Generating Single-Domain Antibodies Against Fibronectin Splice Variants

by

Alexis W. Peterson

B.S., Biochemistry
Boston College, 2012

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN BIOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2017

© 2017 Massachusetts Institute of Technology. All rights reserved

Signature of Author ____ Signature redacted _____
Department of Biology
December 16, 2016

Certified by _____ Signature redacted _____
Richard Hynes
Daniel K. Ludwig Professor for Cancer Research
Thesis Supervisor

Accepted by _____ Signature redacted _____
Michael Hemann
Associate Professor of Biology
Co-Chair, Biology Graduate Committee

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JAN 12 2017
LIBRARIES
ARCHIVES

MITLibraries

77 Massachusetts Avenue
Cambridge, MA 02139
http://libraries.mit.edu/ask

DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

The images contained in this document are of the best quality available.

Generating Single-Domain Antibodies Against Fibronectin Splice Variants

by

Alexis W. Peterson

Submitted to the Department of Biology
on December 16, 2016 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Biology

Abstract

Here, I describe the process of generating single-domain antibodies which bind to splice variants of fibronectin containing EIIIA or EIIIB. Alpacas were immunized with either purified antigen cocktails, or from the ECM of tumor samples, and antibody libraries were generated. Using these libraries to pan against, I selected for VHH which bind to EIIIA and EIIIB. Since these splice variants are upregulated in tumor angiogenesis and are rarely seen elsewhere in adult tissues, antibodies targeting EIIIA or EIIIB may be of use for imaging tumors and metastases.

Thesis Supervisor: Richard Hynes
Title: Daniel K. Ludwig Professor for Cancer Research

Table of Contents

List of Figures

1. Introduction

Metastatic spread is responsible for the overwhelming majority of cancer-related deaths. There are many aspects of metastasis one could study, including how the extracellular matrix (ECM) contributes to the metastatic cascade. Our lab has used LC/MS proteomic analysis to analyze the ECM signatures of mouse xenograft models of melanoma cell lines[1], breast cancer cell lines of differing metastatic potential[2], as well as human patient colon carcinomas and their metastases to liver[3], in addition to ongoing studies.

These data have given us lists of ECM proteins which we think may be used as targets to image tumors. Among the ECM proteins of interest are two splice variants of fibronectin, termed EIIIA and EIIIB. Fibronectin has been shown to play a role in cell adhesion, growth, differentiaton, and migration[4]. It plays an important role in linking the extracellular environment to the inside of cells, via acting as an integrin family ligand. Thus, processes dependent on sensing the extracellular enviroment generally rely, in part, on fibronectin. Fibronectin generally exists as a ~500 kDa unit, a dimer of two very similar ~250 kDa subunits[5], and it can be localized to the ECM or circulate in blood, termed cellular and plasma fibronectin, respectively. Through alternative splicing, there are multiple isoforms of fibronectin which exist. Two exons which may either be skipped or included are EIIIA and EIIIB. Both EIIIA and EIIIB are type III repeats, meaning they are approximately 90 amino acid residues long and do not contain any disulfide bonds[5].

EIIIA is very highly conserved among species, with only one amino acid different between humans and alpacas. This amino acid change is an aspartic acid to glutamic acid, which likely results in either no change or a very small change to the overall domain structure and

function. EIIIB is also very highly conserved, and is identical in human, rat, rabbit, dog, monkey, and mouse. As with EIIIA, in alpaca, there is one amino acid different from the human form of EIIIB. Compared to each other, however, EIIIA and EIIIB show only 28% similarity in humans. EIIIA and EIIIB are present during embryogenesis, and can be seen during tumor angiogenesis, and wound healing[6]. Thus, in adult tissues, EIIIA and EIIIB are rarely seen other than during tumorigenesis. Fibronectin can contain EIIIA, EIIIB, both, or neither, and the prevalence of EIIIA compared to EIIIB in tumors differs, with 50% of invasive ductal carcinomas showing EIIIA positivity and 33% showing EIIIB positivity in an in situ hybridization study[7]. These values can vary depending on the exact tumor type as well as the method used to detect EIIIA and EIIIB, but indicate that both EIIIA and EIIIB could be important biomarkers, and may have biological functions in tumorigenesis.

Dario Neri's laboratory has generated antibodies targeting EIIIA and EIIIB. These antibodies have been used for multiple purposes, including to image tumor vasculature[8,9] provide localized chemotherapy [10], provide localized immunotherapy for tumors[11,12], deliver localized cytokines to modulate and study tumor angiogenesis[13], monitor chronic cardiac rejection[14] and inhibit the progression of collagen-induced arthritis[15]. The antibodies have been engineered to be in multiple formats, including as single-chain Fv antibodies (scFv). ScFv antibodies are fusion proteins consisting of the variable regions of the heavy and light chains linked together via a short 10-25 amino acid linker[16], resulting in a ~27kDa antibody. Different types of antibodies can have different uses, and, of critical importance, the size of an antibody can affect pharmacokinetics and pharmacodynamics. Small antibodies may be particularly useful in imaging, as one would expect a greater tissue penetration by smaller antibodies compared to

traditional IgG antibodies, which are approximately 150 kDa[17]. Likewise, antibodies which are smaller than the renal molecular weight cutoff of around 60kDa, can be rapidly excreted in urine, a property which would be useful for imaging studies[18]. Conversely, for treatment, increasing circulation time could be beneficial, so larger antibodies could be more useful in this setting.

To generate antibodies which would be useful for imaging, I generated nanobodies, using alpacas to create the initial antibodies. Members of the camelid family, as well as certain other species including sharks, have naturally occurring single-chain antibodies. In alpacas, heavy-chain-only antibodies constitute half of the serum IgG content, and, like conventional antibodies, undergo affinity maturation[19]. Using molecular biology techniques, it is possible to create the single-domain-only equivalent of an scFv, which consists of only the variable region of the heavy chain. The variable region of the heavy chain is sufficient to confer antigen specificity to the antibody, and the resulting nanobodies, or VHH, are only 12-15 kDa.

Currently, ^{18}F-FDG PET/CT, which utilizes a radioactive glucose analog and relies on increased glycolysis by malignancies, is often used to localize tumors. Yet, because ^{18}F-FDG PET/CT relies on differential glucose uptake, non-malignant pathologies and highly metabolic tissues can affect specificity, and molecular characteristics of tumors cannot be assessed this way. To overcome these limitations, single-domain antibodies (sdAb) against ECM peptides or proteins of interest could be used to image tumors. This would first be done using fluorescence imaging, and could then be extended to more clinically relevant imaging modalities. Clinically, this could be useful for staging and monitoring therapeutic response. In future studies, these sdAb could also have a use as targeted drug delivery systems, or as intra-operative guides.

In the following chapters, I will describe the methods I used [1] to generate purified EIIIA for antibody selection and purification, [2] to pan antibody libraries to select for antibodies binding to antigen of interest, and [3] to sub-clone the antibodies into vectors suitable for high levels of VHH expression.

2. Overexpression and purification of fibronectin EIIIA

To generate purified EIIIA, I used the pET SUMO expression system (Thermo Fisher). This vector uses TA cloning to insert the DNA sequence of interest, and has a N-terminal 6x his tag followed by SUMO. SUMO, an 11kDa ubiquitin-related protein, stabilizes and localizes proteins in vivo. After purification, SUMO (ULP-1) protease can be used to cleave the SUMO/ His tag off the N-terminal of the peptide. This cleavage, which is at the C-terminal of SUMO leaves no excess resides at the N-terminal of the recombinant peptide. In this system, isopropyl β-D-1-thiogalactopyranoside (IPTG) can be used to drive high and regulated expressing through IPTG's inhibition of the lac repressor. For selection, the pET SUMO confers kanamycin resistance.

a. Cloning EIIIA into pET SUMO

TA cloning works by polymerase chain reaction (PCR) amplification using Taq polymerase, which preferentially adds a 3' adenine. This product then can be ligated into a linearized vector containing a complementary 3' thymidine overhang. In order to PCR amplify EIIIA, I used human umbilical vein endothelial cell (HUVEC) cDNA as the DNA template, along with the following primers, which I designed to specifically amplify the DNA encoding the EIIIA portion of human fibronectin.

EIIIA forward: 5' – AACATTGATCGCCCTAAAGGACTGG – 3'

EIIIA Reverse: 5' – AGGAATAGCTGTGGACTGGGTTCC – 3'

To set up the PCR reaction, I combined >100ng HUVEC cDNA, 1X PCR buffer, 0.5mM dNTPs, 2μM forward and reverse primers, and 1 unit of TaKaRa-Taq polymerase (R001A). The cycling parameters consisted of an initial 30 second denaturation at 94°C, followed by 30 cycles of amplification (30 second denaturation at 94°C, 50 second annealing at 60°C, and 30 second extension at 72°C). This was followed by a 7 minute extension step at 72°C, and storage at 4°C. Figure 2.1 shows a DNA gel electrophoresis of the amplified EIIIA (A), as well as a non-template control, in which HUVEC cDNA template was omitted from the PCR reaction mix.

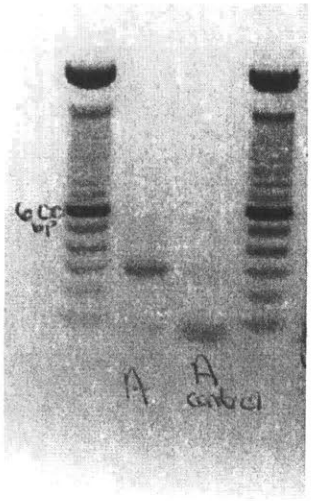


Figure 2.1. Gel electrophoresis of PCR amplified EIIIA. Amplified EIIIA is labeled as (A) and non-template control labeled as (A control), using 1% agarose and ethidium bromide staining.

To ligate the amplified DNA into pET SUMO, I combined 1.5μL PCR product, 1x ligation buffer, 50 ng pET SUMO vector, and 4 Weiss units of T4 DNA ligase in a total volume of 10μL using sterile water. This reaction was allowed to incubate overnight at 15°C. A map of the expected pET SUMO vector containing EIIIA is shown in Figure 2.2.

