The Analysis of Folate-Dependent Transcription Factor Zinc Finger Protein 410

Feifan Xu

A Senior Thesis submitted in partial fulfillment
of the requirements for graduation
in the Honors Program
Liberty University
Spring 2021

Acceptance of Senior Honors Thesis

This Senior Honors Thesis is accepted in partial
fulfillment of the requirements for graduation from the
Honors Program of Liberty University.

_____
Gary D. Isaacs, Ph.D.
Thesis Chair

_____
Mathew Becker, Ph.D.
Committee Member

_____
Cynthia Goodrich, EdD
Assistant Honors Director

_____
Date

Abstract

A previous study that introduced dietary folate to mice in the form of folic acid to determine if gene activity would be altered based on this biological molecule demonstrated that mice without folic acid had cognition deficits, and this phenomenon was correlated with altered gene expression in their brains. The included bioinformatic analysis revealed two main transcription factors that bind to proteins in the nucleus, and one is known as the Zinc Finger Protein 410 (Zfp410). Due to lack of literature explaining the function of this transcription factor, this project is intended to analyze Zfp410 in detail from scratch. Zfp410 cDNA sequence was purchased along with a plasmid vector, transferred to *E.* coli, and used to purify the recombinant protein. The purified protein was then injected into a rabbit for antibody production, and the collected Zfp410 antibody was characterized using several molecular techniques. Within high efficiency of targeting its antigen, Zfp410 antibody was used to identify any interacting proteins of Zfp410 by co-immunoprecipitation along with silver stain and western blot. The results demonstrated several Zfp410 interacting partners which suggested transcription factor Zfp410's role in folate dependent genes' regulations.

Keywords: recombinant protein purification, antibody, co-immunoprecipitation, western blot, folate, gene expression, silver stain

The Analysis of Folate-Dependent Transcription Factor Zinc Finger Protein 410

1. Background

Transcription factors play critical roles in regulating genetic information. By acting as switches, these proteins are capable of activating or disabling DNA transcription, which is the process of turning DNA sequences into RNA, the core genetic material for providing the instruction of protein synthesis. Due to its importance in controlling different gene expressions within cells, transcription factors are highly regulated throughout the cell cycle so that specific genetic information can be expressed at the right moment and the right amount (Latchman, 1997). Furthermore, individual genes are controlled by specific transcription factors, and groups of transcription factors together can control various cell activities such as cell division, cell growth, migration, and apoptosis (programmed cell death).

The mechanism behind the gene regulation function of transcription factors is quite complicated, but the general idea is that transcription factors tend to cooperate with other proteins in a complex and bind to a specific DNA sequence in order to upregulate or downregulate its transcription (Narlikar et al., 2002). There are various factors that can influence the efficiency of transcription factors such as the tightness of a DNA sequence in the form of chromatin and the number or arrangement of transcription factor-binding sites in the regulatory region of variants, also the availability of coactivating proteins (Lillycrop et al., 2005). Like other proteins and enzymes, the majority of transcription factors require cofactors to fully function during transcription process (Robertson et al., 2000). Most of those cofactors are not directly made in human body but rather supplied from daily diet, and folate is one of those critical cofactors that enables the maintenance of transcription factors' functions.

Folate, or vitamin B-9, is a well-known component that is critical for red blood cell formation and brain development. Folate can exist in several ways, and the two most common forms are folic acid and 5-methyltetrahydrofolate. Since folate can transfer one carbon atom required for purine and thymidylate syntheses and assist methylation of different types of essential biological substances like nucleotides, phospholipids, DNA, and neurotransmitters, it serves important role in stabilizing DNA (Robert, 2000). Because of the inverse relationship between gene methylation and gene expression, the maintenance of proper folate levels is extremely important for preventing abnormal DNA and histone methylation patterns and ensure regular control of transcription factors (Cooney et al., 2002). In other words, insufficient folate intake can lead to altered gene expression and nucleic acid metabolism, and this can lead to various folate deficiency disorders such as anemia and cognitive defects. In additions, low folate status can also lead to the elevation of plasma homocysteine which is associated with cardiovascular disease (Selhub et al., 2009).

According to a previous research study started in Spring 2015 funded by the Department of Biology and Chemistry at Liberty University, scholars have confirmed that deficient dietary folate in the form of folic acid could lead to cognitive deficits in mice, and this outcome was proven to be linked with altered hippocampal gene expressions (Lawton et al., 2020). The study investigated the genetic effects of folate deficiency of the adult mice throughout their lives by using a novel object test. Two groups of mice were used for comparison: one group was nourished with a regular diet that contained folic acid and the other group was not. The result demonstrated significant memory impairment in folate deficient group compared to the control group. In additions, hippocampal tissue samples from each group were obtained for gene expression analysis after 6, 12, and 18 months respectively, and a microarray assay was followed

afterwards. A total of 363 significantly downregulated and 101 upregulated genes were determined from the folate deficient group due to this cognition impairment association based on gene ontology analysis (Lawton et al., 2020). Since gene expression is highly regulated by transcription factors, the study was then moved on to analyze those altered genes in the folate deficient hippocampal tissues through bioinformatics analysis in order to discover the factors that contribute to this outcome. The result discovered two main transcription factors that serve as enriched motifs of binding sites for the previous identified folate-dependent genes: V-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MafB) and Zinc finger protein 410 (Zfp410) (Figure 1). Although expression levels of these two transcription factors were not significantly different between dietary conditions based on microarray analysis, they were associated with regulating a number of folate-dependent genes expression and suggested possible future treatment's target for folate deficiency related disorders and cognition defects (Lawton et al., 2020).

While MafB is a relatively well-known transcription factor that regulates lineage-specific hematopoiesis (Artner et al., 2007), Zfp410, on the other hand, is rarely studied, and less publications can be found. This might be associated with lack of reagents available for studies such as purified Zfp410 and its polyclonal antibody. Therefore, this project was conducted to seek the specific role of folate-dependent transcription factor Zfp410 and its mechanism for interaction under molecular level. The whole project is divided into the following three major aims: the production and purification of Zfp410 protein, the production and characterization of a polyclonal ZFP410 antibody, and the utilization of Zfp410 antibody to identify its interacting partners.

Motif	Position Weight Matrix	Type	*p*-Value	% Target	% BG	Fold	Match / Detail
1.		*de novo*	1e-12	5.54	0.31	17.9	MafB Homer (0.591)
2.		*de novo*	1e-12	6.27	0.49	12.8	Zfp410 Jaspar (0.643)

Figure 1. Sequence motifs enriched in folate-regulated genes (Lawton et al., 2020).

2. Aim #1: Production & purification of Zfp410

2.1. Overview

Since purified Zfp410 was not available, the first aim of the project was to express and purify the recombinant protein for in depth study. *Escherichia coli* was chosen for producing the recombinant protein since bacteria are easy to grow and can replicate rapidly in a short amount of time (Del Solar, 1998). Since purification of a recombinant protein requires a special tag, the project used a plasmid vector called pET45b (+) that contained a poly-histidine tag attached to the sequence of Zfp410, which can be recognized by Ni-NTA resin. The vector was then transformed into *E. coli* cells so that the recombinant protein can be generated together with bacterial cells replications after IPTG induction. Once the protein was produced and purified, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to verify the preparation outcome.

2.2. Methods:

2.2.1. Zfp410 production:

The cDNA sequence that codes for Zfp410 was purchased along with a plasmid vector pET-45b (+) from Novagen (Figure 2). As shown in figure 2, pET45b (+) encoded an ampicillin resistant gene that allows the survival of the vector, and this is important during later bacterial growth in terms of increasing the purity of Zfp410. In additions, pET45b (+) also had a sequence

encoding a poly-histidine tag that was consisted of six histidine residues attached to Zfp410 cDNA, which enabled the purification of Zfp410 recombinant protein by Ni-NTA resin. A detailed explanation is given in the purification section below.

