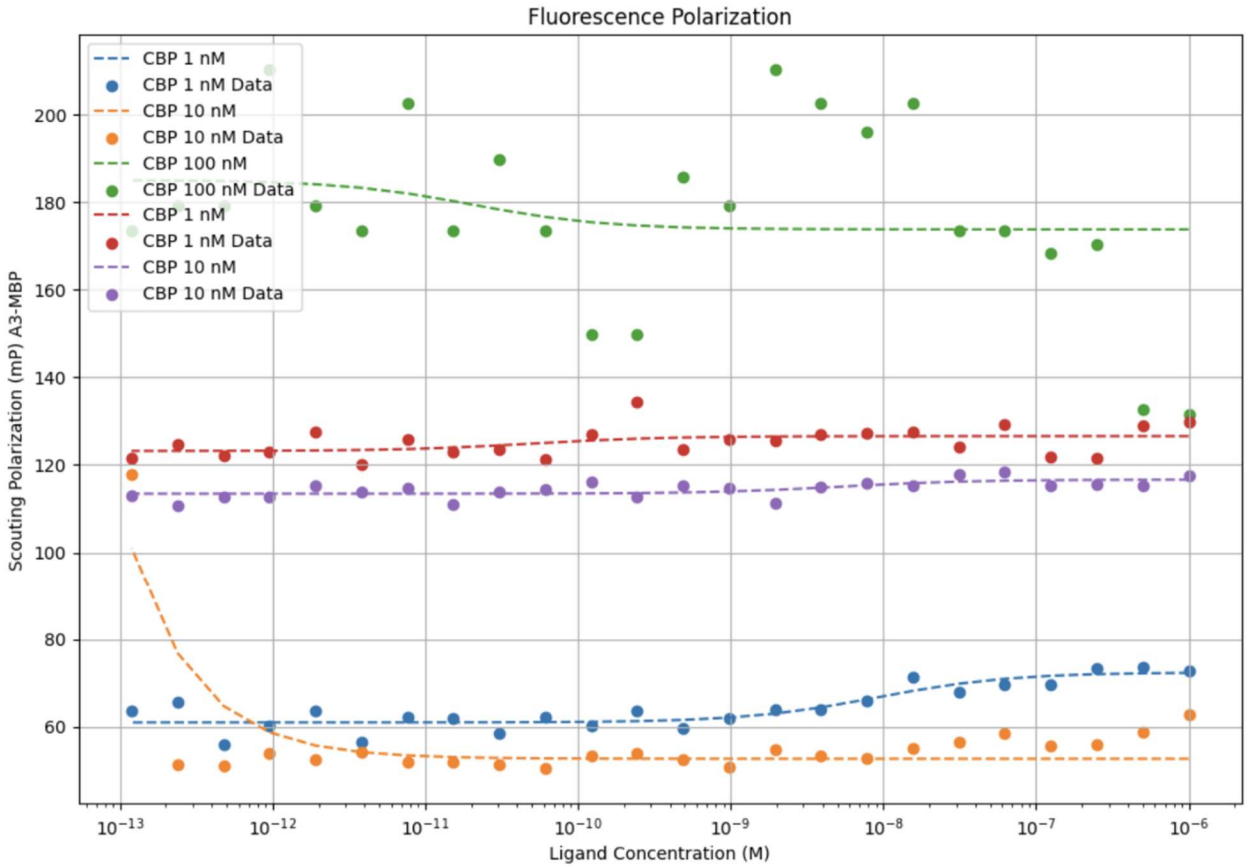
A1:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
Scouting Polarization (mP) A1-MBP
Ligand Concentration (M)
250
200
150
100
50
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}

