Fall 2021

Determining the Primary DNA Substrates of SHLD2's OB-fold Domains

Hari Patchigolla
University of Connecticut - Storrs, hari.patchigolla@uconn.edu

Determining the Primary DNA Substrates of SHLD2's OB-fold Domains

Hari Patchigolla
Mentor: Dr. Dmitry Korzhnev

Abstract:

Failure to repair DNA double-stranded breaks leads to cell death. Radiation therapy is commonly used to kill cancer cells by inducing these breaks. However resistance to radiation therapy, due to a hyperactive DNA double-stranded break repair pathway, is a common occurrence that makes cancer patients more prone to relapse. The Shieldin complex is shown to promote DNA-double stranded break repair by binding to DNA at sites of damage. Thus, the objective of this project is to understand the affinity and type of DNA that Shieldin binds to, through gel-shift assays and computational modeling, for the eventual creation of an inhibitor for this protein to stop the emergence of radioresistant cancer cells.

Introduction:

More than 14 million new cases of cancer are diagnosed globally each year, and over 50% of these patients will undergo radiation therapy (RT) [1]. RT induces DNA double-stranded breaks (DSBs), a cut through both sides of the DNA backbone resulting in DNA overhangs (Fig. 1) [2]. By creating DSBs, RT overwhelms cancer cells by blocking the progression of DNA replication and causing chromosomal translocations which creates genomic instability [2].

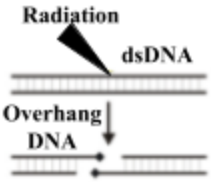

Fig 1: DNA DSB overhangs created by RT

Unfortunately, resistance to RT is observed within cancer cells due to an upregulation of DSB repair pathways [3]. As a result, the creation of chemotherapeutics that target DSB repair pathway machinery can inhibit cancer cell's ability to become radioresistant.

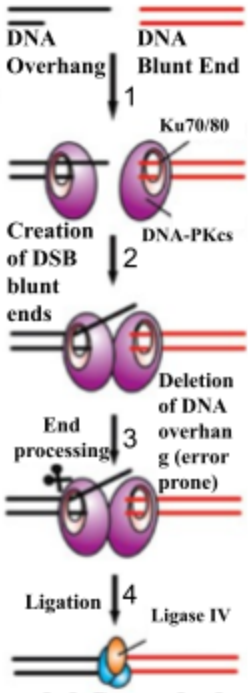

Fig 2: Schematic of error prone NHEJ

Our cells have two predominant pathways to repair DSBs: Homologous Recombination (HR) and Nonhomologous End-Joining (NHEJ). NHEJ accounts for almost 75% of DSB repairs and works by ligating/joining DSB blunt ends (Fig. 2). However, NHEJ inaccurately repairs DSBs by deleting DNA ends (Fig. 2), thus inducing mutagenesis that may contribute to carcinogenesis [4].

NHEJ plays a key role in radioresistant cancer cells since many proteins specific to NHEJ are over-expressed [5]. Furthermore, due to its error-prone repair of DSBs (Fig. 2), NHEJ enables cancer cells to accumulate mutations, further contributing to a radioresistant phenotype. Consequently, proteins involved in NHEJ are potential chemotherapeutic targets to impair the hyperactivity of this pathway in radioresistant cancer cells, thus making them more susceptible to RT. This has

previously been attempted by creating inhibitors for major NHEJ proteins such as DNA-PKcs or DNA Ligase IV. However, these inhibitors had poor specificity and *in vivo* toxicity hence preventing clinical application [6]. Therefore, identifying new chemotherapeutic targets within NHEJ is necessary.

A newly discovered protein, the Shieldin (SHLD) complex, promotes NHEJ by binding to DNA at sites of DSBs through three OB-fold domains on its SHLD2 subunit [7]. By interacting with DNA, SHLD acts downstream of 53BP1 and RIF1 to create the DNA blunt ends necessary for NHEJ. SHLD prevents exo- and endonucleases involved in HR from resecting (cutting away nucleotides) DNA ends which precludes HR from dominating [8].

SHLD is overexpressed in prostate and triple-negative breast cancer patients allowing for a hyperactive NHEJ pathway [10]. SHLD also plays a key role in the fusion of deprived telomeres characterized by a repetitive sequence of DNA at the ends of chromosomes that look like damaged DNA, due to the presence of DNA overhangs. This has a causative role in promoting genomic instability in colon and ovarian cancer tumors. Genetic screens have identified SHLD as a promoter of radioresistance in cancer cells since it localizes to DSBs induced by RT. Furthermore, knockout studies report SHLD depleted breast cancer cells are hypersensitive to RT induced DSBs [9]. Together, this data suggests that inhibiting SHLD's ability to bind to DNA could create cancer cells sensitized to RT.

SHLD binds DNA through the SHLD2 subunit. In order to create small molecule inhibitors for SHLD2, to impair NHEJ and prevent the onset of radioresistance in cancer cells, it must be further understood. One gap in knowledge is SHLD2's preferred DNA substrate (the type of DNA SHLD2 binds to at DSBs). Studies that have tried to identify this are inconsistent in their findings. Findlay *et al.* report that SHLD2 is proficient in equally binding to both dsDNA

and ssDNA [9], while Gao *et al.* claim SHLD2 binds to dsDNA with a higher affinity for ssDNA [11]. More importantly, overhang DNA, typically found at DSBs (Fig. 1), and sequence repetitive DNA is yet to be tested.

Methods:

Cell Culture -

4 mL of LB was inoculated with a colony of cells from the LB agar plate and 4 μL of the respective antibiotic. The tube was placed in a shaker at 37°C overnight.

Plasmid Prep -

Cells grown from a 4 mL culture were formed into a cell pellet in a 2 mL eppendorf tube using a microcentrifuge set to 11,000 rcf for 1 minute. The cell pellet was resuspended in A1 Buffer from the NucleoSpin Plasmid Prep Kit and lysed using the A2 buffer. The A2 lysis buffer solubilizes the cell membrane allowing the cell material (hence the plasmid) to be released. The lysis buffer is neutralized using the A3 buffer and spun down in the microcentrifuge for 20 minutes to separate the soluble and insoluble factions of the cell lysate. The soluble fraction (which contains the plasmid) is loaded onto a silica membrane column and washed several times with the AW and A4 wash buffers and finally 50 μL of elution was captured using the AE buffer. A 1:10 dilution was created by diluting 10 μL of the elution into 90 μL of Milli-Q water. The 280 nm absorbance of the dilution was reate using an OD spectrophotometer to read the contectration of the DNA in ng/μL and to assess the purity of the elution.

DNA Sequencing -

In order to determine if the plasmid contains the right sequence, teo samples per plasmid stock were created each with a concentration of 80 ng/μL. One sample added 5 μL of petUpstream primer (SEQ) and the other sample added the T7-term primer (5’

GCTAGTTATTGCTCAGCGG 3'). These primers bind to the respective parts of the plasmid to help with sequencing. The two samples were sent to Genewiz for DNA sequencing.

Primer Dilutions -

Primers were ordered for IDT in a lyophilized pellet. The amount of Milli-Q water added to resuspend the primers was determined using the following equation: [(mg*1000)*6000]/MW. A 1:10 dilution of the primer stock was done for later use.

Polymerase Chain Reaction (PCR) -

200 µL PCR tubes consisted of the following reagents: 25 µL of Master Mix, 1.5µL of forward primer, 1.5µL of reverse primer, 1µL of template at a 5ng/µL concentration, and 21µL of Milli-Q. To check if PCR worked, a 0.85% agarose gel was prepared by dissolving 2µL of SYBR Safe and 0.3 grams agarose into 35 mL of 0.5X TAE buffer (Included cycle settings). The gel was then poured into a tray with a gel comb and polymerized at room temperature for 15 minutes. Samples for the agarose gel were prepared by mixing 10 µL of the final PCR mix with 2 µL of 6X Purple Loading Dye. The samples were then loaded onto the agarose gel and run for approximately 20 minutes after which the gel was imaged under a UV light.

Cell Transformation -

5 µL of a 5ng/uL plasmid stock were pipetted into competent BL21 cells and incubated on ice for 25 minutes. The cells were then heat shocked for 45 seconds at 42°C and then placed on ice for 2 minutes. The heat shocked cells were then mixed in 500 µL LB for 1 hour at 37°C with shaking and then spread onto an LB agar plate with the respective antibiotic. The plate was incubated at 37°C overnight for cell colonies to grow.

Plasmid Cloning -

The insert is amplified through PCR. Both the insert and the empty plasmid are digested using reaction enzymes with the following regents: 20 µL CutSmart Buffer, 10 µL plasmid or insert, 4 µL Xhol, 4 µL BamH, 62 µL Milli-Q for a total of 100 µL. Then the sample was cleaned using a PCR prep kit and only the plasmid was dephosphorylated using 2 µL of rSAP and incubated at 37°C for 30 minutes and then at 70°C for 20 minutes. The dephosphorylated plasmid sample was cleaned using a PCR Prep kit and then both the insert and vector were ligated using 2 µL of T4 DNA Ligase Buffer, 1 µL of the insert, 2 µL of the vector, 1 µL of T4 DNA Ligase, and 14 µL of Milli-Q. The ligation was run overnight and all 20 µL of the ligated plasmid sample was transformed into BL12 cells using the transformation protocol. Colonies from the transformation were cultured overnight in 4 mL and then plasmid prepped to extract the ligated plasmid and sent for DNA sequencing to confirm correct sequence.

Site-Directed Mutagenesis -

Primers were designed to either delete parts of a protein sequence or to create a mutation within a protein sequence by targeting specific sequences within a plasmid (more detail about the specific primers is in the results section). First the PCR protocol was used to amplify the specific part of the plasmid that was of interest. The amplified DNA was phosphorylated using 20 µL of Milli-Q, 2.5 µL of T4 DNA ligase buffer, 2 µL of the PCR mix, and 0.5 µL of T4 PNK. This reaction was incubated at 37°C for 30 minutes and then at 70°C for 20 minutes. The phosphorylated was then ligated using 13.5 µL of Milli-Q, 3.5 µL of the phosphorylated DNA, 2 µL of T4 DNA Ligase buffer, and 1 µL of DNA ligase. The ligation was run for 3-5 hours and all 20 µL of the ligated plasmid sample was transformed into BL12 cells using the

transformation protocol. Colonies from the transformation were cultured overnight in 4 mL and then plasmid prepped to extract the ligated plasmid and sent for DNA sequencing to confirm correct sequence.

Protein Purification -

Transformed cells were cultured into 1L and then spun down at 6000 rpm with a centrifuge to pellet the cells. The cell pellet is then resuspended into a 1x Lysis buffer. 300µL of Triton-X, 300µL of PMSF, and a dash of lysozyme is added to the cell resuspension which is then sonicated at 15 seconds on and 45 seconds off on ice. The resulting cell lysate is spun down at 15000 rpm with a centrifuge. The supertanted is then filtered through a 0.45µM filter into a cobalt matrix column. The column is then placed at 4^{o}C for 25 minutes with shaking. The solution then flows through the column and the matrix is washed several times (approx. 4) with 1x Lysis buffer and then the protein is eluted from the matrix using an elution buffer. The elution is then concentrated down to 5mL using the respective Amicon tube and loaded onto the FPLC with an FPLC buffer (20 mM Tris, 100mM NaCl, 1mM DTT, 2mM EDTA, pH 8.5) for size exclusion chromatography.

EMSA-

12 binding reactions of varying concentrations of protein were created into 20µL. The concentration of DNA remained the same while the concentration of protein increased. Varying amounts of 5x Binding Buffer (50 µL of 1 M Tris-HCl, pH 7.5; 10 µL of 5 M NaCl; 200 µL of 1 M KCl, 5 µL of 1 M MgCl2, 10 µL of 0.5 M EDTA, pH 8.0; 5 µL of 1 M DTT; 25 µL of 10 mg/mL BSA, and 695 µL of ddH2O [12]). and glycerol were added to maintain the 20µL total volume. Before loading onto a polyacrylamide gel, 4µL of 6x EMSA loading dye was added to solution and 20µL of the total solution was loaded into the gel. The gel was then run for 5-6

hours at 110V in a 1x TBE buffer (89 mM Tris base, 89 mM boric acid, 1 mM EDTA, pH ~8.0). The gel was then stained for DNA diluting 10000X Sybr Safe in 1x TBE buffer with 5µL and was left shaking for 45 minutes. The gel was then imaged under UV light.

Results:

Plasmids -

In order to carry out any structural studies on SHLD2's OB fold domains, plasmid construct encodes the sequence of the OB-folds were created.