The modified pET45-45b (+)/Zfp410 was then transformed into a type of *E. coli* cell called BL21 for protein production since *E. coli* is preferred as the protein expression model due to its ease of growth and genetic manipulation (Wingfield, 2015). A total of 0.8 ug of zfp410 plasmid complex (8 µl of 0.1 µg/µl DNA sample) was mixed with 500 µl of BL21 cells and incubated on ice for thirty-minuets to allow interactions between the bacteria and the DNA. The whole mixture was then set in warm water bath at 42 degrees Celsius for 90 seconds in order to heat shock the semi-frozen *E. coli* cells so that the bacterial membrane can be more permeable for DNA molecule's entrance (Gietz, 2014). Without shaking, the complex was then carefully moved on the ice to chill for a few seconds and then incubated in 37 degrees Celsius water bath for one hour. A total of 500 µl of sterilized nutrient media LB broth that was previously prepared was then added as the major food source for BL21 cells to enable the growth. Once the incubation was completed, 250 µl of the mixture was plated on the LB-ampicillin (LB-Amp) plate (10 g tryptone, 5 g yeast extracted, 5 g NaCl, 10 mL of 1M $MgSO_4$, 3.75 mL of 1M NaOH, 100 mg/L ampicillin, water add up to 1L). Ampicillin was added in the plate because it acts as an antibiotic which inhibits the growth of extra bacteria and fungus on the same plate, therefore increases the purity of the protein of interest (Sturød, 2018). Since Zfp410 and the plasmid vector were ampicillin resistance, they can therefore grow on the LB-Amp plate without problems. The plate that contained *E. coli* with Zfp410 was stored in an incubator at 37 degree of Celsius overnight for bacterial growth.

The plate was retrieved after overnight incubation with some visible white-yellow colonies in the middle of plate (Figure 3). Sample of colony was taken with a sterile micropipette tip and placed into a clean 15 mL conical tube that contained 5 mL of fresh LB-Amp solution. The tube was then placed in the shaker for overnight incubation at 37 degree of Celsius at 250 rpm to generate the overnight starter culture. This step was necessary for growing large scale *E. coli*/Zfp410 because the amount of time that leads to sufficient protein expression is significant long if starting from scratch with the colony sample. Right before the second incubation stage for large scale bacterial expression, 750 µl of the 5 mL overnight starter culture was stored into a separate 1.8 mL storage tube with 750 µl of 30% glycerol-LB broth for long-term storage in the -80 °C freezer. The rest of the culture was added into a 300 mL Erlenmeyer flask that holds 200 mL fresh sterilized LB broth with 200 µl of 1000x ampicillin. The similar process was repeated as the previous incubation except that one additional factor was addressed. Due to the condition that sufficient amount of bacterial growth needed to be reached in order to produce enormous amount of Zfp410, the expression level of BL21 cells had to be checked through optical density test before recombinant protein generation. Different from absorbance, optical density measures the refractive medium' ability to slow or delay the transmission of light. Therefore, the stronger the ability of a medium can decrease the light's transmission, the higher the optical density is (Flournoy, 2018). This concept is important for the application in microbiology and molecular biology since optical density can evaluate the quantity of bacteria presented based on the amount of light that can be refracted. The ideal optical density (OD) range of *E. coli* was between 0.4 and 0.6, which means that the incubation time for BL21's growth in the 200 mL LB-Amp flask was between 2.5 hours and 3.5 hours in order to reach the range. The initial OD reading was recorded after 2 hours incubation counted from the addition of overnight starter culture into the

flask. Two milliliters of culture were obtained into a separate clean test tube and set inside the OD reader. The number was lower than 0.4 at that moment, which indicated that the bacterial quantity was insufficient, but the range hit after another 20 to 30 minuets of incubation. Again, the amount of incubation time required for qualifying the optical density test depends on various factors such as the bacterial cells' health condition and the LB-Amp broth's quality.

Once enough BL21 cells were produced, the Zfp410 was then ready to be made in large scale follow the addition of a molecular reagent known as Isopropyl β-D-1-thiogalactopyranoside (IPTG). This compound acted as a "twin" version of another lactose metabolite molecule called allolactose and has the ability of triggering transcription of the *lactose* operon, which is encoded in the pET-45 vector and responsible for the translation to bacterial genes that can enable the bacteria's metabolism of lactose under the situation of deficient glucose level (Marbach & Bettenbrock, 2011). Like some other operons, *lac* operon is normally repressed by a protein that is encoded by *lacl* gene so that the RNA polymerase is not able to bind to the promoter (a DNA binding site that leads to transcription) (Lewis, 2013). The only molecule that can remove this inhibitory protein is the lactose metabolite molecule mentioned above. Because IPTG is the mimic molecule of allolactose, it can therefore bind to *lac* operon and kick off the repressive protein in order to turn on gene transcription, which also indirectly activates protein production. After the addition of 200 μl of 1 M IPTG, the whole culture was incubated in the shaker at 37 °C for two hours to enable the expression of Zfp410. The culture was then spun down at 5200 RPM for 15 minuets afterwards using the JA-14 rotor. A large, dark-creamy color solid pallet was formed at the bottom of the rotor that demonstrated the dense culture (Figure 4). The clear yellow solution on the top was then discarded in the sink since it was simply the used LB-Amp broth.

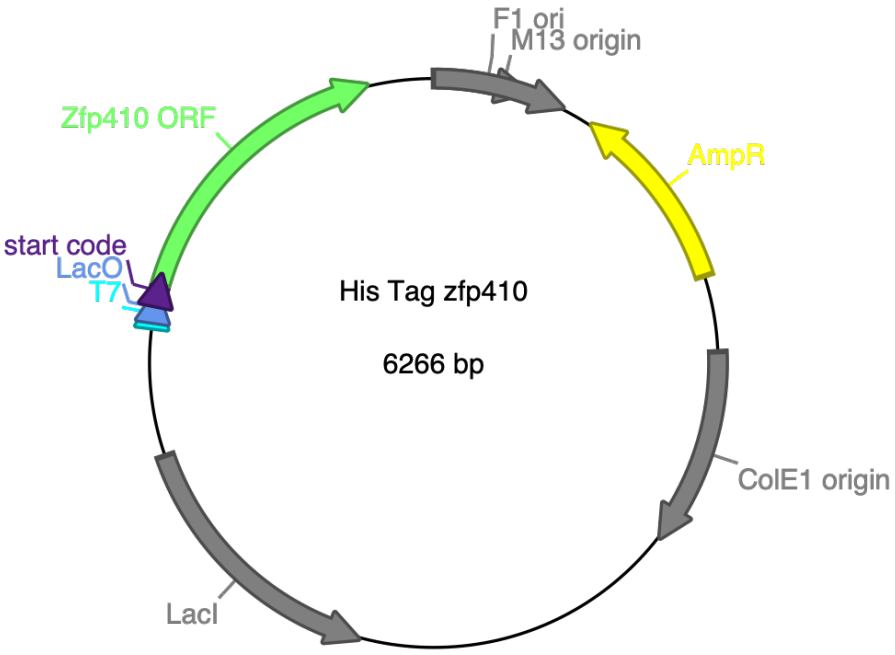

Figure 2. The genomic map of plasmid vector pET-45b (+) that was inserted with Zfp410.

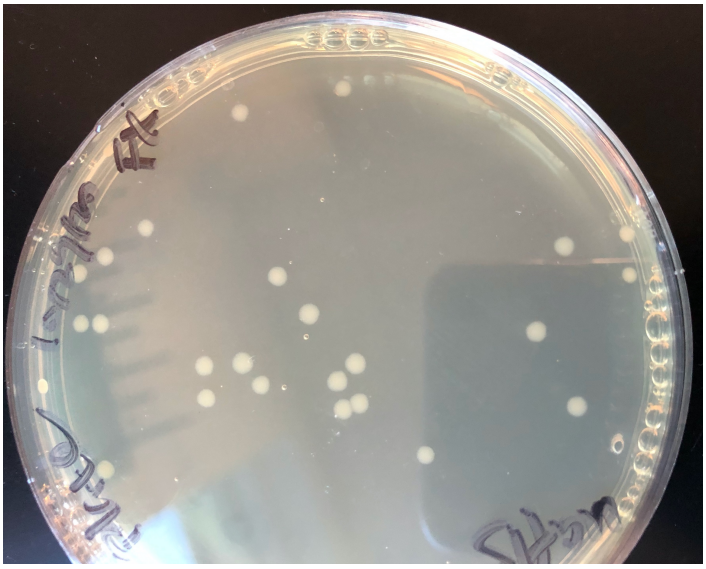

Figure 3. The LB-Amp plate that has *E. coli*/Zfp410 colonies grown on it.