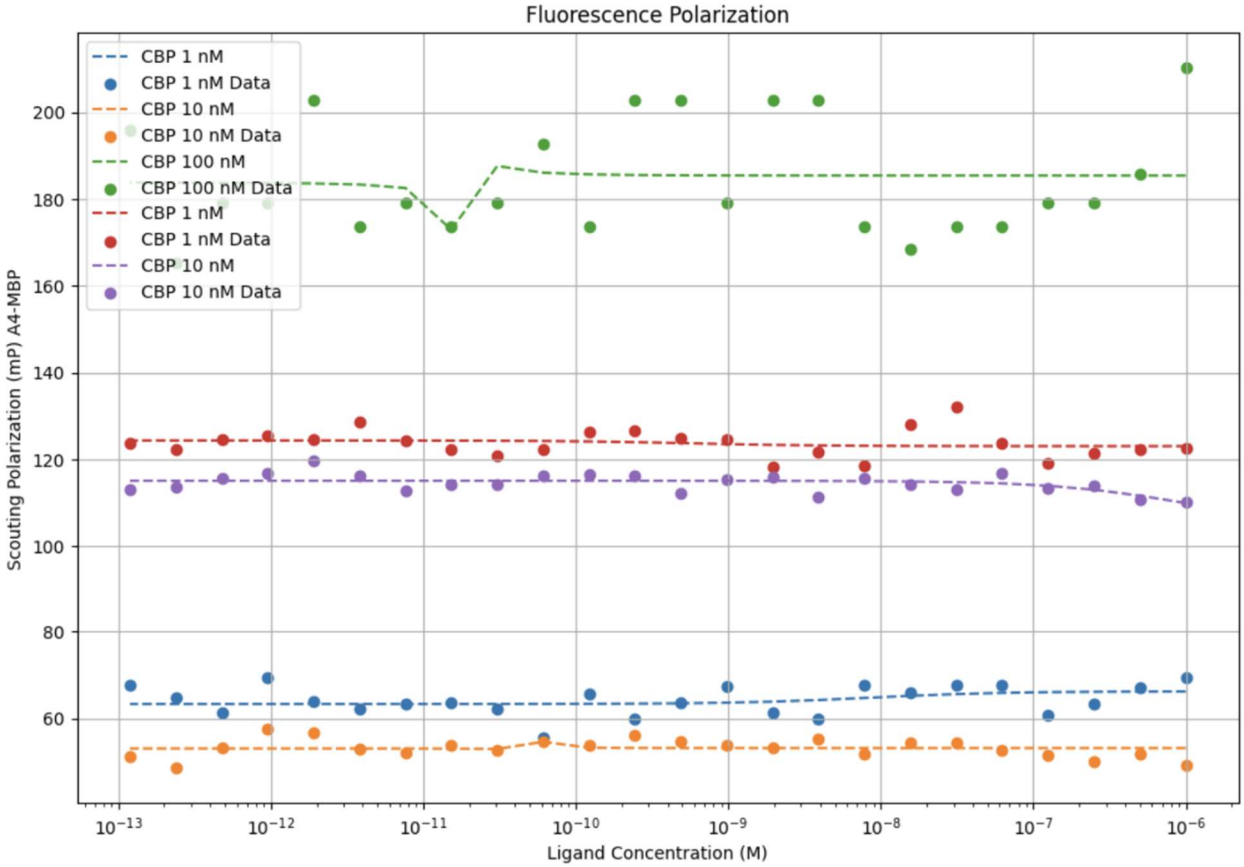
A2:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
Scouting Polarization (mP) A2-MBP
Ligand Concentration (M)
225
200
175
150
125
100
75
50
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}