A plasmid is a small circular double-stranded DNA that is separate from chromosomal DNA. Plasmids contain different parts as shown in Figure 1 below. The most important aspects of a plasmid relevant for this study are the inserted gene and the antibiotic resistance gene. The inserted gene encodes the sequence of a subdomain of SHLD2 C-terminus OB-folds. Because this gene is inserted downstream of the promoter, the gene will be expressed within the BL21 E. Coli cells that are transformed with the plasmid. The antibiotic resistance gene allows for the selection of cells that have this plasmid. Though cells are transformed to take up the plasmid, not all of them might have it, as a result growing cell cultures with the presence of an antibiotic that the plasmid confers

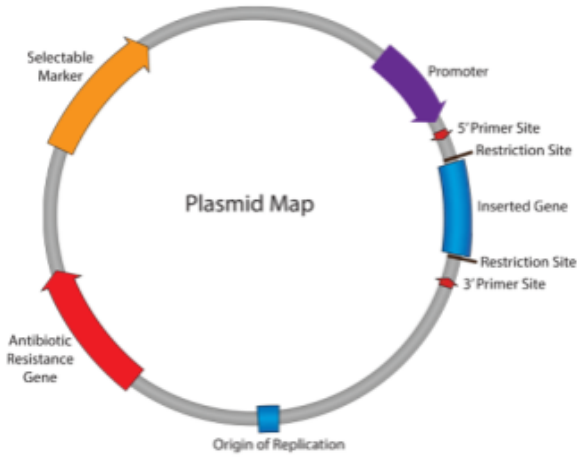

Figure 1 - Plasmid Schematic

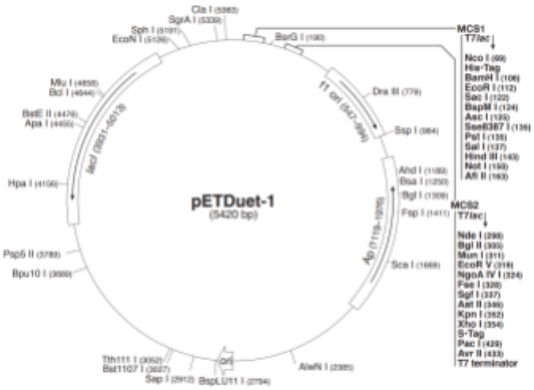

Figure 2 - pETDuet Vector

resistance to can kill any cells that do not have the plasmid.

The plasmid used in this study is the pETDuet vector shown in Figure 2. This plasmid has an ampicillin resistance gene allowing ampicillin to be used as a selective marker. The pETDuet vector expresses protein with a His-tag on the N-terminus (six histidines at the very beginning of the protein's primary sequence). The significance of this His-tag is discussed in the protein purification section of this report.

OB1:

The plasmid that encodes SHLD2 OB1 was already in the lab, however it was resequenced. In order to accomplish this the OB1 plasmid was transformed into Top10 E. Coli cells.

Colonies from the plate of transformed Top10 cells were cultured into 4 mL and plasmid prepped using the NucleoSpin Plasmid Prep Kit. The concentration of OB1 extracted was 968 ng/μL. To prepare samples for DNA sequencing, two samples each with a concentration of 80 ng/μL of plasmid were made. When reading the sequence of the plasmid, only the sequence of the inserted gene is of interest since this is where the sequence encoding OB1 is. To make sure not the whole plasmid is sequenced and only the inserted gene is, one of the two samples read the sequence in the forward (5'-3') direction and the other sample would read the sequence in the reverse (3'-5') direction. 5μL of a pETUpstream primer at a concentration of 166.7 μM was added to the forward sample. The sequence of the pETUpstream primer is 5' ATGCGTCCGGCGTAG 3' which binds right before the sequence of the OB1 gene allowing for sequencing results to be read in the 5'-3' direction. The T7-term primer is added to the reverse sample which binds near the end of the OB1 gene on the plasmid allowing for sequencing results to be read in the 3'-5' direction. The two samples were sent to Genewiz for sequencing results.

Figure 3a and 3b show the forward and reverse sequencing results respectively. Both sequences were translated into an amino acid sequence using the ExPASY translate tool. The amino acid sequence of the inserted gene is shown in Figure 3c. Using the blastp multiple protein sequence alignment software, this amino acid sequence was mapped to the range 427 to 596 on the full SHLD2 isoform 1 sequence.

a) >3_for-petUpstream_C09.ab1

NNNNNNNNNNNNNNNNCCNNNNNAATNNNTACNACTCACTATAGGNNNNNNGTGAGCGGATAACAATTCCCCTCTAGAAATAATNTTNTTTAACTTTAAGAAGGAGATATA
CCATGGGCAGCAGCCATCACCATCATCACCACAGCCAGGATCCAAAGAAAACGTCATTAATCAAGAACTGCGATTCAAAGAGCCAGAAGTACAATTGCCTGGTCATGGTAT
TGAGCCCTTGCCATGTAAAAGAGATTAACATTAAATTCGGACCAAATTCTGGCTCTAAAGTGCCCTTGGCTACTGTGACCGTAATCGATCAATCCGAAACCAAGAAAAAGG
TATTCCTTTGGCGCACCGCCGCATTCTGGGCGTTTACAGTCTTCCTTGGTGACATCATTTTGTTGACCGACGTCGTGATCCATGAAGATCAATGGATTGGAGAGACTGTAT
TACAATCCACCTTTTCCAGTCAATTGTTGAATCTTGGCTCCTATTCCTCCATTCAGCCTGAGGAATATTCGAGCGTGGTCTCCGAGGTAGTGTTGCAGGACCTTCTTGCAT
ATGTGTCTTCAAAGCATTCATACTTACGCTAACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCTTAATTAA
CCTAGGCTGCTGCCACCGCTGAGCAATAACTNGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCNGATTGGCGAATG
NNACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCANCGTGACCGCTACACTTGCCAGCGCCCTANCGCCCGCTCCTTTCGCTTNCTTCCCTTC
CTTTCTCGCCACNNTCGCCGGCTTTNCCCGNNNAGCTCNAAATCGGGGGCTCCCTTTAGGGTTCCNATTTANTGCTTTACGGCACCNCGACCCCAAAAAATTTGANTNAGG
GTGANGGTTCACGTAATGGGCCA

b) >4_rev-T7-Term_D09.ab1

NNNNNNNNNNNNNNNNNNNCTGCGCTAGTAGACNAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTNACCNGNNTCGAGTTAGCGTAAGTATGAATGCTT
TGAAGACACATATGCAAGAAGGTCCTGCAACACTACCTCGGAGACCACGCTCGAATATTCCTCAGGCTGAATGGAGGAATAGGAGCCAAGATTCAACAATTGACTGGAAAA
GGTGGATTGTAATACAGTCTCTCCAATCCATTGATCTTCATGGATCACGACGTCGGTCAACAAAATGATGTCACCAAGGAAGACTGTAAACGCCCAGAATGCGGCGGTGCG
CCAAAGGAATACCTTTTTCTTGGTTTCGGATTGATCGATTACGGTCACAGTAGCCAAGGGCACTTTAGAGCCAGAATTTGGTCCGAATTTAATGTTAATCTCTTTTACATG
GCAAGGGCTCAATACCATGACCAGGCAATTGTACTTCTGGCTCTTTGAATCGCAGTTCTTGATTAATGACGTTTTCTTTGGATCCTGGCTGTGGTGATGATGGTGATGGCT
GCTGCCCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGA
GATCNNTCTCGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTC
GCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGNTTGGGCGCCATCTCCTTGCATGCACCATTCCNTTGCGGCGGC
GGTGCTCAACGGNCTCAACCTACTACTGGGCTGCTTCCTAATGCAGAGTCGCATANNNNNAGCGTCNAGATCCCGGACACCATCNAATGGCGCAAAANCNTTCGCGGTANN
NGCATGATNNNCGCCCGGAANAAANTCAATTN

c)MGSSHHHHHHSQDPKKTSLIKNCDSKSQKYNCLVMVLSPCHVKEINIKFGPNSGSKVPLATVTVIDQSETKKKVFLWRTAAFWAFTVFLGDIILLTDVVIHEDQWIGET
VLQSTFSSQLLNLGSYSSIQPEEYSSVVSEVVLQDLLAYVSSKHSYLR

Figure 3 - OB1 plasmid Sequencing Results

Figure 4 shows the secondary structure prediction of SHLD2 from the residue range of 427 to 595. This prediction was done using the Quick2D set on the MPI bioinformatic Toolkit software. Quick2D uses ten algorithms to predict the amino acid locations of α-helices (red) and

β-sheets (blue) which are the two primary components of a protein's secondary structure.

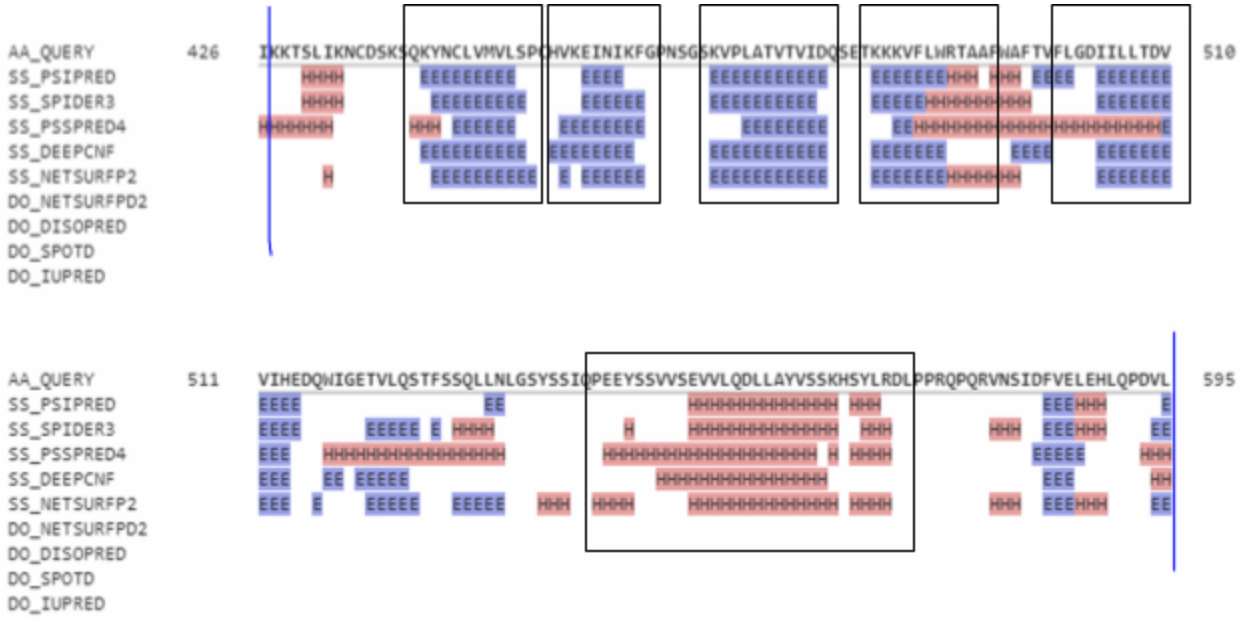

Figure 4 - Secondary Structure Prediction of OB1

From Figure 4, it is clear that there are five β-sheets followed by an α-helix marked down in a black box. These ranges are important because they represent the secondary structure components of an OB-fold at the general level which is a five-stranded β barrel with one end capped by an α-helix.

Both the sequencing results and the secondary structure prediction of the mapped region on SHLD2's amino acid sequence show that the plasmid correctly encodes the sequence of first DNA binding domain on SHLD2.