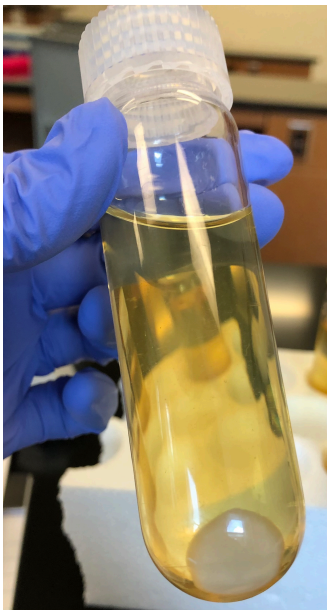

Figure 4. The BL21 culture with Zfp410 pellet inside a JA-14 rotor after the spin.

2.2.2. Zfp410 purification:

Although the bacterial transformation and amplification were successful, the recombinant protein has to be separated from the BL21 cells and other non-specific proteins for the purpose of analysis. As a result, the purification of Zfp410 was introduced as the next major stage of the project. The pellet formed from centrifugation was suspended with 15 mL of lysis buffer (10mM Tris pH 7.5, 0.3M NaCl, 0.5mM EDTA, 10% glycerol, 0.1% Igepal, 23mM Imidazol, water up to 50 mL, add 1mM DTT and one protease inhibitor tablet per mL at time of use) so that the cell membrane could be opened and allowed the release of Zfp410 and other cytoplastic contents. Lysis buffer is often composed with salt, protease inhibitors and mild detergent material like Igepal to rapture bacterial cell membranes (Massiah, Wright, & Du, 2016). The buffer used in the current project also contained 10 mM imidazole because it can remove nonspecific proteins' binding without disturbing the protein of interest, therefore increasing the purity of Zfp410. In order to further rapture the BL21 cells, the lysed culture was then moved to another clean 50 mL conical tube and placed on ice for sonication as it used energy from ultrasonic frequency to

intensively pop open the cell. Sonication is widely used in many fields including extraction of plants, detecting pathogens and even diagnosis of infectious diseases (Bürger, 2019). The process of sonication was repeated 3 times within 45 seconds per each sonication cycle. Zfp410 containing lysate was then spun down using a JA 25.50 rotor (smaller volume compare to JA-14) for 10 minuets at 10,000 RPM. The resulting pellet was a little smaller than the previous one before lysis and sonication but appeared as similar physical appearance. The top clear creamy color solution ideally contained the majority non-purified soluble Zfp410 while the pellet was just the left-over *E. coli* cell components; however, the pellet was kept in the freezer for the purpose of safety in case that some Zfp410 were still insoluble and left in the pellet. This step was actually proved to be really necessary later, which will be discussed in the next section.

Since the supernatant part of Zfp410 was still in its unpurified stage, Ni-NTA resin was used to capture Zfp410 out from the pool by targeting the attached poly-histidine tag. Ni-NTA resin, or Nickel-nitrilotriacetic acid, functions as chelator agarose that forms bonding of ions to metal ions such as nickel and cobalt (Spriestersbach, 2015). The complex is able to form a strong micromolar affinity with poly-histidine tagged proteins, which is any protein that has at least six histidine residues in it. Based on the fact that plasmid vector pET45 b (+) encoded a sequence of poly-histidine tag that attached to the Zfp410 cDNA, this implied that the Zfp410 can be easily purified using Ni-NTA resin as mentioned earlier (Figure 5). After adding 200 µl of the Ni-NTA resin, the whole solution was set on a rotator in a cold fridge room for two hours at 4°C to allow fully interaction between the resin and Zfp410. A quick centrifuge was followed by the end of incubation to spin down the resin/Zfp410, and the top solution that came through the resin was removed to a separate tube and labeled for storage. The purpose of this step was also making sure to prevent loss of any Zfp410 since the resin could reach its maximum binding capacity, and

a small aliquot (15 µl) of this "flow through" solution was kept for future sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis in order to show resin's fishing efficiency. The resin was washed five times with a total of 50 mL of wash buffer (10mM Tris pH 7.5, 0.2% Igepal, 0.3M NaCl, 10% glycerol, 15mM Imidazol, water add up to 50 mL, add 1mM DTT at time of use) (10 mL each time) to rinse off any nonspecific binding materials. Another 15 µl aliquot was saved from the washed solution for SDS-PAGE analysis to demonstrate the washing efficiency. Lastly, 200 µl of elution buffer (10mM Tris pH 7.5, 0.1% Igepal, 0.2M NaCl, 10% glycerol, 250mM imidazol, water, add 2 µl/mL DTT at time of use) was applied to separate Zfp410 and the resin since the buffer contained 250 mM imidazole, which competed the binding of Ni-NTA with the poly-histidine tag (Bio-Rad, 2021). The same step was repeated four more times so that a total of five elution samples (200 µl each) were retrieved from the resin. 15 µl of aliquot from each elution was saved for SDS-PAGE analysis in order to see the amount of Zfp410 obtained and its purity. All five elution along with the wash and flow-through samples were flash frozen using liquid nitrogen and kept in the -80 °C freezer. Aliquots of all samples were added with equivalent amount of 2x SDS loading buffer (purchased from Bio-Rad) and stored in the -20 °C freezer.

SDS-PAGE is one of the most commonly used molecular techniques in cellular biology, biochemistry, and other scientific areas. It is used to separate proteins based on molecular weights that are between 5 and 250 kDa. The most critical reagent used in this task, SDS, serves as a strong detergent that denatures proteins to their unfolding state due to its ability of masking protein's intrinsic charge (Matsumoto, Haniu, & Komori, 2019). The polyacrylamide, on the other hand, eliminates the influence of proteins' charges and structures so that each individual protein solely travels based on its size, or molecular weight. The actual separation step takes

place in a gel apparatus that contained 1x SDS running buffer (100 mL of 10x SDS buffer purchased from Bio-Rad, water up to 1 L) with suitable electrolytes (applied voltage is usually around 100-200V). Because of the negative charge, proteins will move towards from the top to the positive site, which is the bottom of the gel (Pavlova, 2018). Based on the negative correlation between molecular size and protein's travel distance, one can determine the size of a specific protein by comparing the protein's final position to the molecular weight standard ladder: the bigger the protein is, the slower it travels down to the gel and therefore stays nearer to the top of the gel.

Before starting the procedure, all samples from Ni-NTA/His purification were first mixed with equivalent amount of 2x SDS solution (15 µl) and boiled 3 minuets to fully denature proteins. 5 µl of each sample was loaded in the 4% polyacrylamide gel (purchased from Bio-Rad) one next to each other along with 5 µl of molecular weight standard ladder. After added about 700 mL of 1x SDS running buffer in the apparatus (Figure 6), the gel was run for one hour to separate all proteins, and the electrophoresis was completed once the aligned dark blue line was close to the bottom of the apparatus. Since proteins that were separated in the gel by size remained invisible except the SDS blue dye at the bottom, the gel was then visualized by Coomassie Blue stain (2.5 g Coomassie Brilliant Blue, 454 mL of Methanol, 92 mL of Glacial Acetic Acid, water up to 1 L). This dye specifically binds to all proteins and simply covered the rest of the gel in dark blue color. Followed by overnight incubation with 100-200 mL of high de-stain buffer (7.5% Glacial Acetic Acid, 200 of mL methanol, water add up to 1 L), any non-protein region's stain was washed off and each protein then appeared as a distinctive blue band. The darker the band, the more concentrated the protein was. The purity of products was determined by the number of bands presented in each lane since each band represented one

specific protein. Lastly, the position of the band from each elution sample would demonstrate whether Zfp410 was successfully pooled out according to the molecular weight since the ideal molecular weight of Zfp410 should be in between 37-40 kDa based on the number of base pairs presented in the cDNA of Zfp410 (Figure 7).

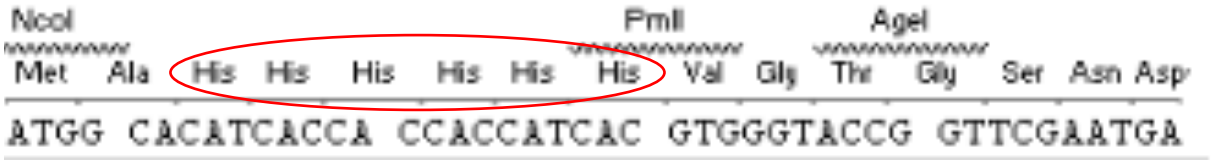

Figure 5. The poly-histidine-tag shown in the cloning expression region of pET-45 b (+).