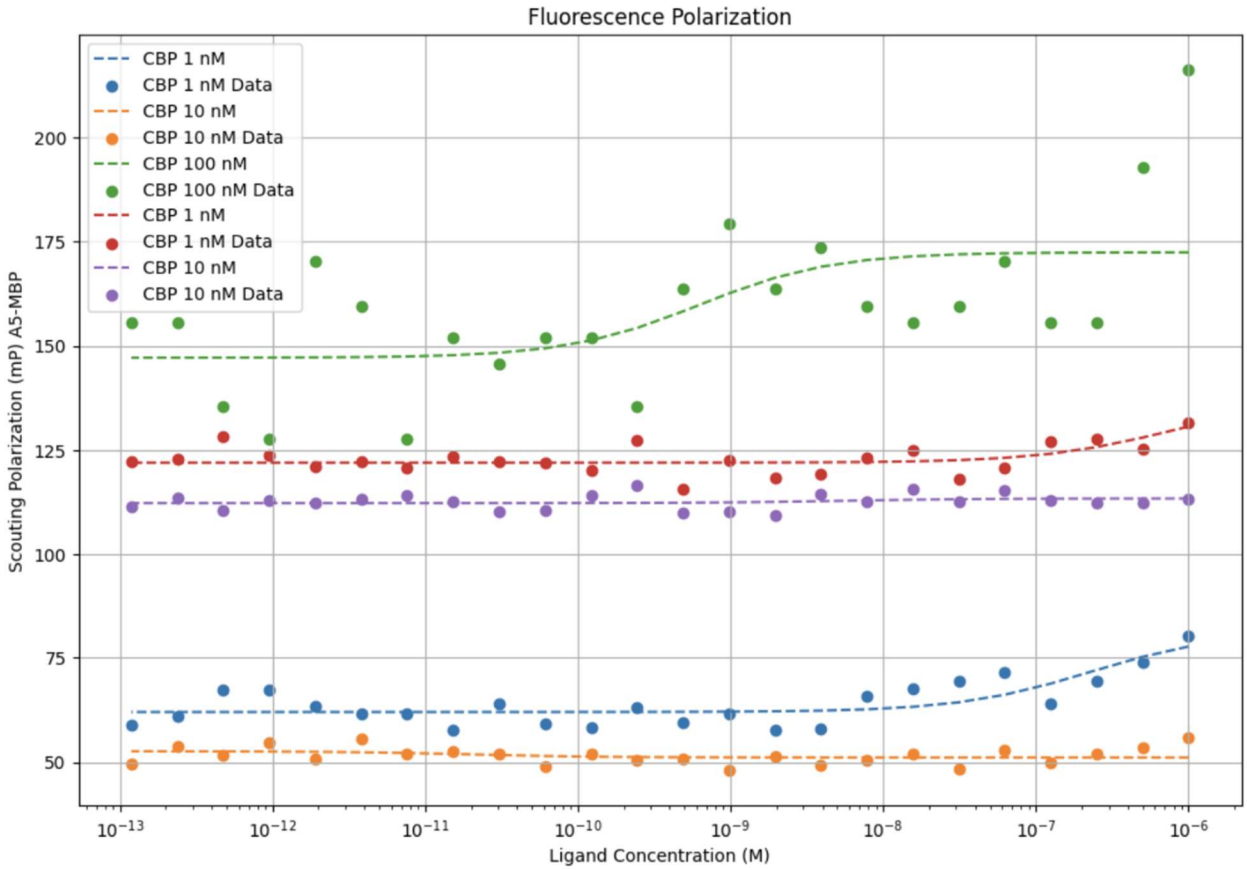
A3:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
Scouting Polarization (mP) A3-MBP
Ligand Concentration (M)