OB2:

Like OB1, the plasmid encoding SHLD2 OB2 was at the lab but was resequenced. The same process used for OB1 was used to sequence OB2. The concentration of the OB2 plasmid was 906 ng/μL and the DNA sequencing samples were prepared and sent to Genewiz the same way as OB1.

a) >8_rev-T7-Term_H09.ab1

NNNNNNNNNNNNNNNNNNNNNCTGCGCTAGTAGACGAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTTACCNTACTCGAGTTATGAGGGGGCGGATTTCT

GGAATTTAGCCTTGAAGGTAATAGCGCGGATATCGTCGTCGAAAAGGCATTCACAGCTAGACCAAGGGGTTGTGTGCAATTCCAAATTCTCCAAAGTGTAATTACATTGGA

CAAAAAGGTACTTAAACTCCCAAATGTATCCCTTTTTGCGTTGCAGTTGCGGATACCACGCTGCCCCAGGGCCCCACAGGACTAAGGCATAGTGTTGGTCCTGCGCCTGCT

CTACAGTCAGCATGACCTTTTTCTGCTTTTGTCCGCGATAGCTATAGACAGCTTCGGTAAGGATTGTGAAGTCTACGACGCGAAGAACAGCGTGAACCAGGACATCTGGCT

GCAGGTGTTCTAACTCCACGAAATCAATCGAATTAACTGGATCCTGGCTGTGGTGATGATGGTGATGGCTGCTGCCCATGGTATATCTCCTTCTTAAAGTTAAACAAAATT

ATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCGATCTCGATCCTCTACGCCGGACGCATCGTGGCCGGC

ATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATG

GTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCNTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATG

CANGAGTCGCATANGGNANAGCGTCGAGATCNNGGACACCATCGAATNGGCGCAAAACCTTTCGCGGTATGGCNTGNTANCGCCCGGAANANAGTCAATTCANGGTGGNNG

AATGTGNNN

b) MGSSHHHHHHSQDPVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQDQHYALVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLEN

LELHTTPWSSCECLFDDDIRAITFKAKFQKSAPS

Figure 5 - OB2 Plasmid Sequencing Results

Figure 5a shows the sequencing results. The forward sequencing results are not shown due to poor sequencing. The amino acid sequence translated from the DNA sequence, shown in Figure 5c, was mapped to the residues 579-707 on the full SHLD2 isoform 1 amino acid

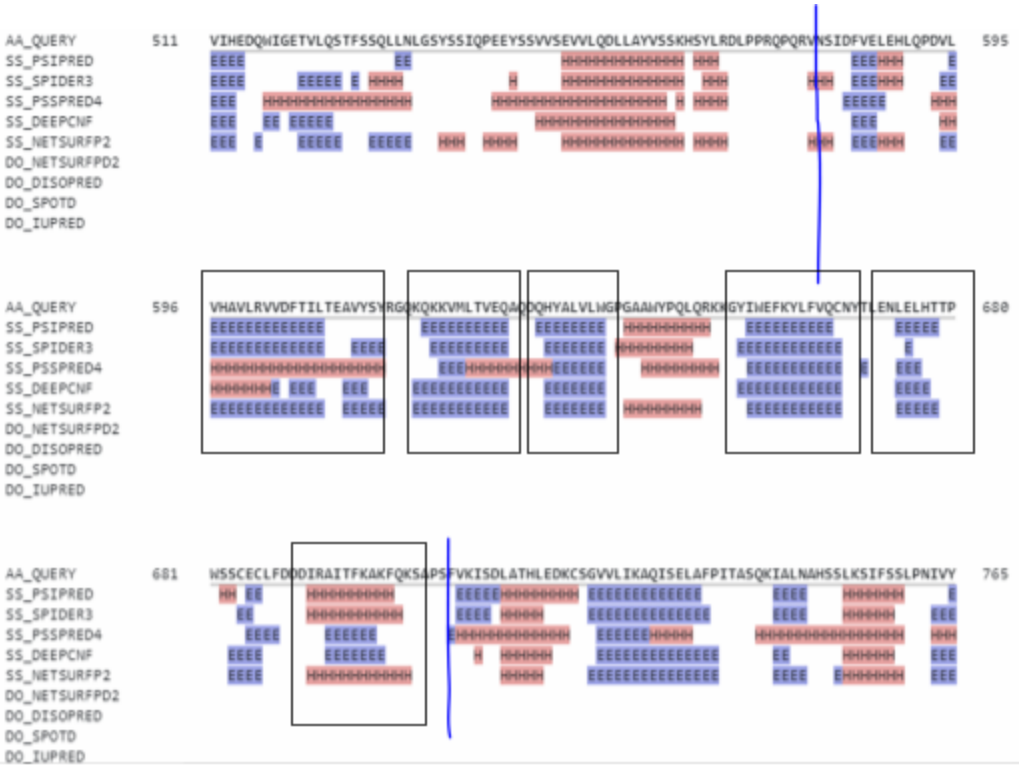

Figure 6 - OB2 Secondary Structure Prediction

sequence. The secondary structure prediction of this residue range is shown in Figure 6 clearly shows a β barrel end capped by an α-helix. All of which make up the structure of an OB-fold

This data shows that the plasmid correctly encodes the sequence of the second OB-fold of the SHLD2 subunit.

OB3:

Like OB1 and OB2, the plasmid encoding SHLD2 OB3 was at the lab but was resequenced. The same process used for OB1and OB2 was used to sequence OB3. The concentration of the OB3 plasmid was 818 ng/μL and the DNA sequencing samples were prepared and sent to Genewiz the same way as OB1 and OB2.

a) >23_for-petUpstream_G11.ab1

NNNNNNNNNNNNNNNNGCNNNNNNNTACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCAT

GGGCAGCAGCCATCACCATCATCACCACAGCCAGGATCCAGTCTACACAGGCTGCGCCAAGTGCGGGTTGGAGCTTGAAACTGATGAGAACCGTATCTATAAGCAATGTTT

TTCCTGCTTGCCATTCACGATGAAGAAAATTTATTACCGTCCGGCGTTGATGACCGCCATCGATGGTCGTCATGATGTGTGTATCCGTGTGGAGTCTAAATTGATTGAAAA

GATCCTTTTGAATATCTCCGCGGATTGCCTTAATCGCGTAATCGTCCCATCCAGCGAAATTACCTACGGTATGGTAGTTGCTGACTTGTTTCATAGCCTTCTGGCTGTTTC

TGCTGAACCATGCGTGTTGAAAATCCAGTCTTTGTTTGTTTTAGACGAGAATTCATATCCTCTTCAGCAGGATTTTAGTCTTCTTGACTTCTACCCCGATATCGTGAAGTA

ACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCTTAATTAACCTAGGCTGCTGCCACCGCTGAGCAATAACT

AGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCG

GCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTANCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGT

CAAGCTCTAAATCGGGGGCTCCCTTTANGGTTCCGATTTAGTGCTTTACGGCNCCTCGACCCCAAAAAACTTGATNNNGTGATGGN

b) >24_rev-T7-Term_H11.ab1

NNNNNNNNNNNNNANNAAGCTGCGCTAGTAGACGAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTNANNNGACTCGAGTTACTTCACGATATCGGGGTAG

AAGTCAAGAAGACTAAAATCCTGCTGAAGAGGATATGAATTCTCGTCTAAAACAAACAAAGACTGGATTTTCAACACGCATGGTTCAGCAGAAACAGCCAGAAGGCTATGA

AACAAGTCAGCAACTACCATACCGTAGGTAATTTCGCTGGATGGGACGATTACGCGATTAAGGCAATCCGCGGAGATATTCAAAAGGATCTTTTCAATCAATTTAGACTCC

ACACGGATACACACATCATGACGACCATCGATGGCGGTCATCAACGCCGGACGGTAATAAATTTTCTTCATCGTGAATGGCAAGCAGGAAAAACATTGCTTATAGATACGG

TTCTCATCAGTTTCAAGCTCCAACCCGCACTTGGCGCAGCCTGTGTAGACTGGATCCTGGCTGTGGTGATGATGGTGATGGCTGCTGCCCATGGTATATCTCCTTCTTAAA

GTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCGATCTCGATCCTCTACGCCGGACG

CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTT

TCGGCGTGGGTATGGTGGCAGGNCCCGTGGCCGGGGGACTGTTGGGCGCCATCNTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTG

GGCTGCTTCCTATNGCANGAGTCGCATAAGGGAGAGCGTCGAGATCCCGGACACCATCGAATGGCGCAAAACCTTTCNCGGTATNGNNN

c)MGSSHHHHHHSQDPVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSEITYGMVVADLFHSLL

AVSAEPCVLKIQSLFVLDENSYPLQQDFSLLDFYPDIVK

Figure 7 - OB3 Sequencing Results

Figure 7a and 7b show the forward and reverse sequencing results respectively. The amino acid sequence resulting from translating the DNA sequence shown in Figure 6c maps to the amino acid range of 764-897 of SHLD2.

Figure 8 shows the secondary structure prediction of this residue range. Though the key features of an OB-fold secondary structure are not as apparent as OB1 and OB2 it is not as obvious in OB3. The implications of this are further discussed during the purification of OB3.

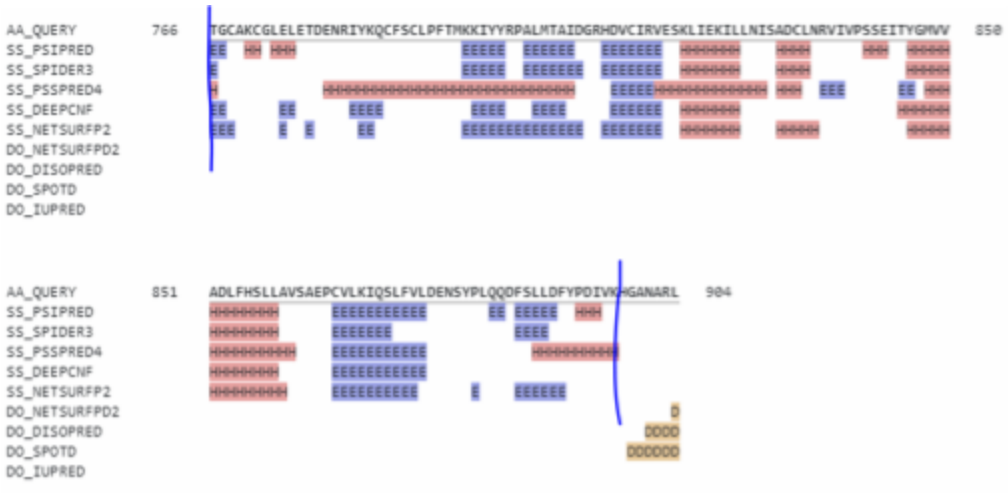

Figure 8 - OB3 Secondary Structure Prediction

OB123:

To make the plasmid that encodes the full C-terminus OB-folds of SHLD2, the plasmid cloning protocol described in the methods section was used. Figure 9a briefly describes how cloning works.

First an empty vector and a target gene are acquired. For the creation of the OB123 plasmid, an empty pETDuet vector was used and a dsDNA fragment that

Figure 9a - Plasmid Cloning Schematic

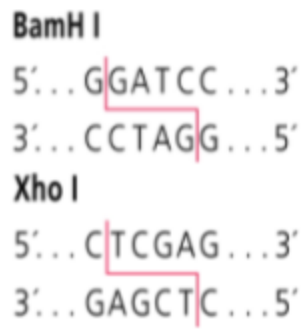

Figure 9b - BamHI and XhoI cutsites

encoded the sequence of SHLD2 OB123 was used as the target gene.

From Figure 2, it is clear that the pETDuet vector has many restriction enzyme cut sites. A restriction enzyme cleaves dsDNA at specific sites/sequences. The restriction enzymes used were BamHI and XhoI. Figure 9b shows the dsDNA cute site that these enzymes recognize and how they cut the dsDNA. BamHI and XhoI generate DNA sticky ends when cutting dsDNA. Sticky ends refers to the presence of a few unpaired nucleotides at the end of dsDNA. This is important because both the empty pETDuet vector and the OB123 DNA fragment each contain a BamHI and XhoI recognition site. As a result, the OB123 fragment and the empty pETDuet vector have compatible DNA sticky ends. Figure 12 better represents this crucial idea. When two fragments of DNA are cut by the same restriction enzyme their sticky ends are complementary.

To accomplish this the OB123 DNA insert was amplified through PCR, and the empty vector and insert were digested with BamHI and XhoI.

The plasmid only was then dephosphorylated using Shrimp Alkaline Phosphatase (rSAP). When digesting the plasmid, the 5' ends still remain phosphorylated, thus allowing the plasmid to re-circulize without the inserted gene. By dephosphorylating the plasmid, rSAP removes the phosphate groups attached to the 5' ends of the plasmid DNA.

However the phosphate groups still remain on the 5' ends of the OB123 fragment after digestion, the fragment can be inserted into the plasmid through the use of T4 DNA ligase.