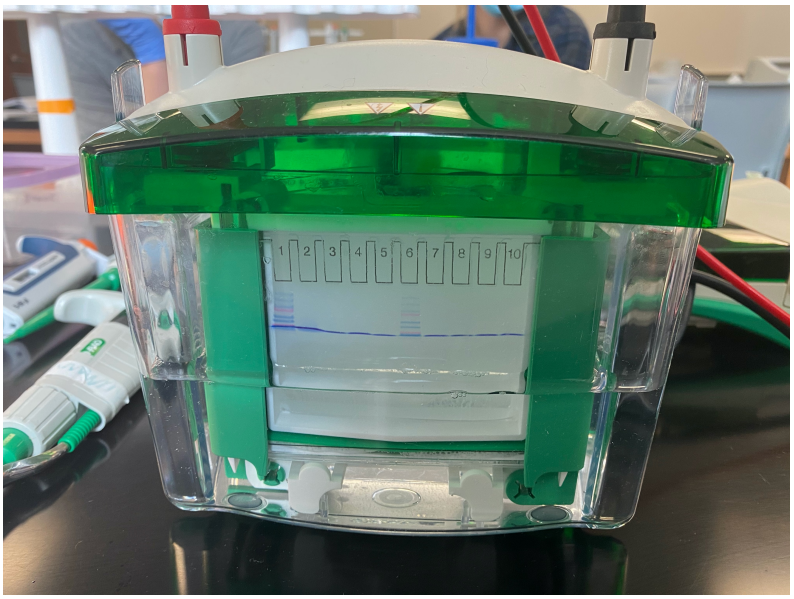

Figure 6. Example of an SDS-PAGE apparatus within protein samples.

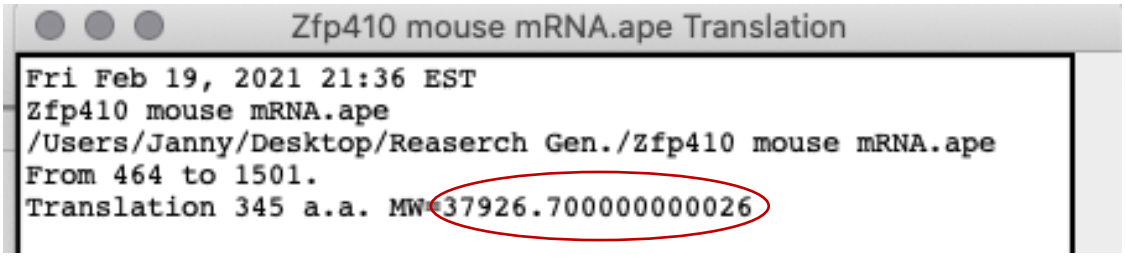

Figure 7. The alternative molecular weight of Zfp410 calculated by ApE software based on its cDNA sequence.

2.3. Results

The gel was visualized with a portal white board as background so that the standard ladder lane could be easily observed; however, all five elution samples were hardly seen on the gel, which indicated that the amount of Zfp410 presented on the gel was significantly low. The lane of washed sample during purification step also showed no band, which proved that the washing step was sufficient. Moreover, the flow through sample obtained from the solution that was not captured by the resin showed no band, which also demonstrated low Zfp410 present. As mentioned in the last section, the only possible explanation for this situation was that the majority of Zfp410 was left in the insoluble form from the second spin after lysis and sonication, and this was caused by a commonly appeared problem during recombinant protein purification known as the inclusion body.

Due to high concentration of protein created, the majority of Zfp410 aggregated together and formed a pellet instead of floating at the top in its soluble form. This phenomenon happened frequently in *E. coli* if the inside recombinant proteins are misfolded or presented in huge quantity, and the easiest way to solve this problem is to dissolve the pellet with a denaturing prep (Wingfield, Palmer, & Liang, 2014). As a result, the pellet that was saved previously was then incubated into 5 mL of urea buffer (30 g urea solid, 50mM Tris pH 7.5, 150mM NaCl, water add up to 50 mL) overnight in the cold room since urea is a strong detergent in comparison to the lysis buffer used previously and can dissolve any insoluble protein. Fortunately, all zfp410 was then successfully dissolved in the solution as expected. The following procedures were the same as the purification step described above, and all samples were run on SDS-PAGE for analysis.

Unlike the previous prep, the denaturing prep worked perfectly in terms of purifying Zfp410 based on the gel result (Figure 8). All elution samples were not only visible, but appeared

as huge thick bands, especially for the first two elutions. The thickness of band decreased as the elution number increased, which was the ideal result since Elution 3, 4, and 5 should contained less Zfp410 compared to the first two. In additions, all elution samples only showed one distinctive band around 37-39 kDa besides the flow through sample. This indicated that the purity of the denaturing prep was significantly high, and the desired Zfp410 was successfully purified by Ni-NTA resin.

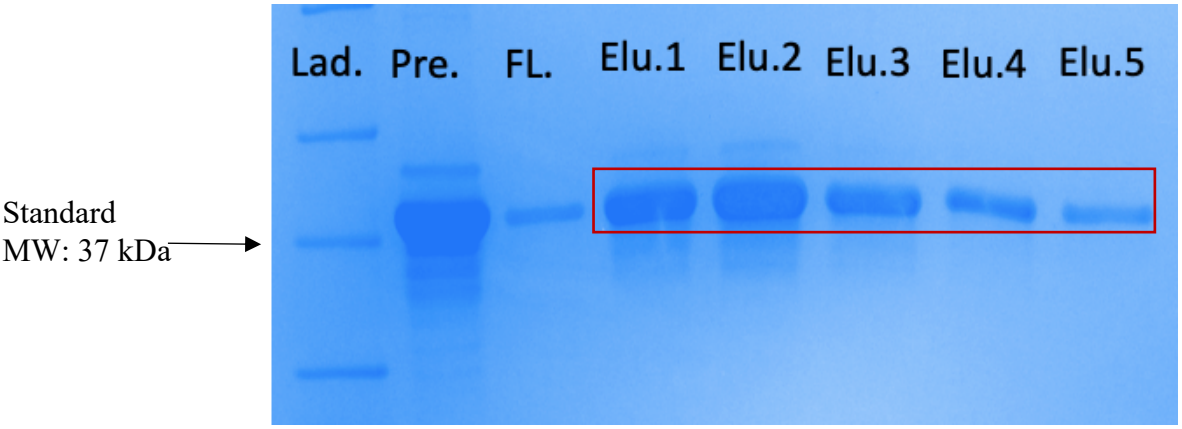

Figure 8. The SDS-PAGE gel of Zfp410 denaturing prep after Coomassie Blue stain.

3. Aim #2: Zfp410's Antibody production & characterization

3.1. Overview:

After successfully expressed and purified the recombinant protein Zfp410 in *E. coli*, custom polyclonal antibody needed to be retrieved through a suitable animal host since antibodies are primarily generated by the host's immune system. Since the concentrated Zfp410 was eluted with a high salt elution buffer (250mM imidazole) and would be lethal to the host, Zfp410 was therefore first dialyzed to get rid of the excess salt. The concentration of Zfp410 was then determined by the Bradford Assay, and the final dialyzed protein was shipped to a rabbit farm for antibody generation. The retrieved antibody was then characterized using western and dot blot to confirm its binding affinity.

3.2. Methods:

3.2.1. Antibody production:

The successful purification of Zfp410 allowed the project to move forward into the second aim, which was the generation and characterization of its antibody so that future molecular experiments such as immunoprecipitation can be performed. One dominant way to retrieve antibody is through the injection of protein of interest into animal models such as mice and rabbits. The mechanism behind this concept is that any non-self or foreign protein will be targeted as an antigen by the host's immune system regardless of harmful or not (Ma & O'Kennedy, 2015). Once the immune cells detect the presence of the foreign proteins, an immediate cellular reaction will take place to process the protein and generate antibodies that specifically target it. This ensures a high specificity and efficiency between the antibody and its antigen. In fact, the development of vaccine is a well-known fascinating application of this concept.