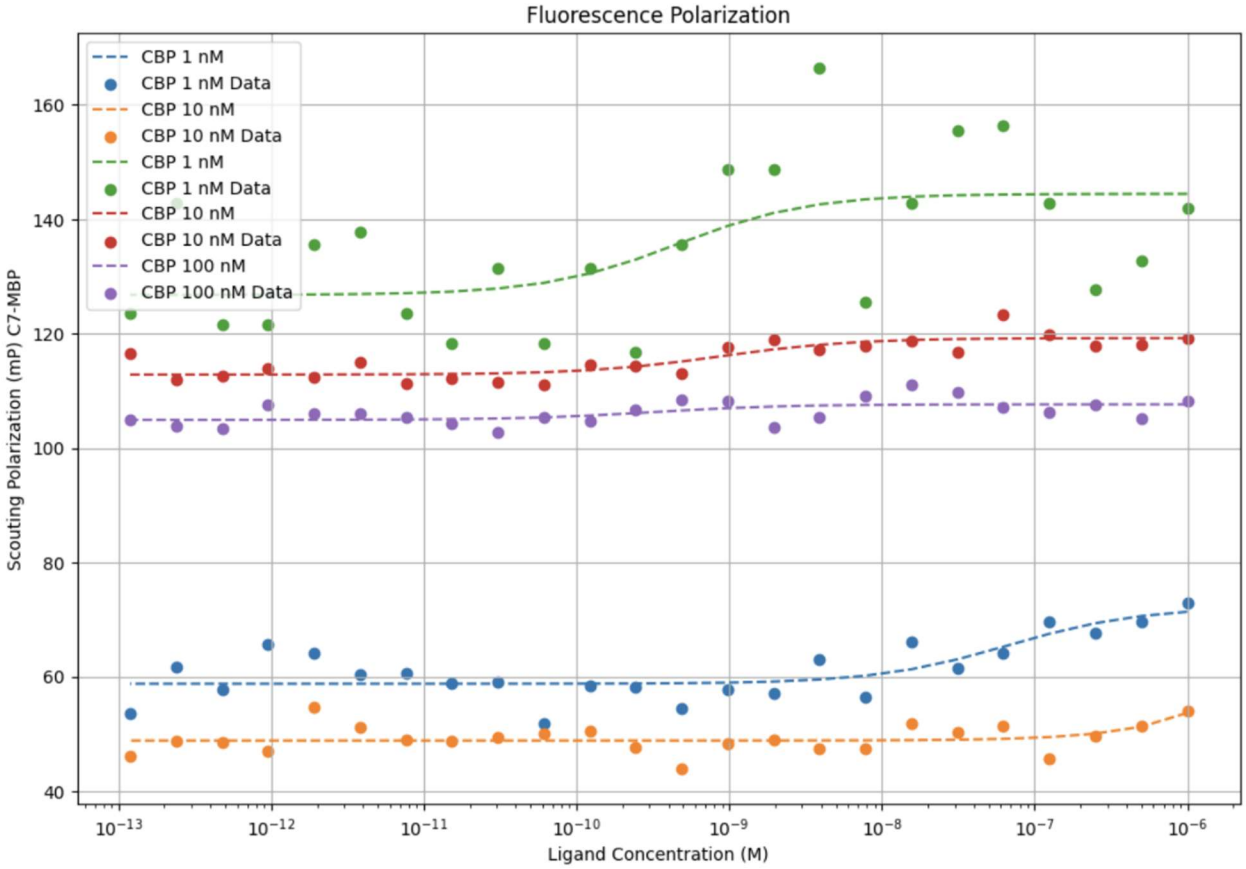
A4:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
Scouting Polarization (mP) A4-MBP
Ligand Concentration (M)

A5:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
Scouting Polarization (mP) A5-MBP
200
175
150
125
100
75
50
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
Ligand Concentration (M)

C7:


Fluorescence Polarization
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 1 nM
CBP 1 nM Data
CBP 10 nM
CBP 10 nM Data
CBP 100 nM
CBP 100 nM Data
Scouting Polarization (mP) C7-MBP
160
140
120
100
80
60
40
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
Ligand Concentration (M)

Bibliography

References:

(1) What is a cell?: MedlinePlus Genetics. https://medlineplus.gov/genetics/understanding/basics/cell/ (accessed 2025-04-26).

(2) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Analyzing Protein Structure and Function. In Molecular Biology of the Cell. 4th edition; Garland Science, 2002.

(3) LaPelusa, A.; Kaushik, R. Physiology, Proteins. In StatPearls; StatPearls Publishing: Treasure Island (FL), 2025.

(4) Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J. Recent Advances in the Development of Protein–Protein Interactions Modulators: Mechanisms and Clinical Trials. Signal Transduct. Target. Ther. 2020, 5 (1), 1–23. https://doi.org/10.1038/s41392-020-00315-3.

(5) Mabonga, L.; Kappo, A. P. Protein-Protein Interaction Modulators: Advances, Successes and Remaining Challenges. Biophys. Rev. 2019, 11 (4), 559–581. https://doi.org/10.1007/s12551-019-00570-x.

(6) Farooq, Q. ul A.; Shaukat, Z.; Aiman, S.; Li, C.-H. Protein-Protein Interactions: Methods, Databases, and Applications in Virus-Host Study. World J. Virol. 2021, 10 (6), 288–300. https://doi.org/10.5501/wjv.v10.i6.288.