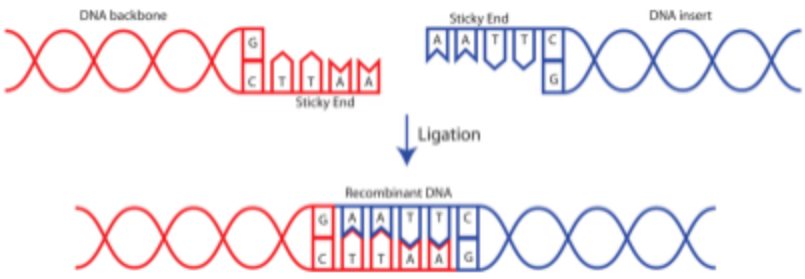

Figure 10 - DNA Ligation Schematic

Figure 10 shows how DNA ligase accomplishes this. DNA ligase forms a phosphodiester bond between the 3' hydroxyl

group of the plasmid and the 5’ phosphate of the insert.

a) >3_for-petUpstream_C05.ab1 (for)

NNNNNNGNNCNNNAATNNNACGACTCNCTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGC

AGCCATCACCATCATCACCACAGCCAGGATCCAAAGAAAACGTCATTAATCAAGAACTGCGATTCAAAGAGCCAGAAGTACAATTGCCTGGTCATGGTATTGAGCCCTTGC

CATGTAAAAGAGATTAACATTAAATTCGGACCAAATTCTGGCTCTAAAGTGCCCTTGGCTACTGTGACCGTAATCGATCAATCCGAAACCAAGAAAAAGGTATTCCTTTGG

CGCACCGCCGCATTCTGGGCGTTTACAGTCTTCCTTGGTGACATCATTTTGTTGACCGACGTCGTGATCCATGAAGATCAATGGATTGGAGAGACTGTATTACAATCCACC

TTTTCCAGTCAATTGTTGAATCTTGGCTCCTATTCCTCCATTCAGCCTGAGGAATATTCGAGCGTGGTCTCCGAGGTAGTGTTGCAGGACCTTCTTGCATATGTGTCTTCA

AAGCATTCATACTTACGCGATCTGCCGCCTCGCCAACCGCAACGCGTTAATTCGATTGATTTCGTGGAGTTAGAACACCTGCAGCCAGATGTCCTGGTTCACGCTGTTCTT

CGCGTCGTAGACTTCACAATCCTTACCGAAGCTGTCTATAGCTATCGCGGACAAAAGCAGAAAAAGGTCATGCTGACTGTAGAGCAGGCGCAGGACCAACACTATGCCTTA

GTCCTGTGGGGCCCTGGGGCAGCGTGGTATCCGCAACTGCAACGCAAAAAGGGATACATTTGGGAGTTTAAGTACCTTTTTGTCCAATGTAATTACACTTTGGAGAATTTG

GAATTGCACACAACCCCTTGGTCTAGCTGTGAATGCCTTTTCGACGACGATATCCGCGCTATTACCTTCAAGGCTAAATTCCAGAAATCCGCCCCCTCATTTGTAAAATCA

GCGATCTGGCANNACACCTNNNNNAANTGNTCCGNTGNNGTATNATCAAGCNCAGATTANCGAATTGGNGTTCCCGATTAN

b)MGSSHHHHHHSQDPKKTSLIKNCDSKSQKYNCLVMVLSPCHVKEINIKFGPNSGSKVPLATVTVIDQSETKKKVFLWRTAAFWAFTVFLGDIILLTDVVIHEDQWIGET

VLQSTFSSQLLNLGSYSSIQPEEYSSVVSEVVLQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQD

QHYALVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKSAIWXXTXXXXSXXVXSSXDXRIXVPDX

c) >4_rev-T7-Term_D05.ab1 (rev)

NNNNNNNNNNNNNNCTGCGCTAGTAGACGAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTTACCAGACTCGAGTTACTTCACGATATCGGGGTAGAAGTC

AAGAAGACTAAAATCCTGCTGAAGAGGATATGAATTCTCGTCTAAAACAAACAAAGACTGGATTTTCAACACGCATGGTTCAGCAGAAACAGCCAGAAGGCTATGAAACAA

GTCAGCAACTACCATACCGTAGGTAATTTCGCTGGATGGGACGATTACGCGATTAAGGCAATCCGCGGAGATATTCAAAAGGATCTTTTCAATCAATTTAGACTCCACACG

GATACACACATCATGACGACCATCGATGGCGGTCATCAACGCCGGACGGTAATAAATTTTCTTCATCGTGAATGGCAAGCAGGAAAAACATTGCTTATAGATACGGTTCTC

ATCAGTTTCAAGCTCCAACCCGCACTTGGCGCAGCCTGTGTAGACAATATTCGGCAAAGAGGAAAAAATGCTTTTAAGTGACGAGTGGGCGTTTAACGCGATCTTCTGACT

TGCGGTAATCGGGAACGCCAATTCGCTAATCTGCGCTTTGATCAATACAACACCGGAACATTTGTCCTCCAGGTGTGTTGCCAGATCGCTGATTTTTACAAATGAGGGGGC

GGATTTCTGGAATTTAGCCTTGAAGGTAATAGCGCGGATATCGTCGTCGAAAAGGCATTCACAGCTAGACCAAGGGGTTGTGTGCAATTCCAAATTCTCCAAAGTGTAATT

ACATTGGACAAAAAGGTACTTAAACTCCCAAATGTATCCCTTTTTGCGTTGCAGTTGCGGATACCACGCTGCCCCAGGGCCCCACAGGACTAAGGCATAGTGTTGGTCCTG

CGCCTGCTCTACAGTCAGCATGACCTTTTTCTGCTTTTGTCCGCGATAGCTATAGACAGCTTCGGTANGGATTGTGAAGTCTACGACGCGAAGAACAGCGTGAACCAGGAC

ATCTGGCTGCNGNNGTTCTAACTCCACGAAATCAATCGAATTAACGCGTTGCGGTTGGNGAGGCG

d)MLTVEQAQDQHYALVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSGVVLIKAQI

SELAFPITASQKIALNAHSSLKSIFSSLPNIVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSE

ITYGMVVADLFHSLLAVSAEPCVLKIQSLFVLDENSYPLQQDFSLLDFYPDIVK

e)MGSSHHHHHHSQDPKKTSLIKNCDSKSQKYNCLVMVLSPCHVKEINIKFGPNSGSKVPLATVTVIDQSETKKKVFLWRTAAFWAFTVFLGDIILLTDVVIHEDQWIGET

VLQSTFSSQLLNLGSYSSIQPEEYSSVVSEVVLQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQD

QHYALVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSGVVLIKAQISELAFPITASQ

KIALNAHSSLKSIFSSLPNIVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSEITYGMVVADLF

HSLLAVSAEPCVLKIQSLFVLDENSYPLQQDFSLLDFYPDIVK

Figure 11 - OB123 Sequencing Results

After the dephosphorylated plasmid was ligated with the OB123 insert, it was transformed into Top10 cells and colonies from the plate were cultured into 4 mL and plasmid prepped using the

NucleoSpin Plasmid Prep Kit. The concentration of the extracted OB123 plasmid was 91 ng/µL. The DNA sequencing samples were prepared the same way as OB1 and sent to Genewiz for results.

The sequence of the OB123 plasmid is shown in Figure 11. An interesting result seen in the sequencing result is that both the forward and reverse sequences (13a and 13c respectively) are different. This is because the DNA sequence is too long for both the pETUpstream and T7-term primers to read the whole sequence. Figure 11e shows the full sequence of the inserted gene which maps to the range 427-897 on the SHLD2 amino acid sequence. This range covers all three OB-fold domains thus confirming the successful creation of the OB123 plasmid.

OB12:

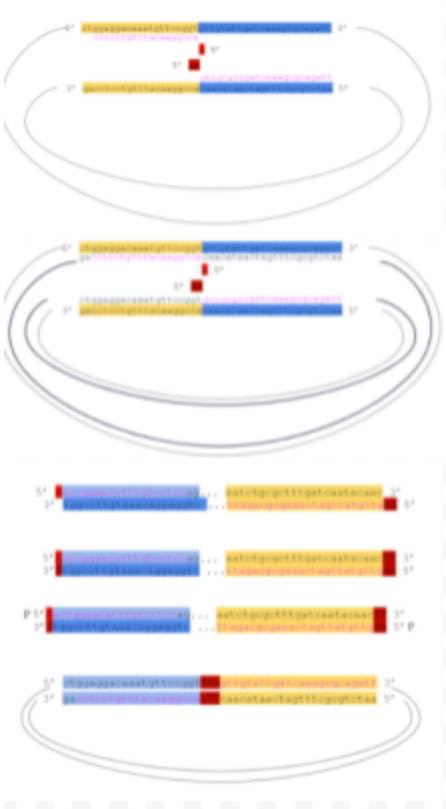

Figure 12 - Site-Directed Mutagenesis

The creation of the OB12 plasmid used a technique called site-directed mutagenesis. This process can add or delete DNA nucleotides from a plasmid. By being able to add or delete nucleotides, the original amino acid sequence of the plasmid is also altered because every three DNA nucleotides, called a codon, directly translates into a specific amino acid. Since the OB123 plasmid previously created has the sequence for the desired OB12 plasmid, this technique was used to add a stop codon to the end of OB12 on the OB123 plasmid (shown in red in Figure 12). By adding a stop codon (5' TAA 3') here, only OB12 will be expressed from the OB123 plasmid since the codon terminates before OB3 can be translated.

To accomplish this PCR was first used to amplify the DNA plasmid that , a forward (5’ AAgttgtattgatcaaagcgcagatt 3’) and reverse (5’ Aaccggaacatttgtcctcc 3’) primer were created both of which bound to the end of OB2 on the OB123 plasmid (shown in red in figure 12). The DNA sequence highlighted in yellow encoding the end of OB2 and the sequence highlighted in blue encoding the beginning of OB3. The primers are designed to bind at this junction. The most important characteristic of these primers is that the DNA sequence (stop-codon) that is being inserted at this junction **does not** bind to the DNA. However, these nucleotides that do not bind get incorporated into the overall plasmid sequence once the DNA is replicated. After the primers bind and are replicated within the PCR, the ends of the primers that initially do not bind are also replicated; this results in a long linear dsDNA strand. To make this linear DNA circular to form a plasmid the 5’ ends were phosphorylated using T4 PNK and ligated using T4 DNA ligase. In the final result, notice that the stop codon is right after the end of OB2 and right before the beginning of OB3, thus indicating that expression of the gene will be terminated after translating OB12.

a) >1_for-petUpstream_A01.ab1

NNNNNNNNNNNNNCNNNNNNNAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGG
GCAGCAGCCATCACCATCATCACCACAGCCAGGATCCAAAGAAAACGTCATTAATCAAGAACTGCGATTCAAAGAGCCAGAAGTACAATTGCCTGGTCATGGTATTGAGCC
CTTGCCATGTAAAAGAGATTAACATTAAATTCGGACCAAATTCTGGCTCTAAAGTGCCCTTGGCTACTGTGACCGTAATCGATCAATCCGAAACCAAGAAAAAGGTATTCC
TTTGGCGCACCGCCGCATTCTGGGCGTTTACAGTCTTCCTTGGTGACATCATTTTGTTGACCGACGTCGTGATCCATGAAGATCAATGGATTGGAGAGACTGTATTACAAT
CCACCTTTTCCAGTCAATTGTTGAATCTTGGCTCCTATTCCTCCATTCAGCCTGAGGAATATTCGAGCGTGGTCTCCGAGGTAGTGTTGCAGGACCTTCTTGCATATGTGT
CTTCAAAGCATTCATACTTACGCGATCTGCCGCCTCGCCAACCGCAACGCGTTAATTCGATTGATTTCGTGGAGTTAGAACACCTGCAGCCAGATGTCCTGGTTCACGCTG
TTCTTCGCGTCGTAGACTTCACAATCCTTACCGAAGCTGTCTATAGCTATCGCGGACAAAAGCAGAAAAAGGTCATGCTGACTGTAGAGCAGGCGCAGGACCAACACTATG
CCTTAGTCCTGTGGGGCCCTGGGGCAGCGTGGTATCCGCAACTGCAACGCAAAAAGGGATACATTTGGGAGTTTAAGTACCTTTTTGTCCAATGTAATTACACTTTGGAGA
ATTTGGAATTGCACACAACCCCTTGGTCTAGCTGTGAATGCCTTTTCGACGACGATATCCGCGCTATTACCTTCAAGGCTAAATTCCAGAAATCCGCCCNCT

b)MGSSHHHHHHSQDPKKTSLIKNCDSKSQKYNCLVMVLSPCHVKEINIKFGPNSGSKVPLATVTVIDQSETKKKVFLWRTAAFWAFTVFLGDIILLTDVVIHEDQWIGET
VLQSTFSSQLLNLGSYSSIQPEEYSSVVSEVVLQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQD
QHYALVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAX

Iso1: 427 - 705

c)NNNNNNNNNNNNNNNANCTGCGCTAGTAGACGAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTTACCAGACTCGAGTTACTTCACGATATCGGGGTAG
AAGTCAAGAAGACTAAAATCCTGCTGAAGAGGATATGAATTCTCGTCTAAAACAAACAAAGACTGGATTTTCAACACGCATGGTTCAGCAGAAACAGCCAGAAGGCTATGA
AACAAGTCAGCAACTACCATACCGTAGGTAATTTCGCTGGATGGGACGATTACGCGATTAAGGCAATCCGCGGAGATATTCAAAAGGATCTTTTCAATCAATTTAGACTCC
ACACGGATACACACATCATGACGACCATCGATGGCGGTCATCAACGCCGGACGGTAATAAATTTTCTTCATCGTGAATGGCAAGCAGGAAAAACATTGCTTATAGATACGG

TTCTCATCAGTTTCAAGCTCCAACCCGCACTTGGCGCAGCCTGTGTAGACAATATTCGGCAAAGAGGAAAAAATGCTTTTAAGTGACGAGTGGGCGTTTAACGCGATCTTC

TGACTTGCGGTAATCGGGAACGCCAATTCGCTAATCTGCGCTTTGATCAATACAACTTAACCGGAACATTTGTCCTCCAGGTGTGTTGCCAGATCGCTGATTTTTACAAAT

GAGGGGGCGGATTTCTGGAATTTAGCCTTGAAGGTAATAGCGCGGATATCGTCGTCGAAAAGGCATTCACAGCTAGACCAAGGGGTTGTGTGCAATTCCAAATTCTCCAAA

GTGTAATTACATTGGACAAAAAGGTACTTAAACTCCCAAATGTATCCCTTTTTGCGTTGCAGTTGCGGATACCACGCTGCCCCAGGGCCCCACAGGACTNANNNATAGTGT

TGGTCCTGCGCCTGCTCTACAGTCAGCATGACCTTTTTCTGCTTTTGTCCGCGATAGCTATAGACAGCTTCGGTAAAGGATTGTGAAGTCTACGACGCGAAGAACNN

d) MLTVEQAQDQHYXXVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSG

Iso1: 624 - 724

e)MGSSHHHHHHSQDPKKTSLIKNCDSKSQKYNCLVMVLSPCHVKEINIKFGPNSGSKVPLATVTVIDQSETKKKVFLWRTAAFWAFTVFLGDIILLTDVVIHEDQWIGET

VLQSTFSSQLLNLGSYSSIQPEEYSSVVSEVVLQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQD

QHYXXVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSG

Figure 13 - OB12 Sequencing Results

Figure 13 shows the sequencing of the plasmid after it was extracted from a 4mL culture. Like the sequencing results from the OB123 plasmid, the forward and reverse sequencing results overlap with Figure 16e showing the full amino acid sequence. The sequence maps to the range 427-724 on the full SHLD2 sequence, thus confirming the successful creation of the OB12 plasmid.