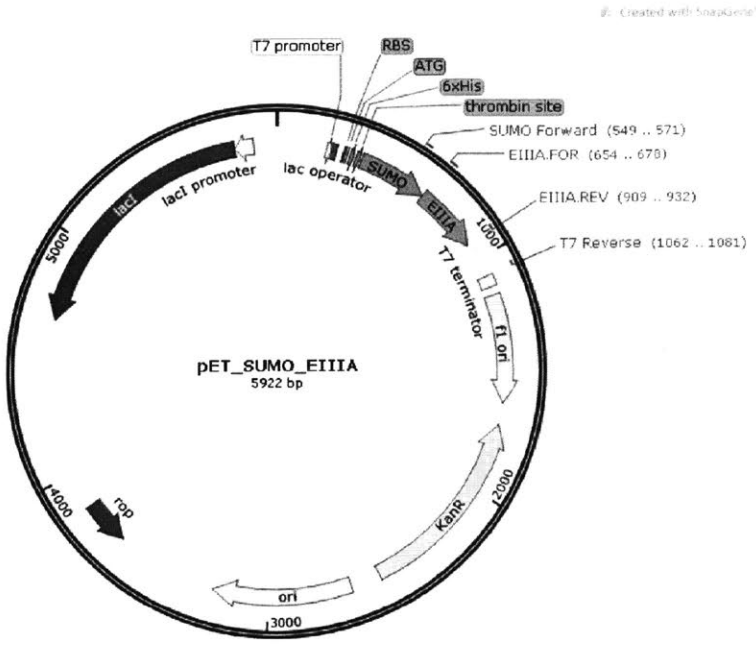


Figure 2.2. Map of expected pET SUMO-EIIIA vector. Figure generated using SnapGene.

I used One Shot Mach1-T1^{R} chemically competent cells supplied with the pET SUMO kit for transformation. The cells were made competent (able to take up extracellular DNA) by the manufacturer using calcium chloride treatment. In this method, calcium is thought facilitate the adsorption of DNA to the core oligosaccharide domain of lipopolysaccharide molecules which are highly abundant on the cell surface of gram-negative bacteria such as *E. coli*[20]. After thawing the cells on ice, I pipetted 2μL of the ligation reaction into a vial of cells, and gently mixed the solution. As a transformation control, I set up a parallel reaction using pUC19 in place of the ligation reaction. The cells were incubated on ice for 30 minutes, heat-shocked for 30 seconds at 42°C, then immediately placed back on ice. After the addition of 250 μL of S.O.C. media, the cells were incubated for 1 hour at 37°C with shaking at 200 rpm. 100-150 μL of the

transformation reactions were plated onto L.B plates containing 50 μg/mL kanamycin and grown overnight at 37°C.

As TA cloning can result in the amplified DNA being inserted in either direction, I selected 10 positive transformants for overnight liquid culture in L.B. + 50 μg/mL kanamycin to check for direction of insertion. Following overnight growth, I made 15% glycerol stocks of each clone, and mini-prepped the plasmid DNA using a standard protocol supplied by the supplier (Qiagen). The samples were sent for sequencing using the following primers:

pET SUMO forward: 5' – AGATTCTTGTACGACGGTATTAG – 3'

T7 reverse: 5' – TAGTTATTGCTCAGCGGTGG – 3'

Of the 10 colonies initially selected, 4 showed the insert in the correct orientation. The sequences of these clones are shown in Table 2.1.

Table 2.1. Sequencing results from clones with the correct DNA insert orientation, using the pET SUMO forward primer. Aberrations from the expected sequence are highlighted.

CLONE #	DNA SEQUENCE
Expected	ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATT CCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTA CAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCC CTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCTCAG ACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATA TGGAGAGCCAGCCCCTGATTGGAACCCAGTCCACAG
1	ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATT CCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTA CAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCC CTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCTCAG ACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATA TGGAGAGCCAGCCCCTGATTGGAACCCAGTCCACAG
5	ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATT CCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTA CAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCC CTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCTCAG ACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATA TGGAGAGCCAGCCCCTGATTGGAACCCAGTCCACAG
6	ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATT CCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTA CAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCC CTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCTCAG ACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATA TGGAGAGCCAGCCCCTGAGTGGAACCCAGTCCACAG
10	ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATT CCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTA CAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCC CTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCTCAG ACCGGGTTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATA TGGAGAGCCAGCCCCTGATTGGAACCCAGTCCACAG

b. EIIIA overexpression

For overexpression and purification, I selected clone # 10, as it was one of the clones which did not have any aberrations from the desired, expected sequence. As a host for expression, I used BL21(DE3) One Shot *E. coli*, which were supplied with the pET SUMO kit. This strain of *E. coli* yields high levels of expression of genes regulated by the T7 promotor. This particular strain of *E. coli* was not used in the initial cloning and glycerol stock because there is some basal level of T7 polymerase in BL21(DE3) cells, which can lead to plasmid instability if the gene of interest contained in the plasmid has toxicity to *E. coli*.

To transform BL21(DE3) cells with pET-SUMO-EIIIA (isolated from clone 10), I thawed a vial of BL21 cells on ice, and added 3ng plasmid DNA with gentle mixing. The cells and plasmid DNA were incubated 30 minutes on ice, heat-shocked for 30 seconds at 42°C, and placed back on ice. 250 μL S.O.C. medium were added and the cells were placed in a 37°C incubator for 1 hour with 200 rpm shaking for a recovery. S.O.C., or super optimal broth with catabolite repression, is a nutrient rich broth with a pH of 7 composed of 2% w/v tryptone, 0.5% w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose. In comparison, L.B., or lysogeny broth, consists of 1% w/v tryptone, 0.5% w/v yeast extract, and 171 mM NaCl. The use of S.O.C. medium following heat-shock results in higher transformation efficiency than if L.B. were used for certain bacterial strains, likely due to the presence of cations such as Mg^{2+} and K^+ and glucose in S.O.C[21]. After the 1 hour recovery period in S.O.C, the entire transformation reaction was then added to 10mL of L.B. medium containing 50 μg/mL kanamycin and grown overnight at 37°C with shaking, generating a saturated transformation culture.

As previously mentioned, the DNA inserted into pET SUMO is under control of the T7 promotor and the Lac operator. The activity of the T7 promotor is inhibited by the lac repressor protein encoded by LacI. In the absence of lactose, the lac repressor protein binds the lac operator sequence and bends the DNA 40°, preventing T7 RNA polymerase from accessing the T7 promotor[22]. IPTG is a molecular mimic of allolactose, which is a lactose metabolite able to bind to the lac repressor. This, in effect, inhibits the inhibition, allowing transcription.

To determine an optimal time-point for IPTG induction, I ran a 5-hour time-course. First, 10mL of L.B. containing 50 μg/mL kanamycin were inoculated with 500 μL of saturated transformation culture, and grown until the OD_{600} was between 0.4 and 0.6, placing the bacteria in mid-log phase growth. The 10mL culture was split into two; one un-induced and one induced with 1mM IPTG. At the time the cultures were split (time 0) and every hour for 5 hours, I collected 500 μL in a microcentrifuge tube, centrifuged the sample at 12,500 rpm for 30 seconds, and aspirated the supernatant. The samples were stored at -20°C. To assess protein induction, thawed samples were resuspended in 80 μL 1X SDS-PAGE sample buffer. After 5 minutes of boiling and brief centrifugation, 10 μL of each sample were loaded on an SDS-PAGE gel and electrophoresed. Proteins were visualized using Coomassie staining, as shown in Figure 2.3.

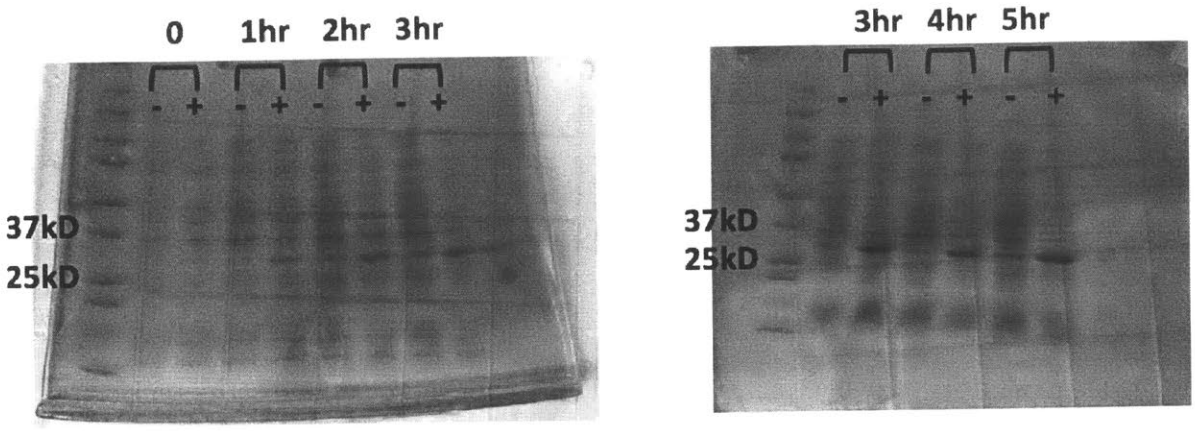


Figure 2.3. EIIIA pilot induction using 1 mM IPTG. (-) are uninduced controls, while (+) were induced for the indicated time.

The expected size of the SUMO-EIIIA fusion protein is around 23.5 kDa. The EIIIA portion should be 10.5 kDa, and the 6x-His-SUMO approximately 13 kDa. A difference in expression was present by 1 hour, as shown by a band just above 25 kDa. Although the band was slightly higher than expected, it ran close to the expected size,. By 5 hours, there was a large increase in expression in the induced sample, and this induction time was used for subsequent, larger-scale overexpression of the recombinant fusion protein.

To scale up overexpression, fresh BL21(DE3) cells were transformed with the isolated pET-SUMO-EIIIA-(10) plasmid, using the same methods as previously described. Following overnight growth, 2 mL of saturated transformed bacteria were added to 100 mL of L.B. medium containing 50 μg/mL kanamycin. After growth to mid-log phase (2 hours), IPTG was added to reach a final concentration of 1 mM. After 5 hours of induction, the cells were harvested by centrifugation at 3000 x g for 10 minutes at 4°C. Samples were stored at -80°C until purification.

c. EIIIA purification

The general steps in the purification protocol are to (1) lyse induced cells, (2) isolate the fusion protein using the 6x-His tag, (3) cleave SUMO from EIIIA using SUMO protease, (4) remove SUMO and SUMO protease.