(7) Tsai, C.-J.; Ma, B.; Nussinov, R. Protein–Protein Interaction Networks: How Can a Hub Protein Bind so Many Different Partners? Trends Biochem. Sci. 2009, 34 (12), 594–600. https://doi.org/10.1016/j.tibs.2009.07.007.

(8) Clark, D. P.; Pazdernik, N. J. Chapter 20 - Aging and Apoptosis. In Biotechnology (Second Edition); Clark, D. P., Pazdernik, N. J., Eds.; Academic Cell: Boston, 2016; pp 627–661. https://doi.org/10.1016/B978-0-12-385015-7.00020-X.

(9) Lin, J.; Yang, Q.; Wilder, P. T.; Carrier, F.; Weber, D. J. The Calcium-Binding Protein S100B Down-Regulates P53 and Apoptosis in Malignant Melanoma. J. Biol. Chem. 2010, 285 (35), 27487–27498. https://doi.org/10.1074/jbc.M110.155382.

(10) Kumar, A.; Kumar, P.; Kumari, S.; Uversky, V. N.; Giri, R. Folding and Structural Polymorphism of P53 C-Terminal Domain: One Peptide with Many Conformations. Arch. Biochem. Biophys. 2020, 684, 108342. https://doi.org/10.1016/j.abb.2020.108342.

(11) Laptenko, O.; Shiff, I.; Freed-Pastor, W.; Zupnik, A.; Mattia, M.; Freulich, E.; Shamir, I.; Kadouri, N.; Kahn, T.; Manfredi, J.; Simon, I.; Prives, C. The P53 C-Terminus Controls Site-Specific DNA Binding and Promotes Structural Changes within the Central DNA Binding Domain. Mol. Cell 2015, 57 (6), 1034–1046. https://doi.org/10.1016/j.molcel.2015.02.015.

(12) Zimmer, D. B.; Weber, D. J. The Calcium-Dependent Interaction of S100B with Its Protein Targets. Cardiovasc. Psychiatry Neurol. 2010, 2010, 728052. https://doi.org/10.1155/2010/728052.

(13) Designing Specific Inhibitors to Target S100B in Melanoma. https://archive.hshsl.umaryland.edu/entities/publication/812c684c-9b71-4791-93df-2a87ea5c48c4 (accessed 2025-04-04).

(14) Giebler, H. A.; Lemasson, I.; Nyborg, J. K. P53 Recruitment of CREB Binding Protein Mediated through Phosphorylated CREB: A Novel Pathway of Tumor Suppressor Regulation. Mol. Cell. Biol. 2000, 20 (13), 4849–4858. https://doi.org/10.1128/mcb.20.13.4849-4858.2000.

(15) Lee, C. W.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. Structure of the P53 Transactivation Domain in Complex with the Nuclear Receptor Coactivator Binding Domain of CREB Binding Protein. Biochemistry 2010, 49 (46), 9964–9971. https://doi.org/10.1021/bi1012996.

(16) Lambert, P. F.; Kashanchi, F.; Radonovich, M. F.; Shiekhattar, R.; Brady, J. N. Phosphorylation of P53 Serine 15 Increases Interaction with CBP*. J. Biol. Chem. 1998, 273 (49), 33048–33053. https://doi.org/10.1074/jbc.273.49.33048.

(17) Peck, B.; Chen, C.-Y.; Ho, K.-K.; Di Fruscia, P.; Myatt, S. S.; Coombes, R. C.; Fuchter, M. J.; Hsiao, C.-D.; Lam, E. W.-F. SIRT Inhibitors Induce Cell Death and P53 Acetylation through Targeting Both SIRT1 and SIRT2. Mol. Cancer Ther. 2010, 9 (4), 844–855. https://doi.org/10.1158/1535-7163.MCT-09-0971.

(18) Li, P.; Zhang, L.; Zhou, C.; Lin, N.; Liu, A. Sirt 1 Activator Inhibits the AGE-Induced Apoptosis and P53 Acetylation in Human Vascular Endothelial Cells. J. Toxicol. Sci. 2015, 40 (5), 615–624. https://doi.org/10.2131/jts.40.615.