OB23:

Site-directed mutagenesis was also used in the creation of the OB23 plasmid. Conceptually it is very similar however instead of adding DNA nucleotides to the plasmid, nucleotides that encoded the OB1 domain were deleted from the OB123 plasmid through specifically designed primers.

The forward primer (5’ CGCGATCTGCCGCCTCG 3’) binds to the OB123 plasmid at the end of OB1 and the reverse primer (5’ CTGGCTGTGGTGATGATGGTGATGGC 3’) binds to the OB123 plasmid to the beginning of OB1. The same process discussed above occurs which effectively deletes OB1 and the resulting plasmid encoded OB23.

a) NNNNNNNNNNNCGCNANNTNATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGG
GCAGCAGCCATCACCATCATCACCACAGCCAGGATCAGGACCTTCTTGCATATGTGTCTTCAAAGCATTCATACTTACGCGATCTGCCGCCTCGCCAACCGCAACGCGTTA
ATTCGATTGATTTCGTGGAGTTAGAACACCTGCAGCCAGATGTCCTGGTTCACGCTGTTCTTCGCGTCGTAGACTTCACAATCCTTACCGAAGCTGTCTATAGCTATCGCG
GACAAAAGCAGAAAAAGGTCATGCTGACTGTAGAGCAGGCGCAGGACCAACACTATGCCTTAGTCCTGTGGGGCCCTGGGGCAGCGTGGTATCCGCAACTGCAACGCAAAA
AGGGATACATTTGGGAGTTTAAGTACCTTTTTGTCCAATGTAATTACACTTTGGAGAATTTGGAATTGCACACAACCCCTTGGTCTAGCTGTGAATGCCTTTTCGACGACG
ATATCCGCGCTATTACCTTCAAGGCTAAATTCCAGAAATCCGCCCCCTCATTTGTAAAAATCAGCGATCTGGCAACACACCTGGAGGACAAATGTTCCGGTGTTGTATTGA
TCAAAGCGCAGATTAGCGAATTGGCGTTCCCGATTACCGCAAGTCAGAAGATCGCGTTAAACGCCCACTCGTCACTTAAAAGCATTTTTTCCTCTTTGCCGAATATTGTCT
ACACAGGCTGCGCCAAGTGCGGGTTGGAGCTTGAAACTGATGAGAACCGTATCTATAAGCAATGTTTTTCCTGCTTGCCATTCACGATGAAGAAAATTTATTACCGTCCGG
CGTTGATGACCGCCATCGATGGTCGTCATGATGTGTGTATCCGTGTGGAGTCTAAATTGATTGAAAAGATCCTTTTGAATATCTCCGCGGATTGCCTTAATCGCGTAATCG
TCCCATCCAGCGAAATTACCTACGGTATGGNAGTTGCTGACTTGNTTCNNAGCCTTCTGGCTGNTTNNTGCNN

b) MGSSHHHHHHSQDQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQDQHYALVLWGPGAAWYPQL
QRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSGVVLIKAQISELAFPITASQKIALNAHSSLKSIFSSLP
NIVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSEITYGMXVADLXXSLLAXXC

c) NNNNNNNNNNNNNNAGCTGCGCTAGTAGACGAGTCCATGTGCTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTTACCAGACTCGAGTTACTTCACGATATCGGGGTAGA
AGTCAAGAAGACTAAAATCCTGCTGAAGAGGATATGAATTCTCGTCTAAAACAAACAAAGACTGGATTTTCAACACGCATGGTTCAGCAGAAACAGCCAGAAGGCTATGAA
ACAAGTCAGCAACTACCATACCGTAGGTAATTTCGCTGGATGGGACGATTACGCGATTAAGGCAATCCGCGGAGATATTCAAAAGGATCTTTTCAATCAATTTAGACTCCA
CACGGATACACACATCATGACGACCATCGATGGCGGTCATCAACGCCGGACGGTAATAAATTTTCTTCATCGTGAATGGCAAGCAGGAAAAACATTGCTTATAGATACGGT
TCTCATCAGTTTCAAGCTCCAACCCGCACTTGGCGCAGCCTGTGTAGACAATATTCGGCAAAGAGGAAAAAATGCTTTTAAGTGACGAGTGGGCGTTTAACGCGATCTTCT
GACTTGCGGTAATCGGGAACGCCAATTCGCTAATCTGCGCTTTGATCAATACAACACCGGAACATTTGTCCTCCAGGTGTGTTGCCAGATCGCTGATTTTTACAAATGAGG
GGGCGGATTTCTGGAATTTAGCCTTGAAGGTAATAGCGCGGATATCGTCGTCGAAAAGGCATTCACAGCTAGACCAAGGGGTTGTGTGCAATTCCAAATTCTCCAAAGTGT
AATTACATTGGACAAAAAGGTACTTAAACTCCCAAATGTATCCCTTTTTGCGTTGCAGTTGCGGATACCACGCTGCCCCAGGGCCCCACAGGACTNANNNATAGTGTTGGT
CCTGCGCCTGCTCTACAGTCAGCATGACCTTTTTCTGCTTTTGTCCGCGATAGCTATAGACAGCTTCGGTAAGNATTGTGAAGTCTACGACGCGAAGAACAGCGTGAACCA
GGACATCTGGCTGNN

d) MLTVEQAQDQHYXXVLWGPGAAWYPQLQRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSGVVLIKAQI
SELAFPITASQKIALNAHSSLKSIFSSLPNIVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSE
ITYGMVVADLFHSLLAVSAEPCVLKIQSLFVLDENSYPLQQDFSLLDFYPDIVK

e) MGSSHHHHHHSQDQDLLAYVSSKHSYLRDLPPRQPQRVNSIDFVELEHLQPDVLVHAVLRVVDFTILTEAVYSYRGQKQKKVMLTVEQAQDQHYXXVLWGPGAAWYPQL
QRKKGYIWEFKYLFVQCNYTLENLELHTTPWSSCECLFDDDIRAITFKAKFQKSAPSFVKISDLATHLEDKCSGVVLIKAQISELAFPITASQKIALNAHSSLKSIFSSLP
NIVYTGCAKCGLELETDENRIYKQCFSCLPFTMKKIYYRPALMTAIDGRHDVCIRVESKLIEKILLNISADCLNRVIVPSSEITYGMVVADLFHSLLAVSAEPCVLKIQSL
FVLDENSYPLQQDFSLLDFYPDIVK

Figure 14 - OB23 Sequencing Results

Figure 14a and 14b shows the sequencing results for the forward and reverse primers. 14e shows the translated sequence and it maps to 555-897 of SHLD2 which contains both OB2 and OB3, thus confirming the successful creation of the OB23 plasmid.

Protein Purification:

OB1:

The purpose of protein purification is to isolate a protein of interest from all other cell material. In order to accomplish this, the plasmid that encodes the protein, in this case OB1, is used. Bacterial cells will take plasmid and express it to produce the protein. One reason why this is important is because human OB1 is not naturally expressed in bacterial cells, thus using a plasmid allows for the production of OB1. More importantly, the OB1 plasmid expresses OB1 with a His-tag on its N-terminus (six histidines). The sequencing results for OB1 shows this in Figure 3. The significance of this will be discussed shortly.

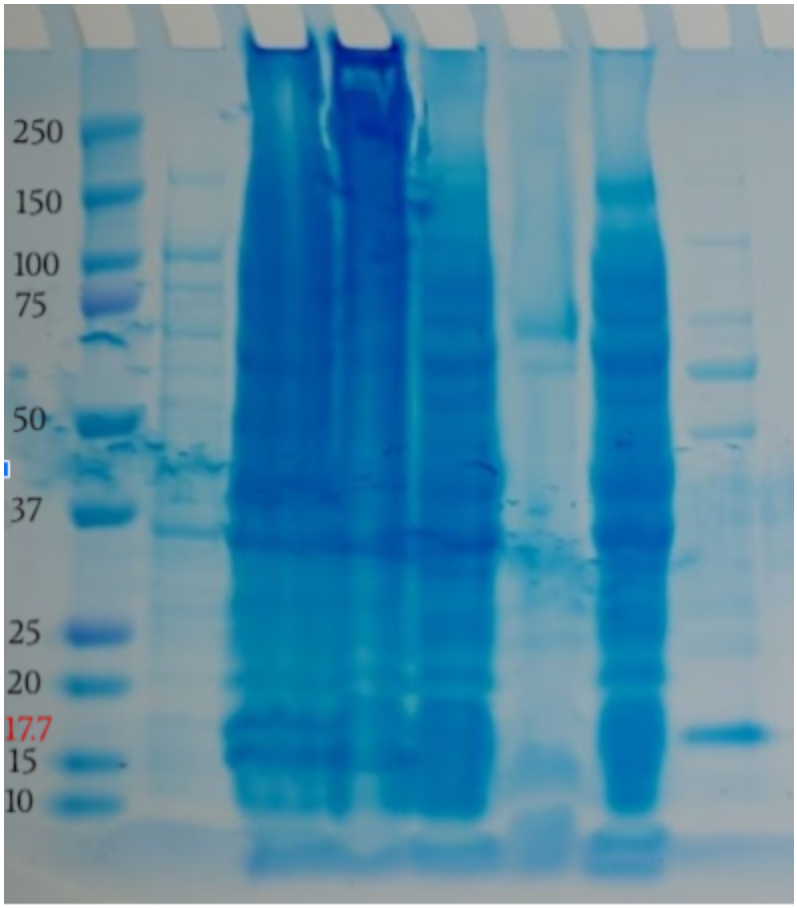

Figure 15 - OB1 Purification

BL21 E. Coli cells were transformed with the pETDuet-OB1 plasmid. BL21 cells are used instead of Top10 cells because BL21 cells transcribe a protein called T7 RNA polymerase [13]. This polymerase helps express other genes, like that encoded on the plasmid, 8 times faster than normal E. Coli cells [13]. Furthermore that gene that encodes T7 RNA polymerase is

under control of the lacUV5 promoter (P lacUV5), which is a strong variant of the wild-type lac promoter [13], thus allowing for tight control over when the OB1 plasmid is expressed.

A colony of the transformed BL21 cells were cultured into 1 L of litrobroth until the OD600 reading reached 0.8 at which point protein expression was induced with isopropyl-β-D-1-thiogalactopyranoside IPTG. By adding IPTG, the gene encoding T7 RNA polymerase is expressed and the polymerase recognizes the T7 promoters on the pETDuet vector, thus initiating the expression of the protein, OB1.

After induction, the cells were grown overnight and spun down into a cell pellet using a centrifuge. The cells were then resuspended into 45 mL of 1x lysis buffer. At this point, the cells should have the OB1 domain expressed. This idea is illustrated in Figure 15. This figure shows an SDS-PAGE gel of the different steps of protein purification described in the methods section. An SDS-PAGE gel is used to separate proteins based on molecular weight, thus allowing for the detection of specific proteins. Since the OB1 domain is 17.7 kDa, presence of OB1 would be seen in the SDS-PAGE gel at that molecular weight. Lane 3 shows the gel sample for the cell culture before adding IPTG. There is clearly no band at the 17.7 kDa mark as expected since protein expression was not induced. However Lane 4 shows a sample of the cell culture once IPTG is added and resuspended in the 1x lysis buffer. There is protein at the 17.7 kDa mark which confirms protein induction.