I elected to use multiple methods of lysis. First, I diluted bacterial cell paste into binding buffer (20mM sodium phosphate, 500mM NaCl, 40mM imidazole, pH 7.4), using 5 mL of binding buffer per gram of cell paste. For enzymatic lysis, I used 1mM $MgCl_2$, ~2 units/mL DNase, 0.2mg/mL lysozyme, and cOmplete mini protease inhibitor tablets (Roche). The resuspended cell paste was placed on a rotating platform for 30 minutes at room temperature. Additional time was added as needed if the solution appeared to be very viscous. Following enzymatic lysis, I moved the solution to an ice bath, and sonicated for 10 minutes on ice. In this procedure, there were 5 rounds of 1 minute sonication in an ice bath, with 1 minute on ice between. Prior to loading the lysed solution onto a column, I checked that the pH was still at 7.4.

Using a HisTrap Crude FF column (GE, 1 mL), I isolated the His-tag-EIIIA fusion protein. These columns are pre-packed with nickel-ion-containing medium, and have a dynamic binding capacity of 40 mg 6x-His-tagged protein per milliliter of medium. To use the HisTrap Crude columns, I first flushed the column with 5mL sterile water. This was followed by equilibration with 5mL binding buffer. The lysate was then applied, and washed with ~20mL binding buffer (20 column volumes). For elution of the fusion protein, I used ~5mL elution buffer (20 mM sodium phosphate, 500mM NaCl, 500mM imidazole, pH 7.4). To check for fractions containing protein, A280 measurements on a nanodrop were taken throughout the procedure. The columns,

which are reusable, were washed with 5mL of 20% EtOH and stored in 20% EtOH at room temperature.

Following elution into elution buffer, I loaded the sample into a dialysis cassette (10,000 MWCO), and dialyzed into SUMO protease buffer (50mM Tris-HCl (pH 8), 150 mM NaCl, 1mM DTT) overnight at 4°C. High imidazole concentrations (greater than 150 mM) can adversely affect SUMO protease activity, necessitating either desalting or dialysis. To cleave SUMO from EIIIA, excess purified SUMO protease (see SUMO protease purification below) at a concentration of 50 μg SUMO protease per mg fusion protein was combined with 0.2% IGEPAL in SUMO protease buffer for 1 hour at room temperature with inversion.

The SUMO protease I used was expressed using a plasmid that added a 6x-His tag. This allowed me to use a second His-trap crude FF (1 mL) column to remove both SUMO and SUMO protease from the EIIIA solution. Since EIIIA does not have a His-tag following SUMO cleavage, the flow-through was collected during the step when the sample containing EIIIA was applied to the column. A portion of the purified EIIIA was dialyzed into phosphate buffer overnight, with 3 total buffer changes, yielding protein in the correct buffer for antibody panning. Nanodrop A280 was used to measure the concentration of purified protein (4.4 mg/ml), and samples were flash frozen in liquid nitrogen prior to storage at -80°C. To check for the presence and purity of EIIIA, samples were taken at various times throughout the purification procedure, and were assessed using SDS-PAGE and Ponceau S staining, followed by western blotting using a commercially available anti-EIIIA antibody (Abcam, ab6328) at a 1:2000 dilution.

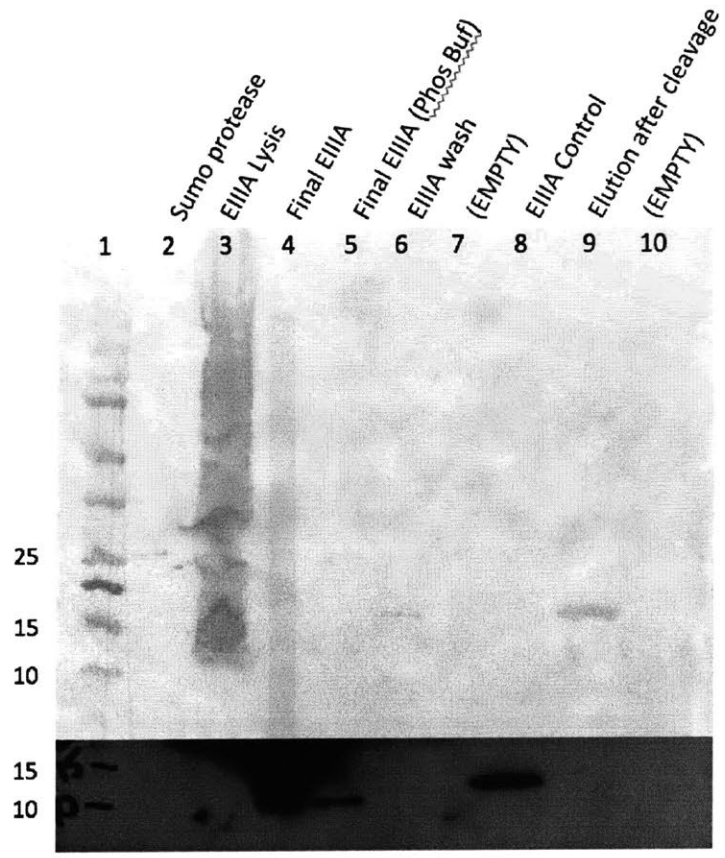


Figure 2.4. Ponceau S staining and western blotting of EIIIA. The final purified EIIIA in lane 4 is in sumo protease buffer, while the final purified EIIIA in lane 5 was dialyzed into phosphate buffer, for use in antibody panning. EIIIA control in lane 8 was previously purified in our lab using a different expression vector.

As EIIIA does not stain well with Ponceau S, I decided to run the sample on LC-MS to check for purity, courtesy of Ethan Evans in the Pentelute Laboratory at MIT. As shown in Figure 2.5(a), there was a prominent peak at 8 minutes, followed by a larger, broad peak at 10.5 minutes. The 8-minute peak was consistent with the molecular weight of the expected EIIIA

(10.5 kD), shown in Figure 2.5(b), and the 10.5-minute peak is characteristic of a detergent, in this case IGEPAL from the sumo protease buffer.

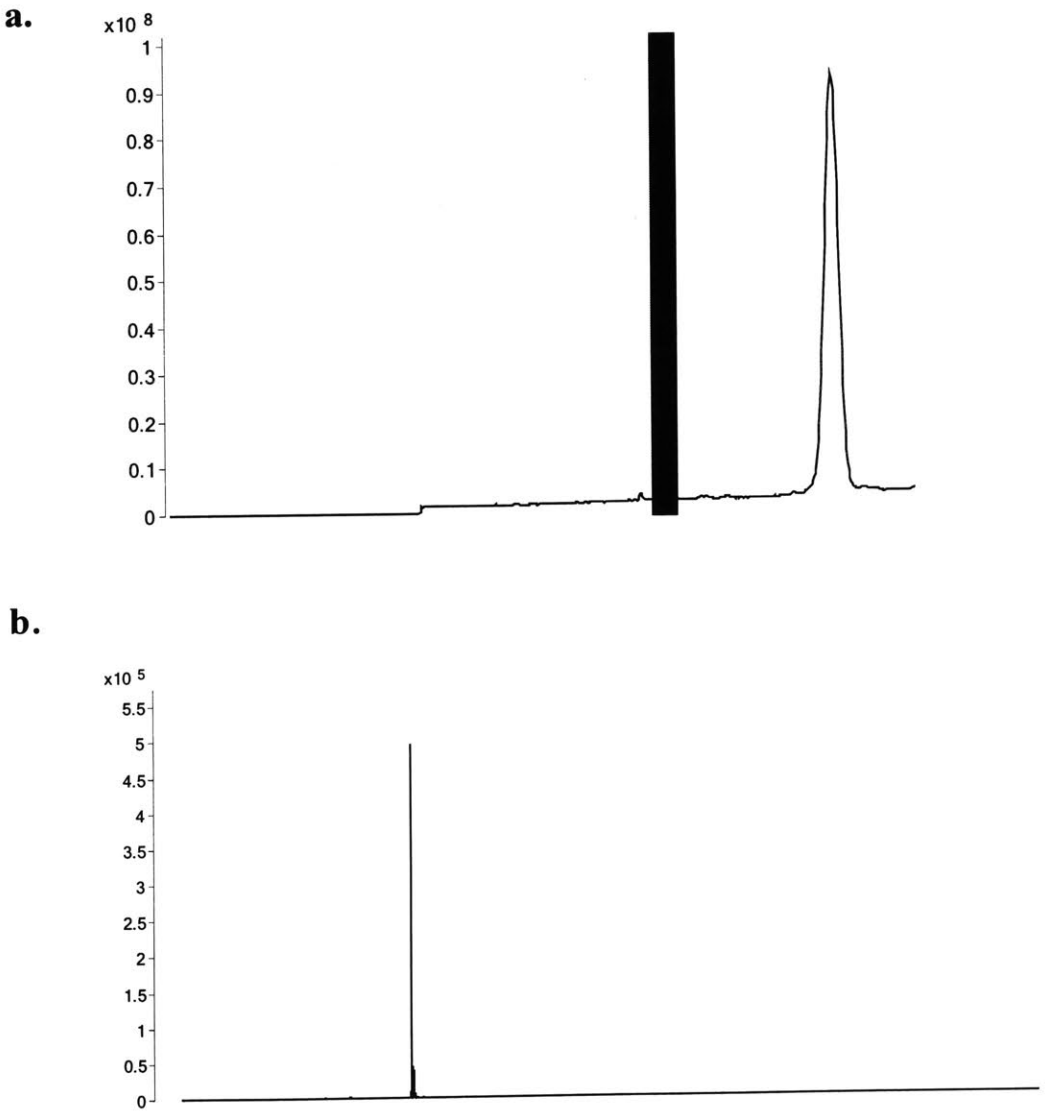


Figure 2.5. LC-MS of purified EIIIA. (a) shows the counts vs. acquisition time, an indication of purity. The scan showed in (b) shows the deconvoluted mass of the peak seen at 8 minutes.