Similarly, Zfp410 would act as a foreign protein so that the antibody could be produced by the host's immune cells. Since certain amount of recombinant protein was required for antibody production, the purified Zfp410 needed to be quantified to ensure enough protein injection in the rabbit. Before the quantification assay, Zfp410 was dialyzed using a semipermeable dialysis tubing that contained small pores, which permits small molecules such as salt to travel through and left big size proteins behind (Li et al., 2019). The concentrated protein sample was carefully pipetted into the tied dialysis tubing and placed in the dialysis buffer (10 mL of Tris pH 7.1, 2 mL of 100% NP-40, 80 mL of 5M NaCl, 1 mL of 1M DTT, water add up to 2 L) for overnight incubation so that salt and other small molecules can pass into the buffer based on osmotic pressure gradient. This procedure was used to remove the excess imidazole

(250 mM) since high salt concentration could be lethal for the rabbit due to alteration of its blood pH level. Dialyzed Zfp410 was then quantified using the Bradford protein assay. Bradford is a well-known molecular technique that determines the concentration of an unknown protein sample by spectroscopic analysis. The basic concept is that the reaction of Bradford reagent with protein causes the absorbance shift of the dye itself. Under normal condition, the Bradford dye appears as red-brown color and maintains the best absorbance at a wavelength of 595 nm (Kielkopf, Bauer, & Urbatsch, 2020). If a protein sample was mixed with the dye, it undergoes an immediate reaction and the color changes to blue, and the color darkens as the concentration of protein gets higher. Therefore, series of known concentration protein standards like bovine serum albumin (BSA) are used first to generate a linear standard curve (Absorbance vs. concentration) (Figure 9). This is extremely helpful to determine the concentration of unknown protein if absorbance value is known and falls in the linear range.

Serial dilutions of BSA standards were prepared by diluting the protein stock (2 mg/mL) with water down to 0.1, 0.2, 0.3, 0.5, and 0.6 mg/mL respectively. A linear standard curve was generated based on its concentration against absorbance value at 595 nm, and the linear equation of the curve was computed by Excel (Figure 10). Moreover, 3 µl of Zfp410 was then mixed with 1 mL of 1x Bradford reagent (made from 5x Bradford stock solution that was purchased from Bio-Rad) and placed inside a cuvette for spectrometry reading. The resulting absorbance value was recorded and plugged into the linear equation to calculate its concentration. The result displayed a high concentration of purified Zfp410 (0.481 mg/mL, or 481 ng/µl), and this granted the shipping of recombinant protein in November 2019 to the Poconal Rabbit Farm & Laboratory (PRF&L).

Antibody production required certain period that lasted approximately three months, and the laboratory company took four injection stages during this period: a first 200 µl injection followed up with 100 µl dose and two more separate doses (50 µl each) as booster shots. Right before the first injection, a small bleed form rabbit was collected and sent back as the pre-immune control since the serum did not contain any antibody. Each small bleed was sent back after one month period and the exsanguinated sample that had full Zfp410 antibody was retrieved after three months (received on February 3, 2020) (Figure 11).

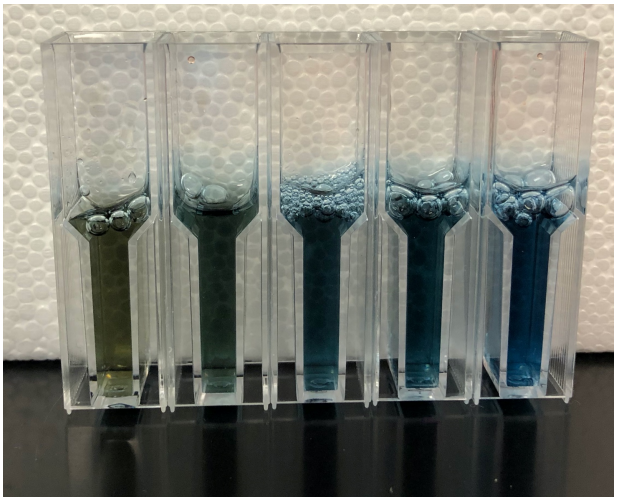

Figure 9. The color change of a series BSA standards after reacting with Bradford reagent.

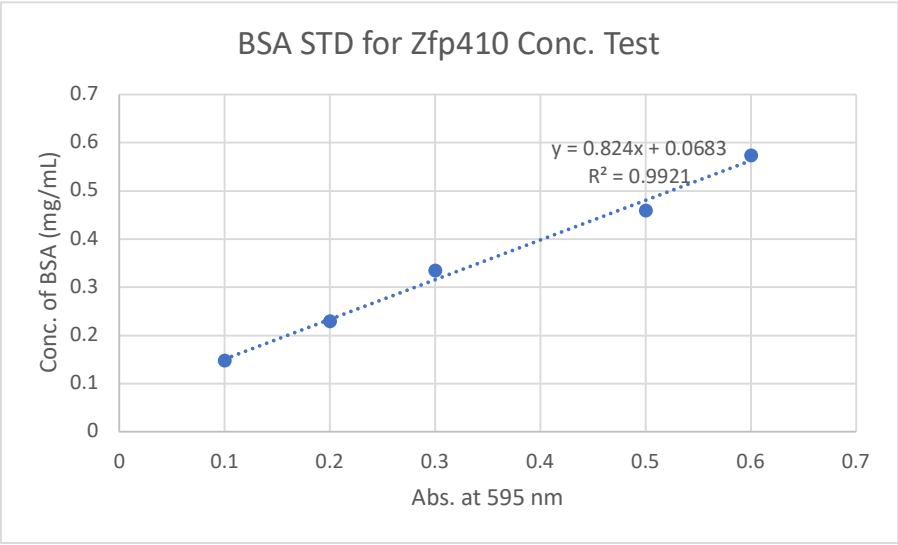

Figure 10. The BSA standard curve used for Zfp410 quantification.

Figure 11. Collected full Zfp410 antibody from the PRF&L on February 3, 2020.

3.2.2. Antibody characterization by dot blot & western blot:

With all stages of Zfp410 antibodies, the quality of the last serum sample was characterized by dot blot and western blot. While Coomassie blue stain only gave the general picture of all proteins presented after SDS-PAGE separation, western blot and dot blot showed specific proteins that can only be recognized with the applied antibodies. These techniques are beneficial for recognizing specific protein of interest from a variety of tissue extracts that contain thousands of proteins. Prior to western blot, the protein has to be transferred from SDS-PAGE gel onto a nitrocellulose membrane using transfer buffer (200 mL of 5x Transfer buffer, 5% methanol, water add up to 1 L) for long-term storage because the gel is very easily broken. After transferring, the membrane that stored proteins was then placed in 100 mL of blocking solution that contained 3% of fat-free milk (3 g of dry milk powder, TBST buffer add up to 100 mL)

overnight at room temperature to cover any non-specific binding region on the membrane. The primary antibody that is solely used for binding its target antigen and leaves any non-recognizable protein species behind were applied once blocking was completed. After rinsed with 100 mL of TBST buffer (1mM Tris Acid, 150mM NaCl, 0.5% Tween-20, water add up to 1 L) to wash off excess primary antibody (1:1000 dilution with 50 mL TBST and 0.5 g dry milk powder) in order to prevent over signal expression, the final step required the incubation with a secondary antibody that only recognizes the primary antibody, which indirectly binds to the protein of interest (Pillai-Kastoori, Schutz-Geschwender, & Harford, 2020). Secondary antibody often contains a fluoresce tag that allows signal detection under certain light wavelength, which results in glowing under special spectrometer, and this is how the protein that was picked up by the antibody can be visualized as a colored band while other non-recognizable proteins remain invisible on the screen.

Unlike western blot, dot blot requires neither SDS-PAGE nor transferring but directly dotting the protein sample on a nitrocellulose membrane instead (Fischer, Heckler, & Boon, 2020). Proteins of interest are usually diluted and dotted on the membrane for detection since it will be invisible once dried up. The dotted membrane will then be placed in blocking solution and incubated with primary and secondary antibody just like western blot.

In this project, four serial dilutions of Zfp410 were prepared by diluting the stock protein (481 ng/µl) down to 5, 0.5, 0.05, and 0.005 ng/µl with TBST buffer respectively. Total of 2 µl from each dilution was dotted on the membrane so that the mass of Zfp410 presented would increase by a factor of 10 from dilution 4 to 1 (0.01 ng, 0.1 ng, 1 ng, 10 ng respectively). The dotted membrane was then incubated with 100 mL of blocking solution for 2 hours and followed up with 1 hour incubation with 50 mL primary antibody (the Zfp410 antibodies that was diluted

in 1:1000) and 1 hour incubation with 50 mL of secondary antibody (*Li-COR IRDye® 680RD goat anti-rabbit IgG secondary antibody diluted in 1:15,000*). The membrane was then imaged by *Li-COR* UV-Vis spectrometer at 700 nm to visualize Zfp410 (all four dots should appeared in red).