After the cells were resuspended, they were sonicated. Sonication rupures cell membranes thus allowing for the cell proteins to become released in solution. Lane 5 of the gel in figure 15 shows this since there seems to be more protein in the lane, and protein is present at the 17.7 kDa mark showing that OB1 is present in the sonicated cell lysate.

This cell lysate was then spun down using a centrifuge set to 15000 rpm for 1 hr. By spinning down this lysate, the insoluble and soluble fractions are separated by pelting the insoluble cell material. This insoluble material includes phospholipids or even insoluble/aggregated protein. Lanes 6 and 7 of the gel show the soluble and insoluble fractions of the spun down cell lysate. Seeing that there is a lot of protein present in the soluble fraction (especially at the 17.7 kDa mark) shows that the OB1 protein is in the soluble fraction. This is important because it allows for the selective purification of OB1 from the soluble fraction. OB1 is expressed with 6 histidines bonded to the N-terminus of its amino acid sequence (refer to sequencing results in Figure 3). No other protein in the soluble cell lysate has this type of tag thus allowing the usage of a method that selects for proteins with the His-tag which is only the OB1 protein.

The method used is called Immobilized Metal Affinity Chromatography (IMAC). IMAC is used because purification of His-tagged proteins by IMAC is based on the affinity of histidine residues for immobilized metal ions. The cobalt metal ion, Co^{2+}, is immobilized onto a chromatographic matrix, as a result, any His-tag that is in the presence of this matrix will bind to the Co^{2+} since histidine has a high affinity with immobilized metal ion matrices, as electron donor groups on the histidine imidazole ring readily form coordination bonds with the immobilized transition metal (Co2+) [14]. Using a Co2+ resin can capture all the OB1 proteins since they all have a His-tag that can bind to the matrix whereas other proteins may not be able to.

To accomplish this, the soluble cell lysate was filtered onto a purification column with 5 mL of Co^{2+} resin with a 0.45 micrometer filter. Then the filtered solution was incubated with the Co^{2+} resin at 4 celsius for 25 minutes with shaking. This allows the his-tagged OB1 to bind to the

matrix. After the 25 min incubation, the remaining solution is allowed to flow through the matrix. This flow through contains all the soluble proteins that do not bind to the cobalt matrix. It is however possible to have proteins bind to the matrix through nonspecific interactions. This is because proteins containing consecutive histidines or metal-binding motifs that are not OB1 can bind to the matrix. To get rid of these proteins the matrix is washed multiple times with the 1x lysis buffer which contains 0.01M of imidazole. Imidazole has a higher affinity for Co^{2+} than histidine, thus displacing these nonspecifically bound proteins from the matrix. This imidazole concentration is too low to displace much of the bound OB1 since the protein is over expressed. To elute OB1 off the column, an elution buffer with 0.5M concentration of imidazole is used. This effectively displaces much of the OB1 to collect.

Lanes 7 and 8 in figure 15 show a gel sample on the flowthrough and elution. In the flowthrough, lane 7, there are a lot of proteins since these are all proteins that do not bind to the matrix, however in the elution, lane 8, there is only one clear band. This band is at the 17.7 kDa mark which is where the OB1 protein should be. Now, the OB1 protein has been isolated from the cell material, however it is clear that the elution (which contains OB1) is contaminated by other proteins. In lane 8 there are multiple bands of proteins not just at the 17.7 kDa mark. To get rid of these proteins of different sizes another technique called size exclusion chromatography (SEM) is used.

SEM is a method which separates the proteins within the elution based on size. To accomplish SEM, Fast Liquid Protein Chromatography (FPLC) was used. The FPLC machine is able to separate a protein mixture based on size because one of its major components is a column that consists of a matrix of beads in which the spaces between them gets smaller further down.

As a result, bigger proteins come off the column first while smaller proteins come off later since they travel further through the matrix.

To use the FPLC machine a FPLC buffer, specific for the protein, was attached to the machine (this is the buffer the proteins eluted off of). The protein elution was then concentrated down to 5 mL using the and then injected into the FPLC machine. As the machine runs, it separates out the proteins within the 5 mL. To quantify the amount of protein coming off the column as time passes, the separated sample runs through a part of the machine that measures UV absorbance at 280 nm. At this wavelength, aromatic residues within the proteins absorb the most amount of light. Thus, the amount of absorbance at 280 nm is directly related to the amount of protein eluted.

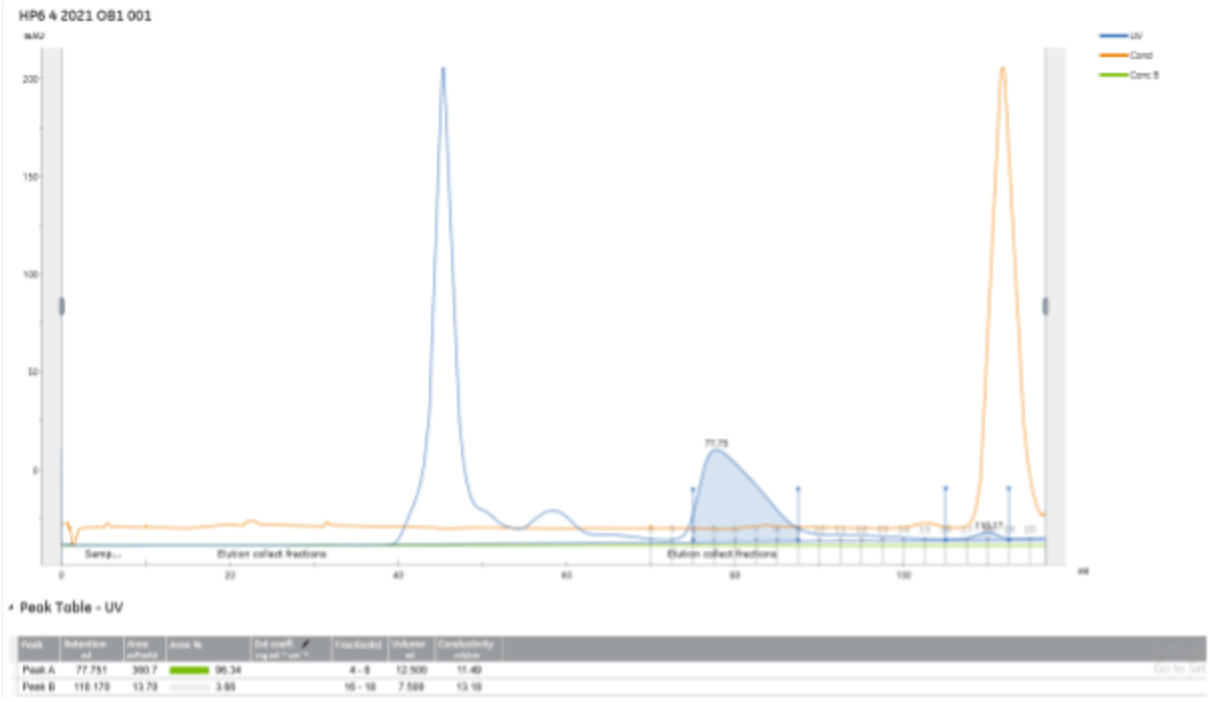

Figure 16 - FPLC Chromatogram of OB1

Figure 16 shows the UV absorbance chromatogram of the FPLC run for OB1. The very first peak of proteins coming off is the aggregate peak which includes any aggregated OB1 and very

large proteins. The elution at 80mL is where OB1 came off the column. The elution was collected in 2mL fractions. Figure 17 shows an SDS-PAGE gel of the aggregate peak and the elution at 80 mLs. The gel crealy shows that the elution contains OB1 since there are multiple bands at the 17.7 kDa mark.

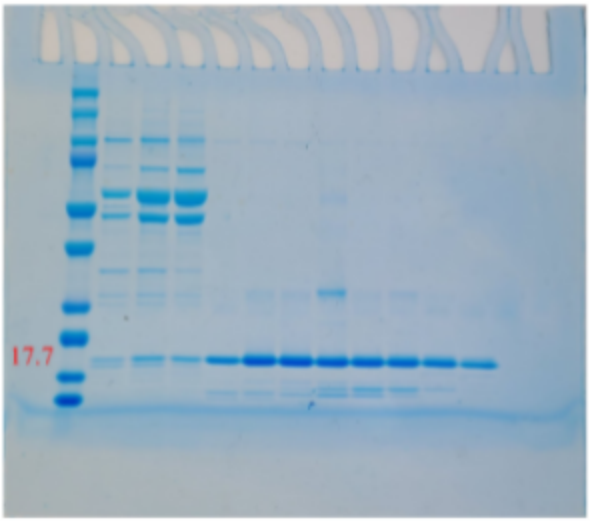

Figure 17 - OB1 FPLC Gel

To calculate the concentration of OB1 that was eluted. The protein was diluted in 8M Guanidine Hydrochloride in a 1:2 ratio and the absorbance at 280 nm was captured. The molar concentration of protein was then calculated using Beer-Lambert's law: A=ecl. The concentration of OB1 is 12µM.

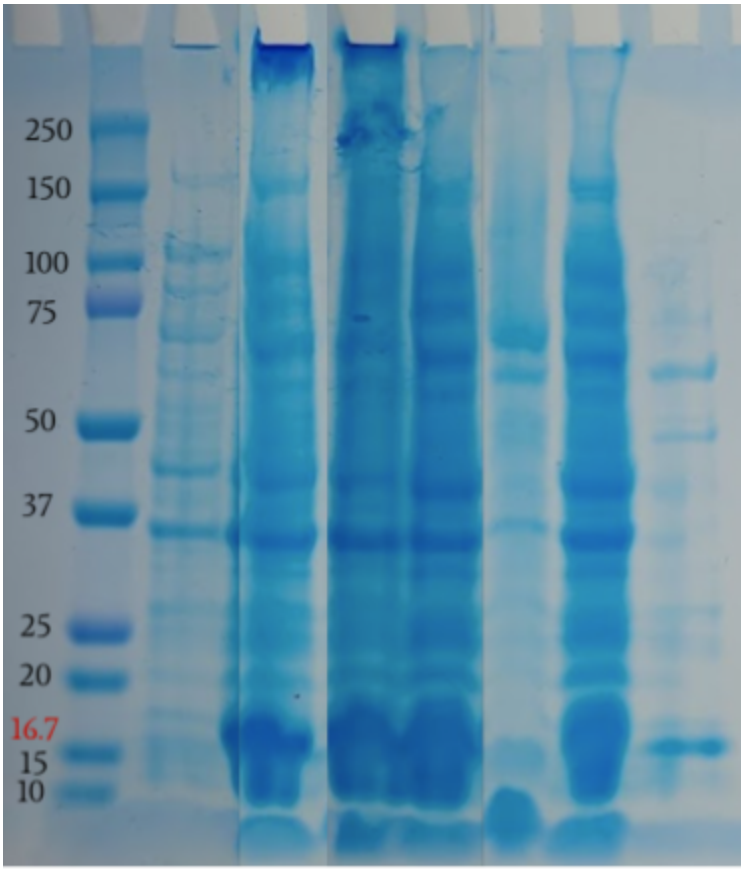

Figure 18 - OB2 Purification

OB2:

The same process to purify OB1 was used to purify OB2 as well. Figure 18 shows the SDS-PAGE gel of the different steps within the purification process. OB2 is 16.7 kDa, and the protein is present in the post induction (lane 3), sonication (lane 4), soluble fraction (Lane 5), and the elution (lane 8).

Figure 19 shows the FPLC chromatogram of OB2, the protein elutes at around 80 mL just like OB1 which makes sense since both proteins are very similar in size.