From the LC-MS as well as vector sequencing, the peptide sequence for the purified EIIIA is below:

D I D R P K G L A F T D V D V D S I K I A W E S P Q G Q V S R Y R V

T Y S S P E D G I H E L F P A P D G E E D T A E L Q G L R P G S E Y T

V S V V A L H D D M E S Q P L I G T Q S T A I P R Q A

d. Expressing and purifying SUMO protease

To generate sufficient amounts of SUMO protease for EIIIA purification, I overexpressed SUMO protease from a glycerol stock of *E. coli* which were previously transformed with a plasmid encoding a version of SUMO protease containing a His-tag for purification, courtesy of Ethan Evans. The bacteria and SUMO protease plasmid confer ampicillin and chloramphenicol resistance. After overnight growth in 10mL of L.B. with 25 μg/ml chloramphenicol and 100 μg/ml ampicillin, the bacteria were transferred to 500 mL of the same medium, and grown to mid-log phase. Expression of SUMO protease was induced with 1 mM IPTG overnight. Following centrifugation at 3000 x g, the cell paste was stored at -80°C until purification. The same purification protocol used for EIIIA was used to purify SUMO protease. As is shown in Figure 2.6, there is a higher intensity >25kDa band in the sumo protease lysis sample, with a band of the same size in the purified sumo protease sample, which is consistent with expression and purification of the expected 26kDa protease.

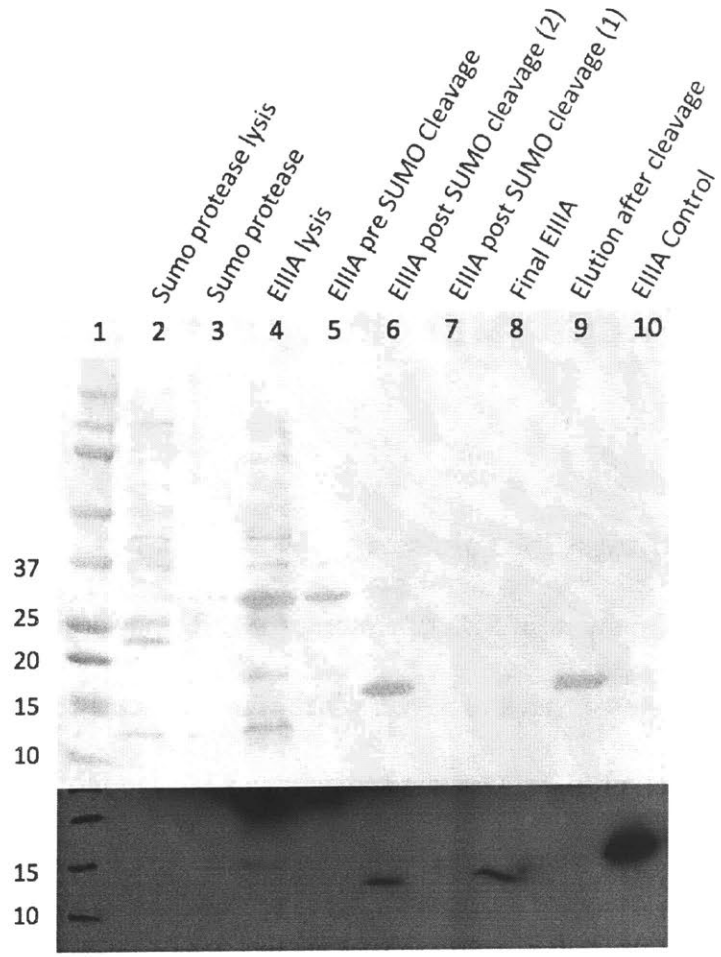


Figure 2.6. Sumo protease expression and purification as shown by SDS-PAGE and Ponceau S staining. EIIIA purification and anti-EIIIA western blotting is also shown.

3. Generating and isolating single-domain antibodies against EIIIA and EIIIB

The steps to creating single domain antibodies include (1) immunizing an alpaca with antigen(s) of interest, (2) generating an antibody library from isolated lymphocytes from the immunized alpaca, (3) panning the library to select antibodies which bind to a particular target, (4) moving the antibody sequence to vectors suitable for high levels of antibody production, (5) purifying the antibodies, and (6) validating the target specificity of the antibodies.

While there are multiple possible methods to generating antibodies that bind to a particular target, the approach described in the following pages uses M13 bacteriophage. M13 bacteriophage is a virus which infects *E. coli*, and is composed mainly of the major coat protein (P8) along with multiple minor coat proteins. One of the minor coat proteins is pIII, which is localized to the surface of phage and is responsible for initiating infecting by binding to a receptor at the tip of the F pilus of *E. coli*[23]. To use this naturally occurring system for biological studies, phagemid plasmids such as pCANTAB have been created which have origins of replication, bacteriophage packaging signals, and encode the pIII protein adjacent to a cloning site. Following cloning of DNA sequences of interest and subsequent expression of pIII, fusion proteins composed of pIII and protein encoded by the inserted DNA are generated. When the other proteins necessary for phage synthesis are also present, such as when helper phage is added, functional phage will display the fusion protein on their surface and the DNA encoding the displayed protein will be packaged inside the viral particle. This allows for selection of DNA sequences based on the resulting protein.

a. VHH library construction[24]

Noor Jailkhani and Steffen Rickelt, postdocs in the Hynes laboratory, immunized an alpaca with a cocktail of ECM proteins and synthetic peptides, which included rat-derived EIIIA and EIIIB. In collaboration with the Ploegh lab at MIT, Noor made the VHH antibody library which I used for panning. After immunizing and boosting the immunization three times, they collected alpaca blood 3-5 days after a final antigen boost. Using centrifugation, they isolated the buffy coat, which is a fraction formed by centriguing blood and which contains lymphocytes and platelets. For creating antibody libraries, the lymphocytes are the cells of interest. Using an extraction kit, they extracted RNA from lymphocytes, and generated cDNA using random hexamers, gene-specific primers [Al.CH2 (ATGGAGAGGACGTCCTTGGGT) and AlCH2.2 (TTCGGGGGGAAGAYRAAGAC)], and poly dT. They amplified VHH DNA using the following VHH-specific PCR primers.

AlVHH-F1	CTTGCGGCCGCTCAGKTGCAGCTCGTGGAGWCNGGNGG
AlVHH-shR1	GATCGGCGCGCCGAGGGGTCTTCGCTGTGGTGCG
AlVHH-lhR1	GATCGGCGCGCCGGTTGTGGTTTTGGTGTCTTGGG

The PCR amplified products were digested with AscI and NotI, and gel-purified. A pCANTAB derivative was digested with AscI and NotI, digested PCR products were ligated into the vector using T4 ligase, and electrocompetent TG1 cells were transformed. The pCANTAB derivative used for library construction has a C-terminal E-tag, and digestion sites for AscI and NotI, which allow VHH sequence insertion between the leader sequence and the main body of

the M13 gene 3. Insertion of a VHH sequence does not alter the function of the proteins generated from the flanking M13 sequences, and the restriction enzyme digests ensure proper orientation of the VHH sequence. The E-tag is used as a way to detect antibody during panning using anti-E-tag antibodies. Following overnight growth, the bacterial cultures were frozen in glycerol, and a portion was mini-prepped and sent for sequencing to determine overall library diversity. In addition to the library made from an alpaca immunized with recombinant peptides and ECM proteins, libraries were also made from alpacas immunized with extracellular matrix preparations derived from human patient samples of lung and liver breast cancer metastases.

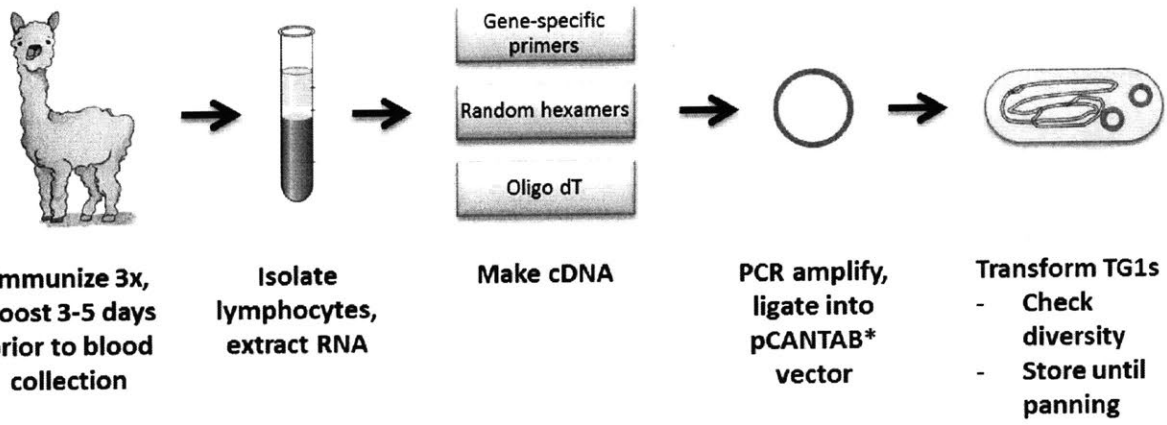


Figure 3.1. Overview of library construction

Table 3.1. Single-domain antibody libraries

Library name	Immunized with	Library diversity
Riuniti	Cocktail of purified ECM proteins and recombinant ECM peptides	***2.7 x 10^5 cfu/ml***
Ducati	Colorectal cancer metastasis to liver ECM prep	***3.7 x 10^6 cfu/ml***
Mac	Triple-negative breast cancer metastasis to lung ECM prep	***8.03 x 10^6 cfu/ml***
Esteban	Triple-negative breast cancer metastasis to liver ECM prep	***2.11 x 10^6 cfu/ml***

b. Panning[24]

Panning is the process used to select antibodies which bind to a specific target, generally using purified antigen. The method I followed uses magnetic beads bound to antigen using biotin-streptavadin conjugation. Starting with library stock, VCSM13 helper phage is added and grown overnight to generate phage displaying the sdAb library on their surfaces as fusions to capsid protein pIII. Using Chromalink Biotin Protein Labeling system, purified antigen is biotinylated at primary amines, and the efficiency of the reaction can be measured by absorbance at A354. This biotinylated antigen is then bound to magnetic streptavadin beads. The phage are negatively selected against beads without antigen, and then positively selected using the antigen-bead mixture. Following significant washing, ER2738 cells, a strain of *E. coli* commonly used for phage propagation, are added and grown overnight. This process, from adding helper phage to adding *E. coli*, is repeated using bacteria from the first round of panning along with more

stringent washing and lower antigen concentrations to yield higher affinity nanobodies. To confirm nanobody expression and antigen-binding ability, an ELISA is used, with the nanobody acting as the primary antibody in the assay. An overview of the panning process is shown in Figure 3.2.