Since dot blot only indicated the efficiency of Zfp410 antibody, western blot was performed at the same time to verify the specificity of Zfp410 antibody's recognition. Same amount of Zfp410 was loaded on the SDS-PAGE in the exact increasing trend as the one applied in the dot blot with the addition of standard molecular weight ladder on the side as size comparison. The gel was transferred for 1 hour with 1x transfer buffer at 400mA and incubated overnight in 100 mL of blocking solution. After incubation with the same primary and secondary antibody, the result was also imaged by *Li-COR* UV-Vis spectrometer at 700 nm.

3.3. Results

Images from both dot blot and western blot confirmed the efficiency and accuracy of Zfp410 antibody's binding ability to its target. The dot blot showed a clear increasing signal trend from left to right as the concentration of Zfp410 increased, and the result also demonstrated strong binding affinity in between the antibody and Zfp410. The quality of the antibody was significantly high since it was able to easily detect the presence of only 0.1 ng of Zfp410, and the dot was almost visible for the 0.01 ng Zfp410 sample (Figure 12a). On the other hand, western blot showed similar signal trend to dot blot, and it proved the accuracy of antibody's detection based on the band's position, which was around 39~40 kDa (Figure 12b). Since both images together confidently implied the positive quality test of Zfp410 antibodies that was made from the purified recombinant protein, therefore, Zfp410 antibodies were ready to be used as the

dominant tool to fish out the interacting partners from the pool, which shifted the project to its last aim.

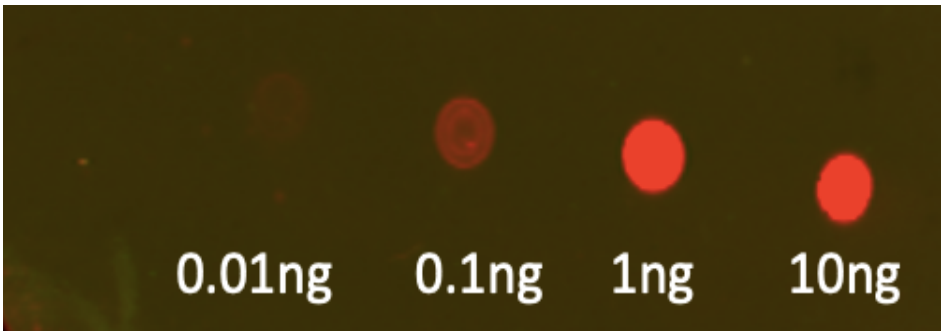

Figure 12a. Dot blot using Zfp410 antibody to detect the recombinant protein.

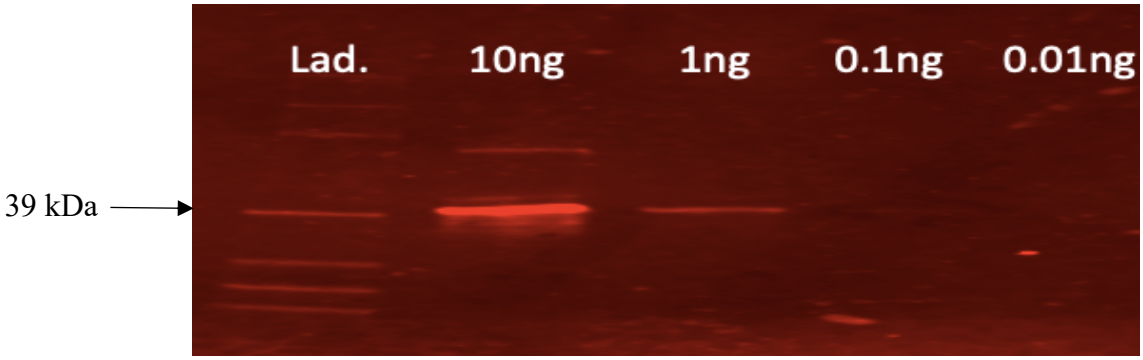

Figure 12b. Western blot using Zfp410 antibody for the determination of its accuracy.

4. Identification of Zfp410's interacting partners via the utilization of its antibody

4.1. Overview:

In order to further discover Zfp410's function, it is vital to identify its binding partners in *vitro*. By demonstrated a strong binding affinity to the recombinant protein, polyclonal Zfp410 antibody was then applied in Co-immunoprecipitation to pull out Zfp410's interacting proteins. The outcome of this technique was analyzed by comparing the result of both silver stain and western blot in order to identify ultimate Zfp410's interacting partners. The detected proteins were collected and sent off for Mass Spectrometry analysis to confirm their actual identities.

4.2.1. Co-immunoprecipitation:

Protein interactions can be determined by Co-immunoprecipitation (Co-IP). Generally, immunoprecipitation is widely used for precipitating a protein out by antibody that binds specifically to it. Similarly, Co-IP uses the exact mechanism besides that it can be described as the advanced version of immunoprecipitation. As described in the name, Co-Ip is capable of pulling out a protein complex instead of just the single protein, and this is what gives Co-IP the ability to find out interacting partners for the protein of interest (Maccarrone et al., 2017). By extracting out a known protein from a pool of proteins in solution using its antibody, any protein that is correlated or interacted with this protein will be extracted at the same time. For this project, Co-IP served as the key procedure to determine the interacting partners of Zfp410 and may further suggest that it is responsible for regulating significant numbers of folate-dependent genes in the hippocampus.

Unlike the purification step, protein A agarose resin, or bead, was used in Co-IP to perform the precipitation. In contrast to the His-resin, this bead bound specifically to the constant region of the Zfp410 antibody instead of an amino acid tag. The variable region of the antibody would bind to Zfp410 and indirectly grabbed its interacting partners. Agarose beads are not as specific in comparison to the Ni-NTA His-resin, which allows for non-specific binding of other proteins to the bead. In order to maximize the binding capacity of the bead to the Zfp410 antibody, 100 µl of the resin was blocked with 10 µl of the standard protein BSA to prevent non-specific proteins from binding. The other reason of using the BSA was that its molecular weight is well-known (~66 kDa) and can be easily identified in various analysis methods (Sigma Aldrich, 2021). The resin was then washed twice (1 mL per time) with wash buffer (150mM NaCl, 0.2% Igepal, 20% glycerol, 20mM Tris pH 7.5, 0.2mM EDTA, 1.5mM $MgCl_2$ water add

up to 10 mL) to remove any excess BSA and divided into two 1.5 mL microcentrifuge tubes (40 µl of resin per tube). A total of 16 µl of either pre-immune serum sample (no-antibody present) or the Zfp410 antibody were added to each tube. The purpose of using the pre-immune serum was solely acting as the negative control since there was no antibody present. Both tubes were then incubated at 4 °C overnight to ensure the beads were fully saturated with the respective antibody.

The covered beads were then used to precipitate Zfp410 out from the previous prepared HeLa cytoplasmic extracts by mixing 420 µl of the extract with 2 µl of purified Zfp410 along with 180 µl of lysis buffer to rupture the cell's plasma membrane. Total of 300 µl of the HeLa extracts master mixture was added to each of the two tubes respectively. Both tubes were then incubated together for 2 hours in the 4 degree of Celsius cold room to encourage maximum binding between the Zfp410 antibody and the recombinant protein along with its interacting partners.

Lastly, both samples were treated with 30 µl of DTT elution buffer (200 µl of Lysis buffer and 8 µl of 0.5M DTT) in order to break the linkage between antibodies and the beads. The reason of using DTT as the major elution solution was that it acted as a reducing agent and is therefore capable of breaking the disulfide bond, which is the dominant linkage in between the antibodies. The samples were all spun down at max speed (15,000 RPM) for 5 min to separate the resin from the protein complexes and followed with addition of 30 µl of 2x SDS loading buffer for preparation of SDS-PAGE analysis.

4.2.2. SDS-PAGE followed with western blot & silver stain:

SDS-PAGE was utilized to separate proteins from the Co-IP by using 8% of polyacrylamide gel. Total of 10 µl of the standard ladder was loaded on the far left, and 10 µl of

the sample from each tube was loaded side by side. Since the result was analyzed by both western blot and silver stain, a total of two gels that had exact same samples were run for one hour at 70 mA. The gel that was used for western was then simply transferred in ~600 mL 1x transfer buffer for another hour and incubated in 3 % fat-free dry milk blocking solution overnight. All other procedures were completed in the next day and the result was imaged using *Li-COR* UV spectrometer.