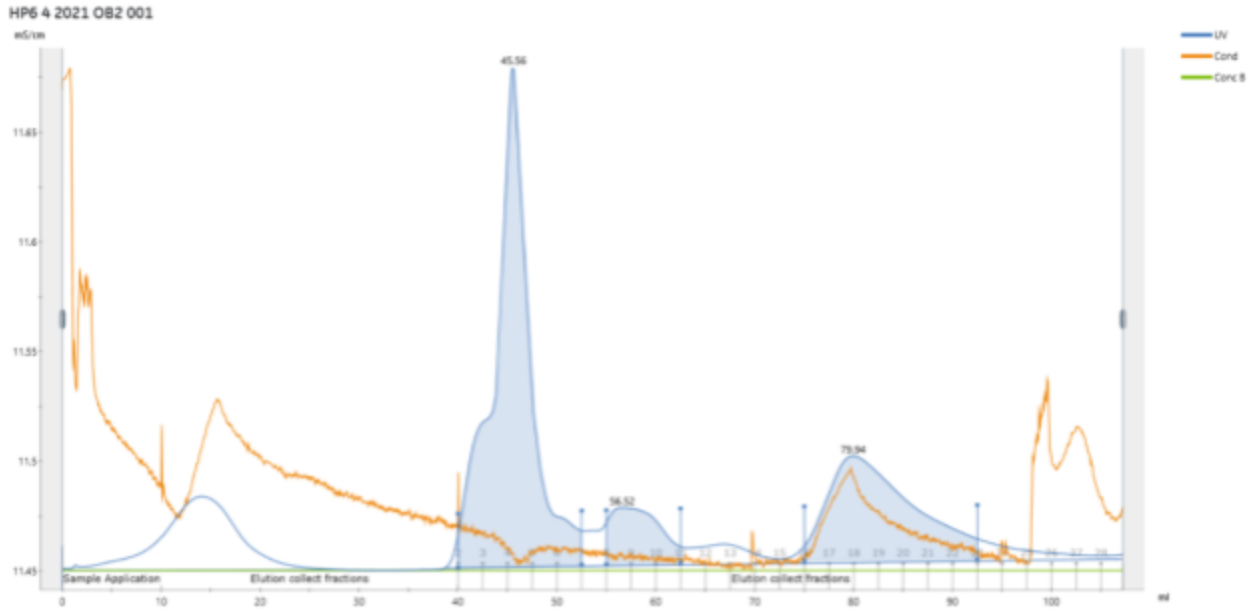

Figure 19 - OB2 FPLC Chromatogram

An SDS-PAGE gel of the aggregate peak and elution is shown in figure 21. The fractions eluted at 80mL show that OB2 has been successfully separated from the rest of the solution.

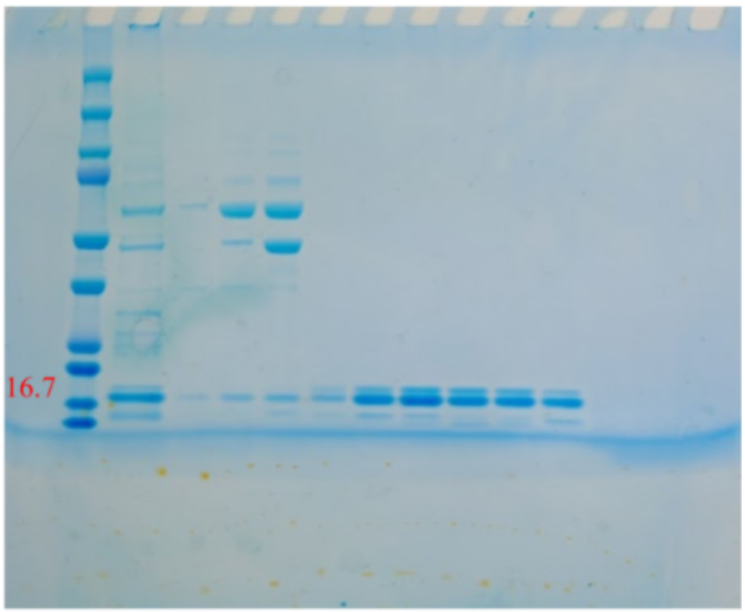

Figure 20 - OB2 FPLC Gel

The concentration of OB2 was calculated the same way as OB1 and is 11μM.

OB3:

OB3 was first purified using the method described above. Figure 21 shows the SDS-PAGE gel for the purification of OB3. An interesting

thing to note about this figure is the OB3 is present in the post induction sample (lane 3) and the sonication sample (lane 4) which is as expected. However OB3 is not present in the soluble or insoluble fractions of the cell latsaye (lanes 5 & 6). The most possible explanation for this is that OB3 is actually in the insoluble fraction. Running a gel sample for the insoluble fraction is tough to go since the insoluble pellet needs to be resuspended in a buffer. Since it is insoluble this is difficult to do. But the fact that OB3 was not present in the soluble fraction shows that it most likely precipitated out of solution. Because OB3 is a sub-domain, it could be that it does not fold properly, and it aggregates out of solution.

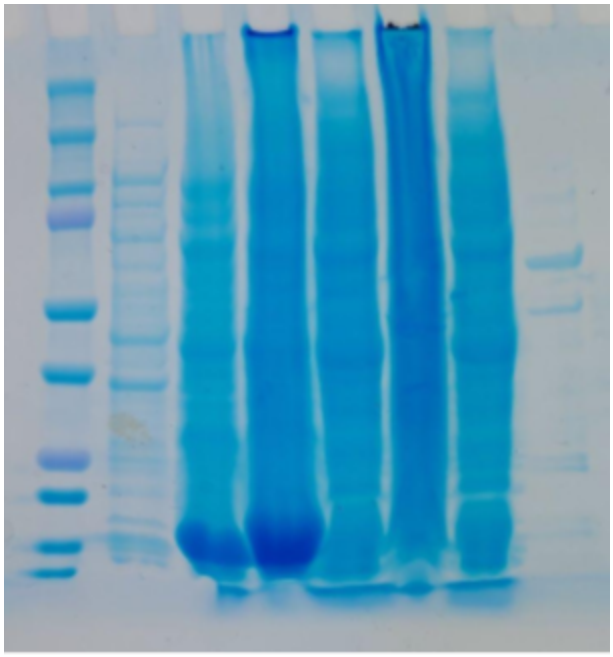

Figure 21 - OB3 Purification

A potential fix for this problem is to go about purifying OB3 using a different method. This new method would involve denaturing OB3 and then purifying it out of solution. Conceptually, this is very similar to a normal purification since a denatured protein still has an exposed His-tag that can bind to the Co2+ resin. Furthermore, the same buffers are used, however both the 1x Lysis Buffer and elution buffer contained 6M GuHCl. GuHCl denature

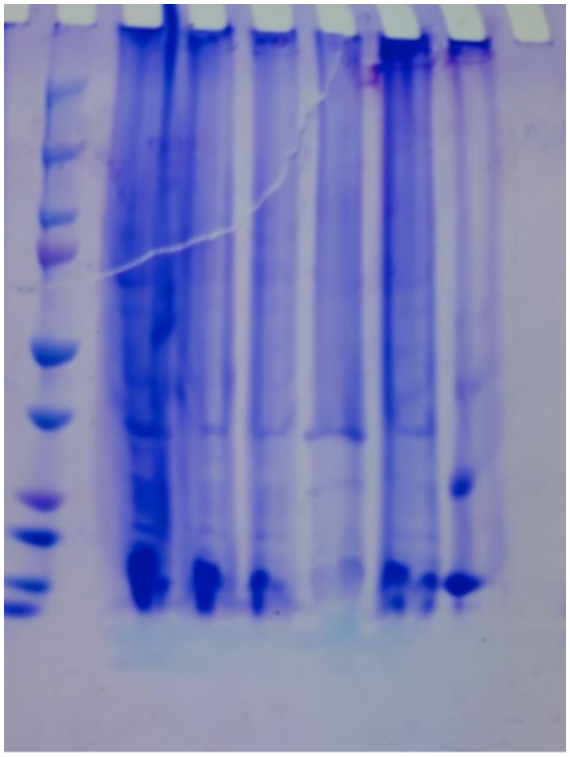

Figure 22 - OB3 Purification with 6M GuHCl

proteins by disputing the hydrogen bonds that hold a protein's secondary structure. As a result, denatured proteins are much more soluble. Figure 22 shows an SDS-PAGE gel of the purification of OB3 under denaturing conditions. OB3 is now present in all the steps it would be expected to

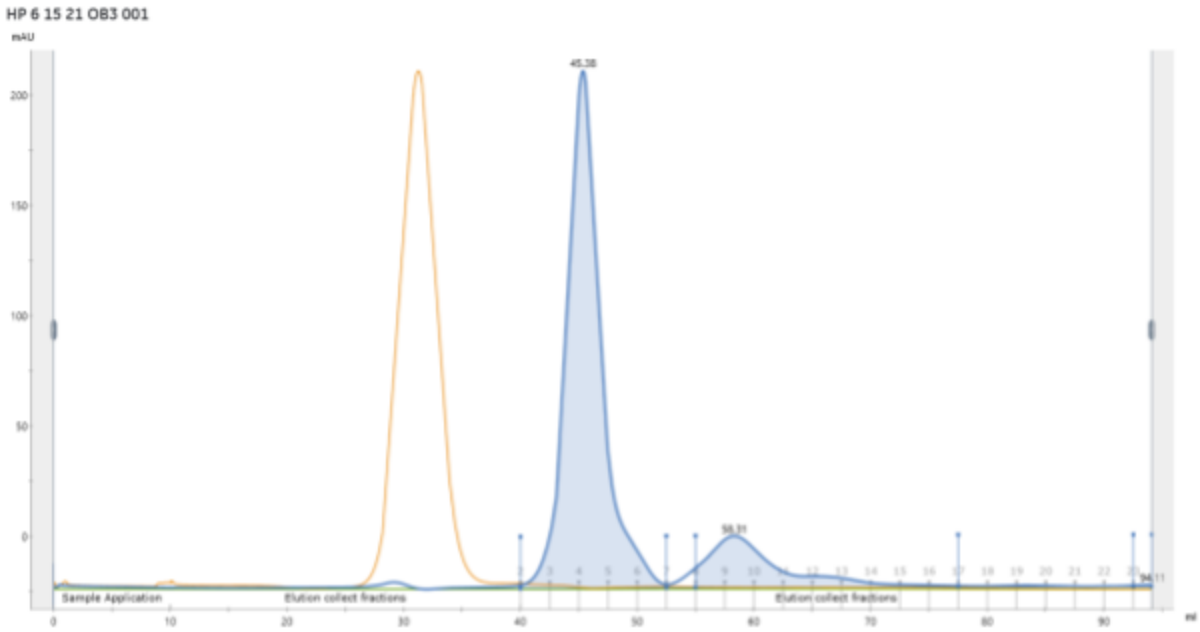

Figure 23 - OB3 FPLC Chromatogram

be in. OB3 is present in post induction sample (lane 2), sonication (lane 3), soluble cell lysate fraction (lane 4), and the lution (lane 7). All of these bands are at the 16.8 kDa mark which is the size of OB3.

Even though OB3 was successfully eluted it is still denatured. To renature OB3, the guanidine was slowly removed from solution using a dialysis tubing.

The elution was then concentrated to 5mL and then loaded onto the FPLC machine for size exclusion. Figure 23 shows the UV absorbance chromatogram of the FPLC run using the OB3 elution from denaturing purification conditions. OB3 should have eluted at around 80 mL (since this is where proteins the size of OB3 elute on

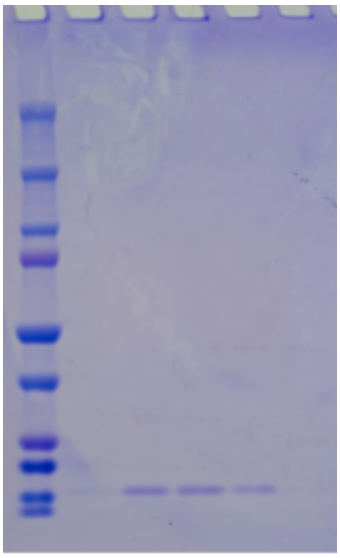

Figure 24 - OB3 FPLC Gel

this FPLC column, however there is no peak there. Figure 24 shows an SDS-PAGE gel of the aggregate peak and clearly OB3 is present within this peak. Since OB3 aggregates under normal and denaturing conditions OB3 is too unstable to be purified.

OB12:

OB12 was purified under normal purification conditions. Figure 25 shows the SDS-PAGE gel of the different steps within the purification process. OB12 is a 35.7 kDa protein and boxed in red, figure 25 shows the presence of OB12 in each step of purification. Lane 5 shows the gel sample of the insoluble fraction and it seems that much of the OB12 protein is present. In fact there is no OB12 present in the elution, showing that the protein was lost before the elution (potentially in the insoluble fraction).