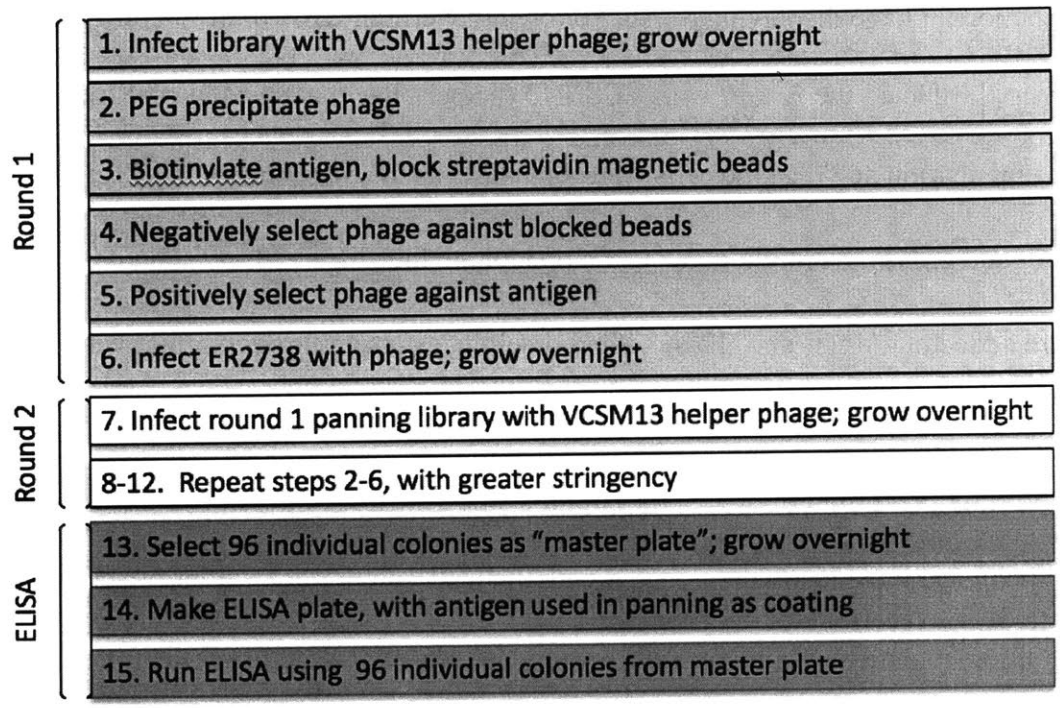


Figure 3.2. Overview of panning process

Noor and I panned against the Riuniti library using purified EIIIB and EIIIA containing a GST tag (GST-EIIIA). EIIIB yielded multiple positive hits on the ELISA and those hits have been sequenced and will be carried forward. Panning for antibodies recognizing GST-EIIIA was unsuccessful, possibly due to the size of the GST tag relative to EIIIA. With purified EIIIA, I repeated the panning against the Riuniti library. The protocols used for library generation, panning, and sub-cloning were given to our laboratory by members of the Ploegh lab at the Whitehead Institute.

c. Experimental details of panning using purified EIIIA and the Riuniti library

1. Helper phage preparation

Since phagemid plasmids do not encode all of the proteins required for phage particle production, helper phage which synthesize the other necessary proteins are needed in order to generate a functional phage displaying SdAb. Following overnight growth of *E. coli* ER2738 in 2YT medium, I infected the *E. coli* with VCSM13 (helper phage) at an MOI of 40:1. After 30 minutes of incubation at 37°C, the cells were spun down and resuspended in 2YT + 50μg/ml kanamycin, and grown overnight at 37°C. To precipitate the phage, 20% PEG 6000 and 2.5 NaCl were added at 1/5 volume. This method precipitates phage through effectively increasing the concentration of solutes through decreasing the effective solvent volume. Following overnight precipitation on ice, the solution was centrifuged to remove excess volume, and the phage were again precipitated with the PEG solution for 1 hour on ice. Following additional centrifugation, the number of phage were assessed using the following equation, which is based on the length of phage DNA and the amount of major coat protein VIII :

$$\#\frac{phage}{mL} = \frac{(6000\times10^{14})\times(A296-A320)}{7200bp\ (genome\ size)}$$

Equation 3.1. Phage per milliliter determination using absorbance measurements.

Using this equation and nanodrop assessment of A296 and A320, I had a yield of 1.4 x 10^{14} phage, to which sterile 50% glycerol was added to reach a final concentration of $1x10^{13}$

phage/mL in 15% glycerol. 1mL aliquots, for use in phage-display panning, were stored at -80°C until use.

2. Phage display panning using EIIIA

In order to select for antibodies which bind a particular antigen of interest, M13 phage antibody libraries can be panned against using purified antigen. In the method I used, biotinylated antigen bound to magnetic streptavidin beads are exposed to the phage-displayed antibody library, and following pH elution, the phage which were bound to antigen are exposed to stationary phase *E. coli*. This procedure, from phage-display and selection against antigen to infection of *E. coli*, is considered one round of panning. After two rounds of panning, phage-displayed antibodies which are able to bind to antigen, as assessed through an enzyme-linked immunosorbent assay (ELISA) are carried forward for sequence analysis and subsequent sub-cloning into appropriate expression vectors.

i. Panning round 1

Day 1: 100ml of S.O.C. containing 50 μg/ml ampicillin was inoculated with 200 μl of thawed Riuniti library stock, yielding an OD_{600} between 0.1 and 0.2. In this step, the glucose in S.O.C. suppresses expression of VHH to minimize growth bias during panning, and ampicillin is for phagemid selection. After 1.5 hours of growth, the OD_{600} was 0.562, and $1x10^{13}$ VCSM13 helper phage were added. Following 2 hours of incubation at 37°C, the first 30 minutes without shaking, the culture was split into two, spun at 8000 rpm for 10 minutes at 4°C, and re-suspended in 100 ml of 2YT containing

0.1% glucose, 50μg/ml ampicillin, and 70μg/ml kanamycin for overnight growth at 30°C. For infection with the phage display library, a 100ml overnight culture of ER2738 in 2YT containing 10μg/ml tetracycline was also initiated.

Day 2: The phage culture was spun down at 8000 rpm for 10 minutes at 4°C, to isolate the supernatant containing phage from *E. coli* in the pellet. As in the helper phage preparation, 20% PEG 6000 and 2.5M NaCl added at 1/5 volume was used to precipitate phage. After two hours of precipitation on ice, the solution was centrifuged at 13,000 rpm for 20 minutes at 4°C, the supernatant discarded, and the resulting phage pellet was re-suspended in 4ml PBS. 1mL PEG solution was added, for a second precipitation step (1 hour on ice). Following centrifugation at 13,000 rpm for 10 minutes at 4°C, the phage pellet was re-suspended in PBS, and the total yield was assessed using equation 3.1. With an A296 of 0.431, and an A320 of 0.094, I had a total phage count of 0.28×10^{13} in 1 ml of PBS.

To biotinylate purified EIIIA, I used Chromalink Biotin Protein Labeling system (Solulink), which contains biotin as well as a chromatophore, for quantification of protein labeling efficacy. Using 100 μg of EIIIA (10.5kDa), as assessed by nanodrop A280, in 100mM phosphate buffer +150mM NaCl, pH 7.4, I added 38.6 μg of NHS biotin, aiming for approximately 5 biotins per molecule. Following 90 minutes of incubation at room temperature, I used a Zeba column (Thermo Fisher) to exchange the solution into PBS. With an A280 measurement of 0.047 and an A354 measurement of 0.0955, 100 μl of

biotinylated protein, and a protein molecular weight of 10.5 kDa, there were approximately 2 biotins/protein, as calculated using the E1% ChromaLink Biotin MSR Calculator available from Solulink.

Two vials of streptavidin beads were prepared, one for a negative selection against the beads themselves, and one for positive selection against antigen bound to the beads. Following two washes in PBS, both sets of beads (100μl of Dynabeads MyOne Streptavidin T1) were blocked in 1ml of 2% bovine serum albumin (BSA) in PBS for one hour at room temperature. Following two additional washes in PBS, the beads for negative selection were set aside. To the beads for positive selection, approximately 20 μg of biotinylated protein was added in 500μl of blocking solution (2% BSA in PBS). This protein-bead mixture was incubated for 30 minutes at room temperature with inversion, and washed twice in PBS prior to use.

To negatively select for antibodies which bind to beads, not antigen, 200 μl of phage (0.056×10^{13}) were added to the beads previously prepared for negative selection (blocked, without added antigen). Following 1 hour of incubation at room temperature with inversion, the supernatant was collected, and added to the EIIIA-biotin-streptavidin beads. This was again incubated for 1 hour at room temperature with inversion. After 15 washes in PBS containing 0.1% tween-20, the beads were washed a final time in this same wash buffer for 15 minutes with inversion at 37°C. To the beads containing phage bound to EIIIA, I added 500 μl of stationary phase ER2738 bacteria, and again incubated the solution for 15 minutes with inversion at 37°C. The cells were pipetted from the beads to a new Eppendorf tube, as elution 1. To the beads, I added 500 μl of 0.2M

glycine, pH 2.2, and allowed the solution to incubate for 10 minutes at room temperature with inversion. The glycine buffer was removed from the beads, and neutralized with 75μl of 1M Tris, pH 9.1, generating elution 2. Elutions 1 and 2 were combined, and this resulting solution was incubated for 15 minutes at 37°C. This combined elution 1 and elution 2 solution was plated onto two 15 cm plates containing 2YT medium supplemented with 2% glucose, 10μg/ml tetracycline, and 50μg/ml ampicillin, and grown overnight at 37°C.

Day 3: Each plate grown overnight was scraped down with 2 mL S.O.C. containing 50μg/ml ampicillin. The resulting solutions were pooled, 50% sterile glycerol was added to reach a final concentration of 15%, and the solutions were stored at -80°C as the first-round panning library.

ii. Panning round 2

Day 1: I used the same protocol for round 2 panning as I did for round 1, using 200 μL of thawed round 1 library stock instead of Riuniti library stock. With a starting OD_{600} of 0.159, the cells reached an OD_{600} of 0.796 after 1 hour and 50 minutes of growth.