As mentioned previously, western blot gave high specificity binding while Coomassie stain gave general outcome of the whole SDS-PAGE separation; however, the western itself cannot determine whether the Co-IP was ultimately successful. Therefore, one important technique, known as the silver stain, was then used as a side-by-side comparison along with the western. Although it is another type of staining that shows all protein bands on the gel just like the Coomassie blue, silver stain is highly sensitive and therefore requires more caution when performing the analysis techniques. Based on the definition, silver stain uses silver to selectively bind to targets due to the chemistry inside the reagents (Wong et al., 2019). The major reagents used for this technique were silver nitrate and formaldehyde. While the former reagent gives the metal that allows visualization of protein band, formaldehyde is the material that allows the formation of metal silver since it reduces silver ion in its metal form, and this is what makes silver stain appeared as yellow-amber color at the end of development. As metal silver precipitated out from the solution, it binds to the protein and sticks there forming dark brown color as the development progresses (Vu et al., 2020). In addition, sodium bicarbonate is also used during silver staining as well since it can act as a buffer of the acidic solution resulted from the reaction between formaldehyde and silver nitrate. Citric acid is usually applied in the end to stop the reaction once the bands can be clearly visualized to prevent over development.

Because of the high sensitivity of silver stain, gels can never be touched directly, and it is best to perform the silver stain in a clean glass container for the purpose of preventing contaminations. The container used for the second SDS-PAGE gel was previously soaked in 10 % SDS buffer overnight prior usage to rinse off any residues that may have been stuck to the container. After rinsing with de-ionized water, the gel was then placed into the clean container along with 100 mL of 50 % methanol solution for 30 minutes. An additional 10 minutes of incubation were performed with 100 mL of 5 % methanol and rinsed 3 times with water carefully, without dropping the gel. 10 μM DTT solution was applied afterwards for 20 minutes, and ~50 mL of 0.1 % silver nitrite solution was carefully added to soak the gel. Once the 20 minutes incubation time ended, the gel was quickly rinsed with water again. 100 mL of Developing solution (formaldehyde & sodium carbonate) was then added to the gel for development of the silver saturated bands. Citric acid was added (5 g per 100 mL of developing solution) when all bands could be clearly visualized. The result was then imaged by the UV-spectrometer and compared to the western blot.

4.3. Results

Both images from western blot and silver stain were compared together to determine the success of Co-IP. Based on the bands shown in the western, it was clear to see that negative control (pre-immune samples) worked since there were no significant bands besides the heavy and light chain of the antibody itself (~50-60 kDa). In contrast, the silver stain displayed multiple bands for each lane regardless of pre-immune or antibodies, which was expected since it was supposed to show all proteins obtained from the HeLa extracts (Figure 13). For the samples that were incubated with the Zfp410 antibody, few bands besides the Zfp410 and antibody chains were shown in the western blot, which indicated that those proteins were most likely just Zfp410

isoforms. In contrast, the silver stain displayed 6 distinctive bands that were at higher molecular weights (65, 78, 90, 100, 150, 250 kDa respectively) and could not be visualized by western blot (Figure 14). This observation could indicate potential Zfp410 interacting partners were isolated. These bands visualized by silver stain were cut with a clean razor blade directly from the gel and put into 1.5 mL centrifuge tubes to be sent off for Mass spectrometry analysis, in order to be identified in the future.

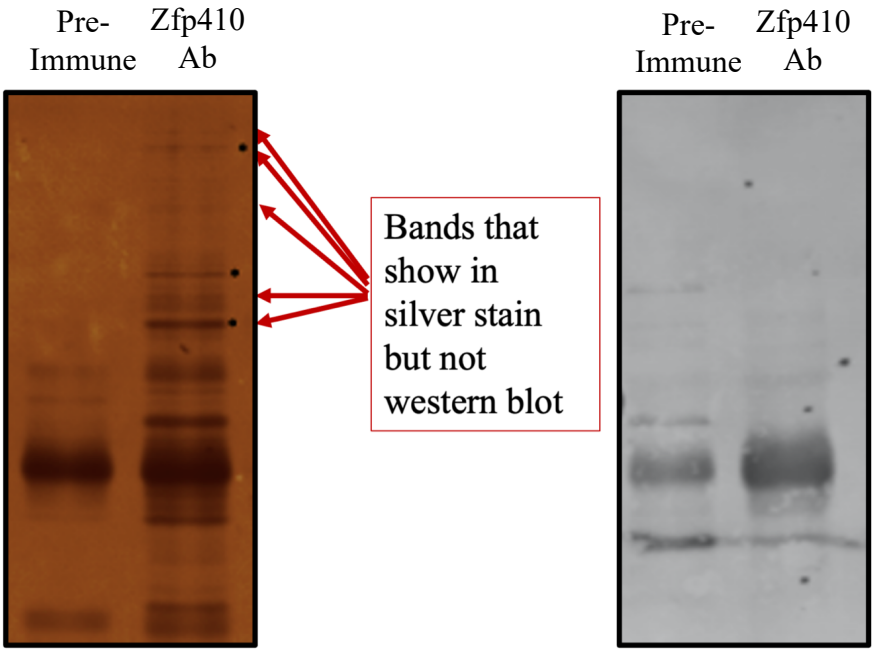

Figure 13. Comparison of western blot (right) and silver stain (left) from Zfp410's Co-IP.

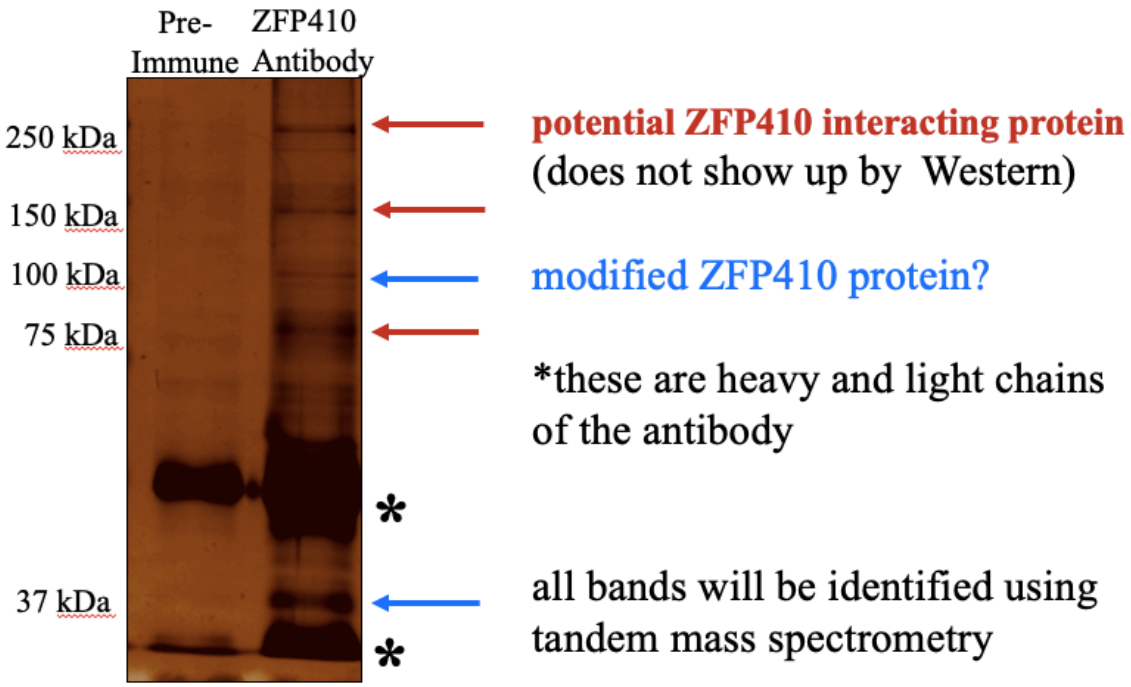

Figure 14. Potential identified Zfp410 interacting partners that would be cut off for MS analysis.