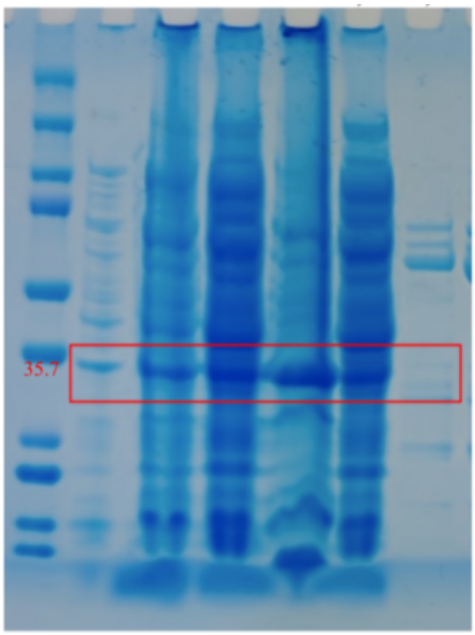

Figure 25 - OB12 Purification

The denaturing protocol using 6M GuHCl was not used because OB12 is too large. This protocol is only recommended for proteins less than 18 kDa since proteins bigger than this are more likely to improperly refold. However, to potentially increase the yield of OB12, the purification was repeated using a 1.5M salt concentration. This is because in the insoluble fraction, larger fragments of DNA are present. Thus, it was hypothesized that OB12 was present in the insoluble fraction because it was binding to DNA. The strength of protein DNA interactions reduces in the

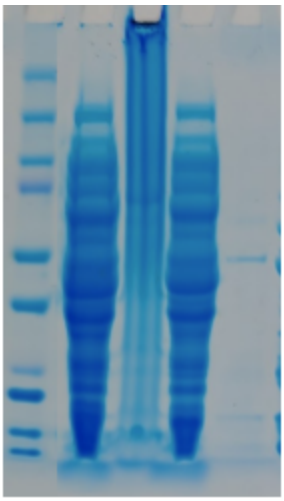
Figure 26 - OB12 Purification with 1.5M NaCl

presence of high salt concentration due to lower electrostatic interactions. However, this did not increase the yield of OB12 as it still was not present in the elution as seen in the last lane of Figure 26.

OB123:

OB123 was first purified under normal conditions. The SDS-PAGE gel in Figure 27 shows the steps of the OB123 purification. This protein is 55 kDa, however there are no clear bands near this area on the gel indicating that OB123 was not successfully purified. In fact, the elution (lane 7) itself does not contain any OB123 protein. This is quite surprising since it was expected that OB123 would be more stable than the other constructs since it contains all individual domains.

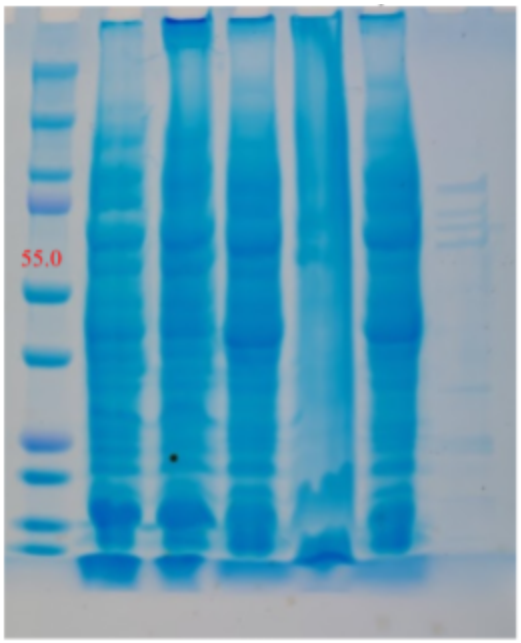

Figure 27 - OB123 Purification

OB123 could have potentially aggregated thus it is not binding to the Co^{2+} matrix. Due to this problem, the denaturing protocol for purifying protein using 6M GuHCl was used. This protocol is only recommended for proteins that are less than 18 kDa in size. However since sub domains (like the individual OB folds) fold independently/separately from one another, it was hypothesized that using this protocol for purifying a protein as big as OB123 was justifiable. From the SDS-PAGE gel shown in Figure 28, it

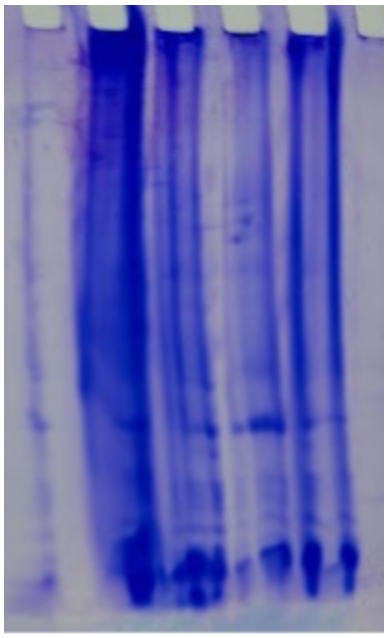

Figure 28 - OB123 Purification with 6M GuHCl

is evident that the denaturation purification protocol was not successful. OB123 was not present in the elution (or any other lane).

OB23:

OB123 was first purified under normal conditions. The SDS-PAGE gel in Figure 29 shows the steps of the OB123 purification. This protein is 55 kDa, however there are no clear bands near this area on the gel indicating that OB123 was not successfully purified. In fact, the elution (lane 7) itself does not contain any OB123 protein. However due to time constraints optimization of

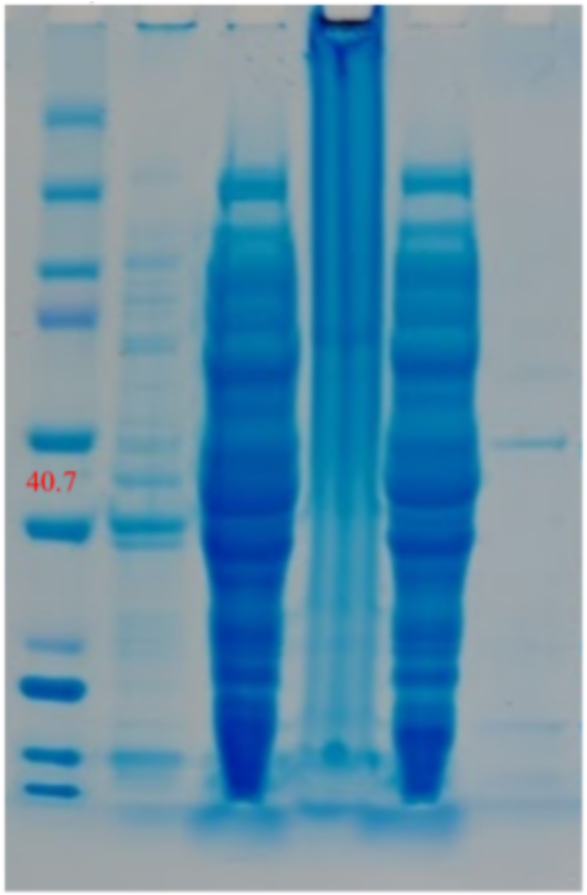

Figure 29 - OB23 Purification

EMSA:

The electrophoretic mobility shift assay tests the DNA binding affinity of a protein for DNA. Figure 30 shows how these assays accomplish this. Each lane is loaded with DNA and increasing concentrations of protein. The protein binds to DNA, and as a result it has a higher molecular weight than the unbounded DNA. Due to this the bands on top showed DNA bounded by protein while those on the bottom show unbounded DNA. Notice that the bands of the unbound DNA

get lighter and the bands of the bounded DNA get darker as the concentration of protein increases. This is what should be expected.

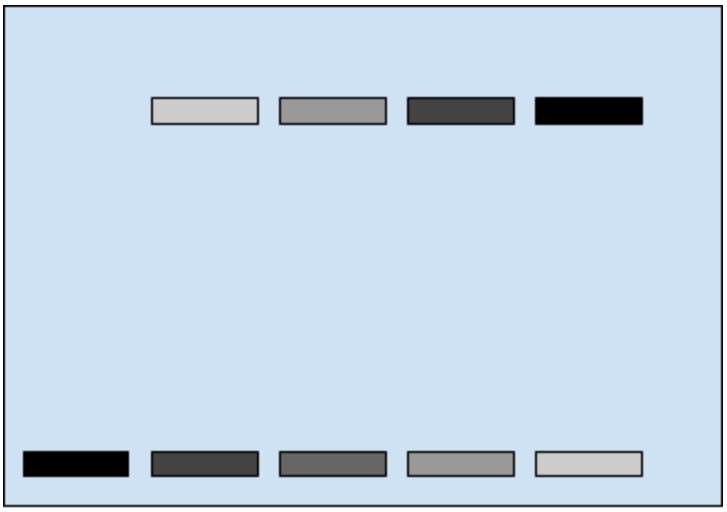

Figure 30 - EMSA Schematic

Figure 31shows the best EMSA gel that this study could produce. Unfortunately there is Not as much separation between the bands and The control band of unbounded DNA and no protein is not present. However the binding reactions used to make this gel can be optimized to produce a better gel with a higher resolution.

Figure 31 - EMSA for OB1 bound to 60nt ssDNA

Well #	DNA (1 µM)	Protein	1x BB		Glycerol	
2 - 250nM DNA, 500 nM OB1	5 µL	1 µL	12 µL		2 µL	
3 - 250nM DNA, 600 nM OB1	5 µL	1.2 µL	11.8 µL		2 µL	
4 - 250nM DNA, 700 nM OB1	5 µL	1.4 µL	11.6 µL		2 µL	
5 - 250nM DNA, 900 nM OB1	5 µL	1.8 µL	11.2 µL		2 µL	
6 - 250nM DNA, 1250 nM OB1	5 µL	2.5 µL	10.5 µL		2 µL	
7 - 250nM DNA, 2000 nM OB1	5 µL	4 µL	9 µL		2 µL	
8 - 250nM DNA, 2500 nM OB1	5 µL	5 µL	8 µL		2 µL	
9 - 250nM DNA, 3000 nM OB1	5 µL	6 µL	7 µL		2 µL	
10 - 250nM DNA, 6000 nM OB1	5 µL	12 µL	1 µL		2 µL	

The table below shows the binding reaction concentrations for this gel. The methods section describes more clearly what each reagent was. Though this gel is not ideal, it does show that OB1 does bind to ssDNA which is important in understanding where near the DSB overhang the protein binds to.

Though there were a lot of optimization steps needed to get the EMSA results shown, improving on these conditions like the runtime and concentrations of the DNA, protein, and binding reagents will all play a role in increasing the resolution of the gel.

Bibliography:

1. Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. Int J Med Sci. 2012;9(3):193-199. doi:10.7150/ijms.3635

2. Mahaney BL, Meek K, Lees-Miller SP. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. Biochem J. 2009;417(3):639-650. doi:10.1042/BJ20080413

3. Toulany M. Targeting DNA Double-Strand Break Repair Pathways to Improve Radiotherapy Response. Genes (Basel). 2019;10(1):25. Published 2019 Jan 4. doi:10.3390/genes10010025

4. Rodgers K, McVey M. Error-Prone Repair of DNA Double-Strand Breaks. J Cell Physiol. 2016;231(1):15-24. doi:10.1002/jcp.25053

5. Srivastava M, Raghavan S. DNA double-strand break repair inhibitors as cancer therapeutics. Chemistry & biology. 2015;22(1):17-29. http://dx.doi.org/10.1016/j.chembiol.2014.11.013. doi: 10.1016/j.chembiol.2014.11.013.

6. Jekimovs C, Bolderson E, Suraweera A, Adams M, O'Byrne KJ, Richard DJ. Chemotherapeutic compounds targeting the DNA double-strand break repair pathways: the good, the bad, and the promising. Front Oncol. 2014;4:86. Published 2014 Apr 22. doi:10.3389/fonc.2014.00086

7. Setiaputra D, Durocher D. Shieldin – the protector of DNA ends. EMBO Rep. 2019;20(5):n/a. doi: 10.15252/embr.201847560.

8. Greenberg RA. Assembling a protective shield. Nat Cell Biol. 2018;20(8):862-863. doi: 10.1038/s41556-018-0152-x.

9. Findlay S, Heath J, Luo VM, et al. SHLD2/FAM35A co-operates with REV7 to coordinate DNA double-strand break repair pathway choice. EMBO J. 2018;37(18):n/a. doi: 10.15252/embj.2018100158.

10. Fackrell K, Bobins L, Parul, Tomida J. FAM35A/SHLD2/RINN2: A novel determinant of double strand break repair pathway choice and genome stability in cancer. Environ Mol Mutagen. 2020;61(7):709-715. doi:10.1002/em.22379

11. Gao S, Feng S, Ning S, et al. An OB-fold complex controls the repair pathways for DNA double-strand breaks. Nature communications. 2018;9(1):3925. doi: 10.1038/s41467-018-06407-7.

12. Hsieh, Y. W., Alqadah, A., & Chuang, C. F. (2016). An Optimized Protocol for Electrophoretic Mobility Shift Assay Using Infrared Fluorescent Dye-labeled Oligonucleotides. Journal of visualized experiments : JoVE, (117), 54863. https://doi.org/10.3791/54863

13. Du F, Liu YQ, Xu YS, Li ZJ, Wang YZ, Zhang ZX, Sun XM. Regulating the T7 RNA polymerase expression in E. coli BL21 (DE3) to provide more host options for recombinant protein production. Microb Cell Fact. 2021 Sep 26;20(1):189. doi: 10.1186/s12934-021-01680-6. PMID: 34565359; PMCID: PMC8474846.

14. Bornhorst, J. A., & Falke, J. J. (2000). Purification of proteins using polyhistidine affinity tags. Methods in enzymology, 326, 245–254. https://doi.org/10.1016/s0076-6879(00)26058-8