Day 2: With slight modifications, the same protocol used in round 1 was used in round 2. Phage were prepared for panning using the same protocols used in round 1. From this phage preparation, I recovered 1.59×10^{13} phage, based on an A296 of 0.330 and A330 of

0.139. The biotinylated antigen prepared for round 1 was also used in round 2 (two days after biotinylation). Instead of using 100 μl of streptavidin beads and 20 μg EIIIA, 40 μl of beads and 2 μg EIIIA were used in round 2, resulting in greater stringency during panning. Likewise, as antibodies binding the beads themselves were already selected against in round 1, I did not have a round of negative selection during the second round of panning. To the blocked EIIIA-biotin-streptavidin beads, 2 μl of pre-cleared phage was added, and incubated for 15 minutes at room temperature with inversion. In order to select for individual colonies, and to determine a titer, 100 μl of the pooled elutions were removed, and six 10-fold serial dilutions were made. 90-100μl of each dilution were plated onto 10cm plates containing 2YT medium supplemented with 2% glucose, 10μg/ml tetracycline, and 50μg/ml ampicillin. As in round 1, the remaining pooled elution was plated onto two 15 cm plates containing 2YT medium supplemented with 2% glucose, 10μg/ml tetracycline, and 50μg/ml ampicillin. All of the plates were grown overnight at 37°C.

Day 3: As before, big plates were scraped, and stored in 15% glycerol as round 2 library stock. Small plates were stored at -4°C and used for picking individual colonies.

d. ELISA

A "master plate" was made by picking 96 individual colonies from round 2 plates. I grew the colonies in 200 μl S.O.C. containing 50 μg/ml ampicillin and 10 μg/ml tetracycline. The 96

well plate was covered with an AirPore sheet (Qiagen), and grown overnight at 37°C with agitation. For storage, 50% glycerol was added to a final concentration of 15% (85μl/well).

For the ELISA, a "soluble plate" was made. 180 μl 2YT with 50 μg/ml ampicillin and 10 μg/ml tetracycline was added to each well of a new 96 well plate. Approximately 2μl of solution was added from each well of the master plate, the plate was covered with an AirPore sheet, and incubated for 4 hours at 37°C with agitation. To induce antibody expression, an additional 80 μl of media (2YT with 50 μg/ml ampicillin and 10 μg/ml tetracycline) containing 10mM IPTG was added to each well, and protein expression was induced overnight at 30°C.

ELISA plates coated with EIIIA were made by adding 100 μl of a 0.5-2 μg/ml EIIIA solution in PBS, and incubating the plate overnight at 4°C. The EIIIA concentrations used in ELISAs were determined using a bicinchoninic acid (BCA) assay. From the BCA assay, the stock concentration of EIIIA was determined to be 0.4 mg/mL, 10 fold different from the 4.4 mg/ml determined by nanodrop. The concentration of EIIIA as determined by nanodrop was used for panning, and thus, throughout both rounds of panning, the stringency could have been very high. This could have led to a loss of active phage, but also could have selected for antibodies with greater binding affinity.

Following overnight incubation with antigen, the ELISA plate was blocked with 200 μl PBS with 4% milk and 1% Tween20 for 2 hours at 37°C. Meanwhile, the soluble plate was spun down at 2500 rpm at 4°C for 10 minutes. 50 μl of supernatant from the soluble plate was added to the ELISA plate containing 50 μl of the above blocking solution, and incubated for 1 hour at room temperature with rotation. To remove unbound antibodies, the plate was washed three

times in PBS with 1% Tween-20, followed by three washes in PBS. Anti-E-tag-HRP at 1:10,000 (Bethyl), in PBS containing 4% milk and 1% Tween-20, was added and incubated for 1 hour at room temperature with rotation. This was followed by the same series of three washes in PBS with 1% Tween-20 and three washes in PBS. The plate was developed by adding 100 μl TMB developing reagents/well. The development was stopped once significant blue color was present, up to 1 hour. To stop the reaction, I used 100 μl 1N HCl, and read the plate immediately at 450 nm. The values obtained from A450 are shown in Figure 3.3 for the ELISA plates coated with 0.05 (a), 0.1 (b) and 0.2 (c) μg EIIIA per well. In Figure 3.3, the positive hits, or those which appeared to have antibody-antigen binding as determined by an arbitrary cutoff, are indicated in grey.

From the three ELISAs, 24 wells appeared to be promising, and were carried forward for sequencing. The master plate was thawed slightly and for each well, 5 ml of L.B. with 50 μg/ml ampicillin was inoculated with a pipette tip dipped in the well to transfer ice crystals containing bacteria/plasmid. Cultures were grown overnight, mini-prepped using a Qiagen mini-prep kit, and sent for sequencing using the primer 5'-TCCGGCTCGTATGTTGTGTGGAAT-3'. Figure 3.4 shows the wells which were mini-prepped (a) and the sequencing results and clustering as assessed using Clustal W (b).

(a)	1	2	3	4	5	6	7	8	9	10	11	12
A	0.10	0.10	0.09	0.10	0.09	0.09	0.11	0.11	0.09	0.09	0.13	0.09
B	0.09	0.12	0.10	0.10	0.11	0.10	0.10	0.09	0.09	0.11	0.08	0.11
C	0.22	0.16	0.11	0.10	0.12	0.12	0.10	0.11	0.11	0.10	0.10	0.09
D	0.23	0.16	0.12	0.16	0.16	0.15	0.16	0.09	0.12	0.19	0.10	0.16
E	0.09	0.08	0.15	0.35	0.19	0.09	0.10	0.09	0.11	0.10	0.11	0.09
F	0.13	0.11	0.09	0.26	0.12	0.12	0.12	0.09	0.10	0.09	0.13	0.08
G	0.08	0.16	0.11	0.09	0.11	0.09	0.12	0.11	0.09	0.10	0.10	0.08
H	0.18	0.09	0.09	0.12	0.09	0.11	0.11	0.11	0.09	0.08	0.09	0.08

(b)	1	2	3	4	5	6	7	8	9	10	11	12
A	0.15	0.21	0.10	0.09	0.10	0.16	0.15	0.15	0.18	0.11	0.32	0.15
B	0.10	0.23	0.14	0.15	0.18	0.13	0.10	0.10	0.18	0.17	0.10	0.13
C	0.18	0.20	0.25	0.10	0.24	0.18	0.13	0.21	0.15	0.12	0.15	0.09
D	0.27	0.16	0.32	0.25	0.14	0.22	0.15	0.12	0.25	0.22	0.41	0.12
E	0.12	0.13	0.32	0.17	0.16	0.11	0.10	0.10	0.13	0.15	0.28	0.11
F	0.19	0.21	0.13	0.23	0.22	0.26	0.21	0.09	0.12	0.09	0.19	0.10
G	0.17	0.21	0.14	0.10	0.13	0.09	0.14	0.16	0.09	0.10	0.19	0.09
H	0.23	0.13	0.18	0.20	0.19	0.18	0.15	0.21	0.12	0.10	0.10	0.09

(c)	1	2	3	4	5	6	7	8	9	10	11	12
A	0.12	0.14	0.10	0.10	0.19	0.25	0.37	0.23	0.20	0.15	0.59	0.11
B	0.10	0.30	0.17	0.10	0.09	0.10	0.10	0.11	0.10	0.47	0.15	0.18
C	0.34	0.50	0.20	0.11	0.22	0.24	0.23	0.37	0.30	0.09	0.11	0.09
D	0.17	0.37	0.28	0.42	0.25	0.39	0.10	0.10	0.13	0.63	0.09	0.20
E	0.14	0.11	0.68	0.13	0.26	0.12	0.10	0.13	0.22	0.16	0.10	0.14
F	0.15	0.85	0.12	0.17	0.28	0.54	0.32	0.10	0.10	0.12	0.24	0.16
G	0.18	0.13	0.11	0.10	0.29	0.10	0.39	0.25	0.09	0.12	0.21	0.09
H	0.25	0.25	0.20	0.47	0.20	0.21	0.34	0.28	0.14	0.33	0.13	0.09

Figure 3.3. ELISA results, quantified by absorbance at 450 nm. Three separate ELISAs were performed using 0.05 (a), 0.1 (b) and 0.2 (c) μg EIIIA per well . Cutoffs for positive hits were arbitrary, and set at 0.1500 for (a), 0.2000 for (b), and 0.300 for (c). Positive hits are indicated in grey, and control wells without added phage supernatant are shown in black.

a. A11, B2, B10, C1, C2, C8, C9, D1, D2, D3, D4, D5, D6, D10, E3, F2, F4, F5, F6, F7, G2, H1, H4, H8

b.

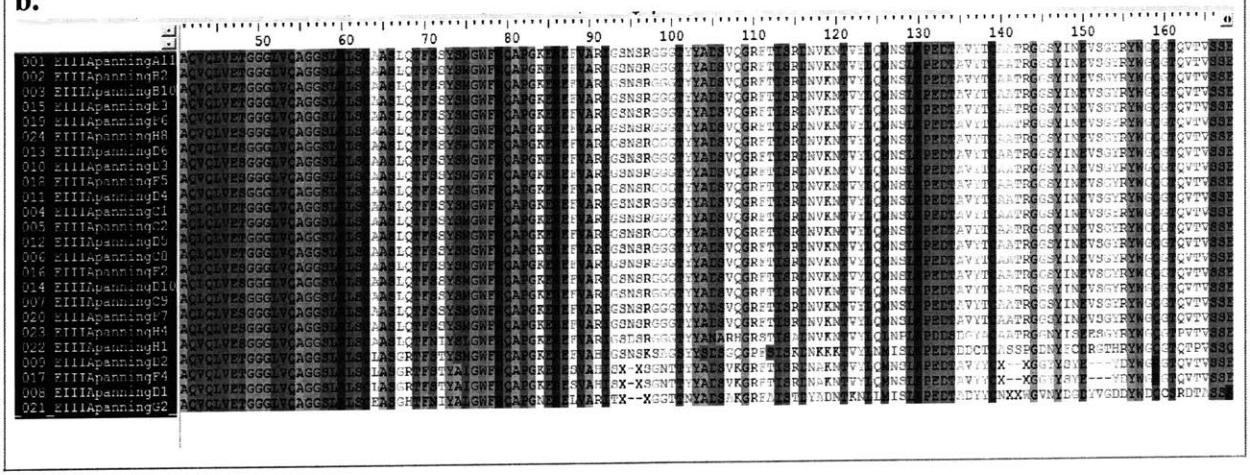


Figure 3.4. Positive hits and translated sequence alignments. Positive hits from the ELISA are listed in (a). In (b), sequences were aligned using the DNA region encoding the QVQL/QLQL motif. After using ClustalW to cluster similar sequences, the sequences were translated.