(19) Yan, C.; Wu, F.; Jernigan, R. L.; Dobbs, D.; Honavar, V. Characterization of Protein–Protein Interfaces. Protein J. 2008, 27 (1), 59–70. https://doi.org/10.1007/s10930-007-9108-x.

(20) Rustandi, R.; Baldisseri, D.; Weber, D. Structure of the Negative Regulatory Domain of P53 Bound to S100B(Bβ). Nat. Struct. Biol. 2000, 7, 570–574. https://doi.org/10.1038/76797.

(21) Dahiyat, B. I.; Mayo, S. L. De Novo Protein Design: Fully Automated Sequence Selection. Science 1997, 278 (5335), 82–87. https://doi.org/10.1126/science.278.5335.82.

(22) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. Science 2003, 302 (5649), 1364–1368. https://doi.org/10.1126/science.1089427.

(23) Simons, K. T.; Bonneau, R.; Ruczinski, I.; Baker, D. Ab Initio Protein Structure Prediction of CASP III Targets Using ROSETTA. Proteins 1999, Suppl 3, 171–176. https://doi.org/10.1002/(sici)1097-0134(1999)37:3+<171::aid-prot21>3.3.co;2-q.

(24) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. De Novo Design of Protein Structure and Function with RFdiffusion. Nature 2023, 620 (7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8.

(25) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust Deep Learning–Based Protein Sequence Design Using ProteinMPNN. Science 2022, 378 (6615), 49–56. https://doi.org/10.1126/science.add2187.

(26) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. Nature 2021, 596 (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(27) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. Science 2023, 379 (6637), 1123–1130. https://doi.org/10.1126/science.ade2574.

(28) Harmalkar, A.; Lyskov, S.; Gray, J. J. Reliable Protein-Protein Docking with AlphaFold, Rosetta, and Replica-Exchange. bioRxiv January 2, 2025, p 2023.07.28.551063. https://doi.org/10.1101/2023.07.28.551063.

(29) Goudy, O. J.; Nallathambi, A.; Kinjo, T.; Randolph, N. Z.; Kuhlman, B. In Silico Evolution of Autoinhibitory Domains for a PD-L1 Antagonist Using Deep Learning Models. Proc. Natl. Acad. Sci. 2023, 120 (49), e2307371120. https://doi.org/10.1073/pnas.2307371120.

(30) relax — PyRosetta 4.0 documentation. https://graylab.jhu.edu/PyRosetta.documentation/pyrosetta.rosetta.protocols.relax.html (accessed 2025-05-16).

(31) Liu, K.; Mao, W.; Shen, S.; Jiao, X.; Sun, Z.; Chen, H.; Shen, C. Floating Anchor Diffusion Model for Multi-Motif Scaffolding. In Proceedings of the 41st International Conference on Machine Learning; PMLR, 2024; pp 31691–31708.

(32) Guo, H.-B.; Perminov, A.; Bekele, S.; Kedziora, G.; Farajollahi, S.; Varaljay, V.; Hinkle, K.; Molinero, V.; Meister, K.; Hung, C.; Dennis, P.; Kelley-Loughnane, N.; Berry, R. AlphaFold2 Models Indicate That Protein Sequence Determines Both Structure and Dynamics. Sci. Rep. 2022, 12, 10696. https://doi.org/10.1038/s41598-022-14382-9.

(33) Bates, T. A.; Gurmessa, S. K.; Weinstein, J. B.; Trank-Greene, M.; Wrynla, X. H.; Anastas, A.; Anley, T. W.; Hinchliff, A.; Shinde, U.; Burke, J. E.; Tafesse, F. G. Biolayer Interferometry for Measuring the Kinetics of Protein–Protein Interactions and Nanobody Binding. Nat. Protoc. 2025, 20 (4), 861–883. https://doi.org/10.1038/s41596-024-01079-8.

(34) Rossi, A. M.; Taylor, C. W. Analysis of Protein-Ligand Interactions by Fluorescence Polarization. Nat. Protoc. 2011, 6 (3), 365–387. https://doi.org/10.1038/nprot.2011.305.

(35) RFDiffusion Potentials. https://huggingface.co/blog/AmelieSchreiber/rfdiffusion-potentials (accessed 2025-05-16).