5. Conclusions

Understanding the mechanisms of how transcription factors function and interact with other proteins are essential for various scientific studies. Furthermore, the association between transcription factors and a number of diseases had shed light on the clinical path for many scientists and medical doctors. Research had shown effective treatments of disorders by targeting correlative transcription factors. One study has developed method that monitored certain transcription factors' expressions and inhibited their interactions with other proteins so that the oncogenes were not able to be activated, which therefore inhibited cancer growth (Lambert et al., 2018). Similarly, folate deficiency is known to be linked with cognition impairment and multiple cardiovascular defects. Under the awareness of certain transcription factor's association with lack of this nutrient, possible treatment can therefore be considered in targeting its expression and interaction. This project took a deep investigation of folate-dependent transcription factor Zfp410 in the purpose of understanding its specific role of regulating gene expressions in the

absence of folic acid. All three aims were achieved as the recombinant protein of Zfp410 was produced and purified, and the characterized polyclonal antibody was utilized in co-immunoprecipitation for the isolation of Zfp410's binding partners.

As the project had successfully isolated and shipped off samples for identification of Zfp410's interacting partners, future studies are needed to analyze those sample proteins for further understanding of Zfp410's interactions. The data of mass spectrometry will be interpreted within each band to estimate possible identities of interacting proteins. Additional co-immunoprecipitation assay is required to confirm those estimations by purchasing each ultimate interacting protein's antibody for western blot usage. Once the identities of Zfp410's interacting proteins are confirmed, the project will move into its final stage, which is the identification of Zfp410's binding sites on a DNA sequence by Chromatin immunoprecipitation (ChIP) assay. This technique will allow measurements of the interaction between Zfp410 and DNA in the cell, in which enables the identification of genetic targets of transcription factor Zfp410 in any given DNA sequences (Ayala et al., 2018). This part of the study is vital since it will fully demonstrate Zfp410's binding sites in the entire genome. Along with the results from co-immunoprecipitation, the project is then able to fully identify Zfp410's interactions both cellularly and molecularly, and these findings can be extremely useful for future investigations in ultimate treatment development of folate-deficiency related defects.

References

Artner, I., Blanchi, B., Raum, J. C., Guo, M., Kaneko, T., Cordes, S., Sieweke, M., & Stein, R. (2007). MafB is required for islet beta cell maturation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 3853–3858.

Ayala, J. C., Benitez, J. A., & Silva, A. J. (2018). Chromatin Immunoprecipitation. *Methods in molecular biology (Clifton, N.J.)*, *1839*, 65–75.

Bio-Rad Laboratories. (February 21, 2021). Retrieved from https://www.bio-rad.com/featured/en/his-tag-purification

Bürger, J., Akgün, D., Strube, P., Putzier, M., & Pumberger, M. (2019). Sonication of removed implants improves microbiological diagnosis of postoperative spinal infections. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *28*(4), 768–774.

Cooney, C. A., Dave, A. A., & Wolff, G. L. (2002). Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *The Journal of nutrition*, *132*(8 Suppl), 2393S–2400S.

Del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., & Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiology and molecular biology reviews: MMBR*, *62*(2), 434–464.

Fischer, J. T., Heckler, I., & Boon, E. M. (2020). SDS-PAGE and Dot Blot Autoradiography: Tools for Quantifying Histidine Kinase Autophosphorylation. *Methods in molecular biology (Clifton, N.J.)*, *2077*, 37–49.

Flournoy, B. (2018, June 04). Difference between optical density & absorbance. Retrieved from https://sciencing.com/difference-between-optical-density-absorbance-7842652.html.

Gietz R. D. (2014). Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods in molecular biology (Clifton, N.J.)*, *1205*, 1–12.

Kielkopf, C. L., Bauer, W., & Urbatsch, I. L. (2020). Bradford Assay for Determining Protein Concentration. *Cold Spring Harbor protocols*, *2020*(4), 102269.

Lambert, M., Jambon, S., Depauw, S., & David-Cordonnier, M. H. (2018). Targeting Transcription Factors for Cancer Treatment. *Molecules (Basel, Switzerland)*, *23*(6), 1479.

Latchman, D. S. (1997). Transcription factors: an overview. *The international journal of biochemistry & cell biology*, *29*(12), 1305–1312.

Lawton, A., Morgan, C. R., Schreiner, C. R., Schreiner, C. G., Baumann, J., Upchurch, B., Xu, F., Price, M. S., & Isaacs, G. D. (2020). Folate-dependent cognitive impairment associated with specific gene networks in the adult mouse hippocampus. *Front. Nutr. 7:574730.*

Lewis M. (2013). Allostery and the lac Operon. *Journal of molecular biology*, *425*(13), 2309–2316.

Lillycrop, K. A., Phillips, E. S., Jackson, A. A., Hanson, M. A., & Burdge, G. C. (2005). Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *The Journal of nutrition*, *135*(6), 1382–1386.

Li, J., Wang, Y., Xu, X., Cao, W., Shen, Z., Wang, N., Leng, J., Zou, N., Shang, E., Zhu, Z., Guo, J., & Duan, J. (2019). Improved dialysis removal of protein-bound uremic toxins by

salvianolic acids. *Phytomedicine: international journal of phytotherapy and phytopharmacology*, *57*, 166–173.

Ma, H., & O'Kennedy, R. (2015). The Structure of Natural and Recombinant Antibodies. *Methods in molecular biology (Clifton, N.J.)*, *1348*, 7–11.

Maccarrone, G., Bonfiglio, J. J., Silberstein, S., Turck, C. W., & Martins-de-Souza, D. (2017). Characterization of a Protein Interactome by Co-Immunoprecipitation and Shotgun Mass Spectrometry. *Methods in molecular biology (Clifton, N.J.)*, *1546*, 223–234.

Marbach, A., & Bettenbrock, K. (2012). lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *Journal of biotechnology*, *157*(1), 82–88.

Massiah, M. A., Wright, K. M., & Du, H. (2016). Obtaining Soluble Folded Proteins from Inclusion Bodies Using Sarkosyl, Triton X-100, and CHAPS: Application to LB and M9 Minimal Media. *Current protocols in protein science*, *84*, 6.13.1–6.13.24.

Matsumoto, H., Haniu, H., & Komori, N. (2019). Determination of Protein Molecular Weights on SDS-PAGE. *Methods in molecular biology (Clifton, N.J.)*, *1855*, 101–105.

Narlikar, G. J., Fan, H. Y., & Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, *108*(4), 475–487.

Pillai-Kastoori, L., Schutz-Geschwender, A. R., & Harford, J. A. (2020). A systematic approach to quantitative Western blot analysis. *Analytical biochemistry*, *593*, 113608.

Pavlova, A. S., Dyudeeva, E. S., Kupryushkin, M. S., Amirkhanov, N. V., Pyshnyi, D. V., & Pyshnaya, I. A. (2018). SDS-PAGE procedure: Application for characterization of new entirely uncharged nucleic acids analogs. *Electrophoresis*, *39*(4), 670–674.

Robert, A. J. (2000). Folate, DNA methylation, and gene expression: factors of nature and nurture. *The American Journal of Clinical* Nutrition, *72*(4), 903–904.

Robertson, K. D., & Jones, P. A. (2000). DNA methylation: past, present and future directions. *Carcinogenesis*, *21*(3), 461–467.

Selhub, J., Morris, M. S., Jacques, P. F., & Rosenberg, I. H. (2009). Folate-vitamin B-12 interaction in relation to cognitive impairment, anemia, and biochemical indicators of vitamin B-12 deficiency. *The American journal of clinical nutrition*, *89*(2), 702S–6S.

Sigma Aldrich. (February 23, 2021). Retrieved from https://www.sigmaaldrich.com/technical-documents/articles/biology/albumin-from-bovine-serum.html

Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., & Maertens, B. (2015). Purification of His-Tagged Proteins. *Methods in enzymology*, *559*, 1–15.

Sturød, K., Salvadori, G., Junges, R., & Petersen, F. C. (2018). Antibiotics alter the window of competence for natural transformation in streptococci. *Molecular oral microbiology*, *33*(5), 378–387.

Vu, P., Faraday, R., Vu, D., & Kim, J. (2020). Silver nitrate stain masquerading as a heterotopic ossification. *Radiology case reports*, *15*(5), 450–453.

Wingfield P. T. (2015). Overview of the purification of recombinant proteins. *Current protocols in protein science*, *80*, 6.1.1–6.1.35.

Wingfield, P. T., Palmer, I., & Liang, S. M. (2014). Folding and Purification of Insoluble (Inclusion Body) Proteins from Escherichia coli. *Current protocols in protein science*, *78*, 6.5.1–6.5.30.

Wong, A., Xie, S., Tang, B. Z., & Chen, S. (2019). Fluorescent Silver Staining of Proteins in Polyacrylamide Gels. *Journal of visualized experiment: JoVE*, (146), 10.3791/58669.