Figure 3.4 shows that there were five unique translated sequences from the plasmid DNA isolated from the 24 wells of interest. One translated sequence was present in 19 of 24 wells assessed. A different translated sequence was in 2 of 24, and 3 wells had unique sequences. The wells which were carried forward from these results were D2, E3, F4, G2, and H1. While there were 5 unique peptide sequences, there were more unique sequences at the DNA level due to the redundancy of the genetic code. However, because the resulting VHH was the same, these DNA level differences were discarded at this point and could be revisited later.

e. Antibody subcloning

After selecting clones of interest, the next step is to move the antibody sequences from the phagemid vector used during library construction and panning to a vector suitable for protein expression and subsequent purification.

To move the antibody sequences to a vector suitable for protein expression, I PCR amplified the DNA region of interest, used a Gibson Assembly reaction to ligate the PCR product into restriction-digested PHEN6 vector, and finally transformed *E. coli* with the plasmid for storage and propagation of the plasmid containing antibody sequences of interest.

Plasmid DNA was isolated from the five wells of interest using a Qiagen mini-prep kit and was used as the DNA template for PCR amplification. To set up the PCR reactions, I used 10 μM primers (Forward: 5'-CGCGGCCCAGCCGGCCATGGCCCAGGTGCAGCTCGTGG-3'; Reverse: 5'-AGTCCTCCTGAGGAGACGGTGACCTGGGTCCCCTGG-3'), 3 ng of DNA template, and 1 unit of Phusion DNA polymerase (New England Biolabs) in 1x HF buffer. Cycling parameters consisted of an initial denaturation at 98°C for 30 seconds, 35 cycles of 10 seconds of denaturation at 98°C, 10 seconds of annealing at 68°C and 40 seconds of extension at 72°C, followed by a final 5 minute extension cycle at 72°C. A representative result from the PCR amplification is shown in Figure 3.5a, indicating successful amplification of all five sequences of interest. Following PCR amplification, the samples were either gel-purified if there appeared to be multiple products, or cleaned-up using a PCR clean-up kit (Qiagen).

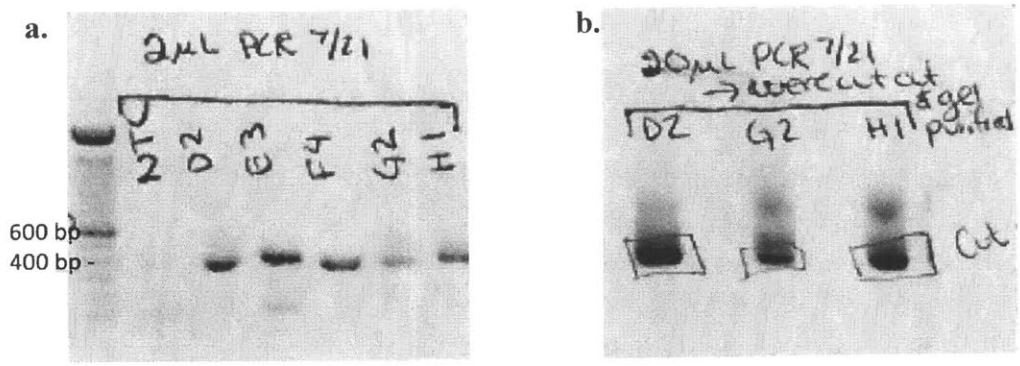
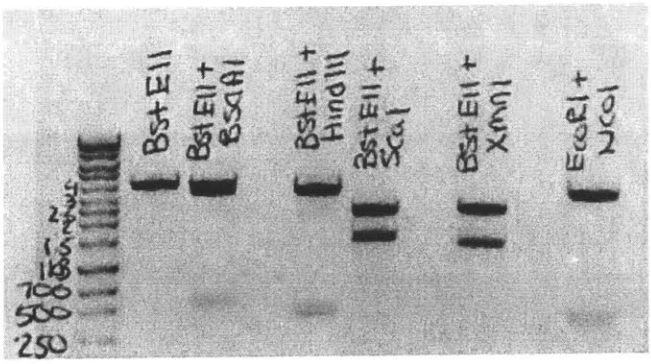


Figure 3.5. PCR amplification of antibody sequences. The no template control (NTC) contained all components of the PCR mixture with the exception of a DNA template. In (a), the PCR products were assessed for size and purity. (b) shows band cut out during gel purification of PCR amplification.

To ligate the purified PCR products into PHEN6, I used Gibson Assembly cloning. For each clone, the PCR products are combined in a 3-fold excess with digested and gel-purified pHEN6 in Gibson Assembly master mix containing Taq ligase, T5 exonuclease, a DNA polymerase, and the four dNTPs. This cloning works by T5 exonuclease chewing back 5' ends to leave a 20-40 basepair overlap between the insert and vector. The overlap allows the insert and vector to anneal. DNA polymerase can then add missing bases, and finally the ligase seals the nick. This cloning method relies on the generation of complementary overhangs in the PCR product, which were added by the primers used in the PCR amplification step. Since there are large variations in the antibody sequences, which could generate new restriction enzyme sites, Gibson Assembly cloning allowed me to use the same method for all of the antibodies, regardless of sequence variation.

To generate a large quantity of PHEN6 plasmid for Gibson Assembly cloning, I transformed DH5α cells with 0.1 ng of stock PHEN6 containing a random VHH, using standard protocols. Single colonies were selected and grown overnight in L.B. containing 50μg/ml ampicillin, then mini-prepped using standard protocols. To assess whether the vector was as expected, I performed a series of restriction digestions, as shown in Figure 3.6. The bands seen on gel electrophoresis for each digest were consistent with the expected band sizes.



Restriction Enzymes	Expected band(s)
BstEII	3610 bp
BstEII + BsaAI	3082 bp, 528 bp
BstEII + HindIII	3152 bp, 458 bp
BstEII + ScaI	2164 bp, 1446 bp
BstEII + XmnI	2283 bp, 1327 bp
EcoRI + NcoI	3186 bp, 424 bp

Figure 3.6. Restriction digestion of PHEN6 vector as shown on gel electrophoresis. Expected band sizes based on the vector map are also shown.

After validating that the mini-prepped PHEN6 contained the expected restriction sites, I prepared digested vector for use in Gibson Assembly cloning. I combined 2 μg of PHEN6, as determined using nanodrop absorbance measurements, with 10 units each of BstEII-HF and NcoI-HF in 1x cutsmart buffer (New England Biolabs). I allowed the reaction to proceed overnight at 37°C, then ran the DNA on a 1% agarose gel for gel purification, as shown in Figure 3.7.

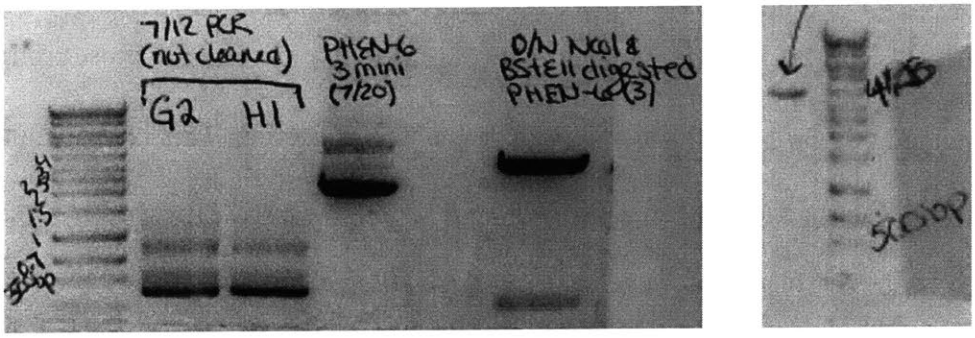


Figure 3.7. PHEN6 restriction digestion and gel purification for use in Gibson Assembly cloning.

I combined 77ng of PHEN6 (3.2 kB) with 29ng PCR product (~400bp) in a 5μl reaction volume using sterile water, yielding the desired 3:1 insert-to-vector ratio. The 5μl mixtures were added to 15μl isothermal assembly aliquots on ice, placed in a 50°C thermocycler for 30 minutes, and stored at -20°C. Fresh or stored Gibson Assemblies were used to transform XL-10 gold ultracompetent cells (Chem-Agilent). For this transformation, 1 to 2 μl of Gibson assembly reaction mix or control pUC18 were added to 50 μl cells, and a standard transformation protocol

supplied by the manufacturer was followed. Following 1 hour of recovery in S.O.C., the entire solution was plated onto L.B. plates containing 50 μg/ml ampicillin, and grown overnight at 37°C.

After overnight growth, I picked individual colonies to grow overnight in 5 ml L.B. with 50 μg/ml ampicillin to miniprep and make 15% glycerol stocks. To check whether there was a VHH within the PHEN6 plasmid, I restriction-digested the mini-prepped plasmids using BstEII and NcoI, using a 1 hour digestion at 37°C. Although cells transformed with plasmids derived from D2, E3, F4, G2, and H1 all exhibited some level of colony growth, only those from D2 showed the expected 400 bp band following BstEII and NcoI digestion. This is shown in Figure 3.8. D2-1, D2-2, and D2-3 were sent for sequencing, and will be carried forward by others in the laboratory.

The colonies which did not have the expected 400 bp band following BstEII and NcoI digestion were likely untransformed bacteria that were able to grow on regions of the plates where antibiotic concentration was very low, given the low ampicillin concentration in the plates. The data do not provide a clear answer as to why sub-cloning failed for clones E3, F4, G2, and H1. Gibson assembly may have been the step which was unsuccessful, as PCR amplification appeared to be successful in all cases, and the transformations were repeated multiple times and were successful for control pUC18. The entire process, from PCR amplification to transformation will need to be redone for E3, F4, G2, and H1, as over-expression and purification of VHH resulting from those sequences cannot be performed until the DNA is in a suitable expression vector.

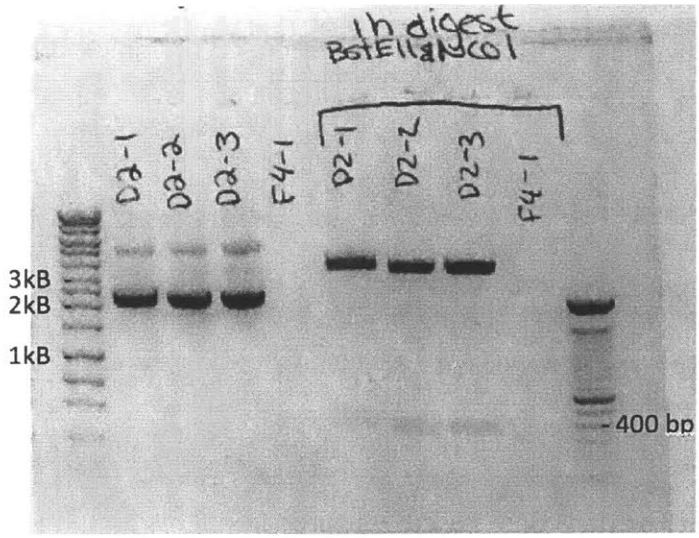


Figure 3.8. Restriction digestion of miniprepped PHEN6-VHH plasmid following Gibson Assembly and transformation of XL-10 gold ultracompetent cells.

4. Conclusions and future directions

Using the pET SUMO expression system and a nickel column purification system, I was able to generate purified EIIIA, as shown by LC-MS and Western blotting. This purified EIIIA was then used in phage display panning to select for VHH that bind to EIIIA. Following two rounds of panning and three ELISAs, 20 wells out of 94 experimental wells indicated high levels of VHH-EIIIA binding. Sequence analysis of the 20 positive hits showed 5 unique peptide sequences, which continue to be sequences of interest. One VHH sequence was successfully sub-cloned into pHEN6. This construct can be used for purification, over-expression, and the resulting VHH can be validated and assessed as described below. Continued work will need to be carried out on the other four VHH sequences which were not successfully sub-cloned into PHEN6.

Typically, when immunizing alpacas to generate VHH, they are immunized with only a few antigens at a time. Here, the VHH library used to develop anti-EIIIA VHH was generated from an alpaca that was immunized with approximately 20 unique proteins and peptides. Successful panning from this library indicates that it may be possible to isolate VHH against novel protein targets from VHH libraries generated by alpacas immunized with complex mixtures. Future directions of the anti-EIIIA VHH are described below.

a. Purification of VHH

Competent cells which have favorable properties for protein overexpression, such as WK6, will need to be transformed with pHEN6-VHH in order to overexpress and purify VHH. After overnight inoculation, expression of the VHH can be induced with IPTG to generate VHH

localized periplasmically in *E. coli*. An osmotic shock using sucrose is sufficient to release the VHH from the periplasm and purification can proceed using nickel or cobalt columns, as the VHH should be synthesized with a C-terminal 6X-His tag which is downstream of the Gibson assembly insertion site in pHEN6. Further purification could also be perfomed using gel filtration.

b. VHH validation

Methods which could be used to validate that the VHH bind to the desired target antigen include ELISA, immuno-localization, Western blotting, and immunoprecipitation. For the ELISA, purified EIIIA which was generated for panning could be used. After coating an ELISA plate with a panel of EIIIA concentrations, a similar protocol as that used during panning could be used. The major exceptions are that instead of using bacterial supernatant, purified antibody would be used, and an anti-His antibody should be used instead of an anti-E-tag antibody.

For immunolocalization, cancer cell lines which are known to express EIIIA and/or EIIIB would be useful. An ideal negative control would be these same cell lines with EIIIA and/or EIIIB genetically removed. A post-doc in the laboratory, Patrick Murphy, has generated MMTV-PyMT lines with fibronectin splice variants AB-/- and AB +/- (MMTV-PyMT AB +/- and -/-, respectively). The AB -/- line would be a negative control, with a lab-based or commerically available anti-EIIIA antibody as a positive control. Following three days of growth on chamber slides in fibronectin-depleted media, the VHH could be added in serum-free media. Following a set incubation time, the cells could then be washed and fixed in paraformaldehyde. This order, of adding primary antibody prior to fixation would help to circumvent the issue of antigens being

hidden following paraformaldahyde cross-linking as well as permeabilzation of the cells during fixation yielding higher background. As the VHH will have a C-terminal 6x-His tag following purification, an anti-His-GFP antibody could be used for microscopy. These antibodies could also be tested on tissue sections, as well as on different human and mouse cell lines.

c. Possible uses

1. In vivo imaging of mice with syngeneic breast tumors

One way to test the VHHs in vivo, would be to use a syngeneic breast cancer model and image using VHH-fluorophore conjugates. Two examples of cells and mouse strains which could be used are MMTV-PyMT AB +/- cells implanted into C57BL/6 mice, and the 4T1 family of cell lines in BALB/C mice. Using luciferase-labeled cells, tumor progression could be monitored using in vivo bioluminescence imaging. The VHH-fluorophore conjugates could be made a number of ways, such as (1) genetically encoding GFP or RFP at the C-terminal of the VHH, (2) sub-cloning the VHH into a vector containing a C-terminal LXPTG motif and using fluorophores with an N-terminal poly-glycine to couple the fluorophore and VHH using sortase[24], or (3) biotinylating VHH and coupling a streptavidin-tagged fluorophore. Sortase would give site-specific labeling, while biotinylation would only be specific to primary amines.

Using mice with implanted tumors, palpation and IVIS in vivo bioluminescence could be used to track tumor progression and burden. At a set time-point or tumor size, VHH could be injected via the lateral tail vein. A reasonable starting dose could be 20 μg

– a dose which was used in a different study using VHH targeting MHC class II and CD11b[25].

Due to the short half-life in circulation of VHH (~10-20 min)[25], imaging should begin quickly post-injection using IVIS in vivo bioluminescence and fluorescence imaging. It is unclear how long the VHHs will continue to stay bound to their targets, as this depends on pharmacokinetics, pharmacodynamics, and ECM remodeling, so imaging time-points most likely would need to be dependent on the results. A control cohort of mice without tumors would check for non-tumor binding. Through comparing the localization of luciferase-positive tumor cells and fluorophore-conjugated VHH, it would be possible to get an idea for the specificity and sensitivity of particular VHHs. If the resolution using whole-mouse live in vivo imaging is insufficient, individual organs could be imaged post-mortem, and tissue slides could also be made as needed. Another alternative is to use radiolabeled VHH, isolate individual organs, and assess bio-distribution through radioactivity quantification[11].

2. Localized immunomodulation of tumors

VHH with high tissue penetration could be used to modulate the local tumor microenvironment. Many studies have conjugated immunocytokines, including IL-2, IL-12, and TNF-α to tumor-targeting antibodies to induce a localized immune response. However, systemic toxicity may be a problem, and the actual localization of these antibody-cytokine conjugates may depend on the relative binding affinities of the two components for their respective ligands[26]. This will have to be tested.

Another option could be to a activate a pre-existing immune response and localize it to the tumor extracellular matrix. This could be done by coupling an antigen which is used in vaccinations to EIIIA-VHH. One example of this would be to couple the B-subunit of diphtheria toxin to the anti-EIIIA-VHH. By itself, the B-subunit is relatively non-toxic, as it is the A-subunit which is responsible for the ADP-ribosylation of eEF-2 that causes the toxicity associated with diphtheria. Mice could be vaccinated with a commercially available DTP vaccine used in humans. Introduction of the diphtheria B-subunit coupled to EIIIA-VHH could induce a rapid immune response in vaccinated mice which theoretically would be localized to the tumor extracellular matrix. The effect of this could recruit large numbers of immune cells while also inducing an increased immune response against tumor antigens, and would be a new approach to cancer immunotherapy.

References

1. Naba, Clauser et al. (2012) *Mol. Cell. Proteomics* 11(4):M111.014647.
2. Naba et al. (2014) *eLife* 3:e01308.
3. Naba et al. (2014) *BMC Cancer* 14:518.
4. Hynes. (1990). Fibronectins. Springer-Verlag, New York.
5. Pankov et al. (2002) *J. Cell Sci.* 115: 3861-3863
6. Astrof et al. (2007) *Dev. Bio.* 311(1):11-24.
7. Matsumoto et al. (1999) *Jpn. J. Cancer Res.* 90:320-325.
8. Santimaria et al. (2003) *Clin. Cancer Res.* 9:571-579.
9. Sun et al. (2014) *Theranostics* 4(8): 845-857.
10. Kumra et al. (2016) *Adv. Drug Dev. Rev.* 97:101-110.
11. Schilemann et al. (2009) *Blood* 113(10): 2275-83.
12. Kiefer et al. (2016) *Immunological reviews* 270:178-192.
13. Pasche et al. (2012) *Angiogenesis* 15(1):165-169.
14. Franz et al. (2013) *J. Heart Lung Transplant* 32(6): 641-650.
15. Doll et al. (2013) *Arthritis Res. Ther.* 15(5): R138.
16. Huston et al. (1988) *Proc. Natl. Acad. Sci . U.S.A.* 85(16):5879–5883.
17. Janeway et al. (2001) Immunobiology, 5e. Garland Science, New York.
18. De Vos et al. (2013) *Expert Opin. Biol. Ther.* 13(8):1149-60.
19. Conrath et al. (2002) *Dev. Comp. Immunol.* 27:83-103.
20. Panja et al. (2008) *Biomacromolecules* 9(9):2501-2509.
21. Chan et al. (2013) *Biosci Rep.* 33(6):e00086.
22. Becker et al. (2013) *Nucl. Acids Res.* 41(1):156-66.
23. Pande et al. (2010) *Biotechnol. Adv.* 28(6):849-858.
24. Pardon et al. (2014) *Nat. Protoc.* 9(5): 674-693.
25. Rashidian et al. (2015) *Proc. Natl. Acad. Sci . U.S.A.* 112(19):6146-6151.
26. Tzeng et al. (2015) *Proc. Natl. Acad. Sci . U.S.A.* 112(11):3320-3325.