THE PENNSYLVANIA STATE UNIVERSITY
THE SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

WESTERN BLOT MARKER WITH UNIVERSAL IMMUNOGLOBULIN BINDING DOMAIN

ANDREW FLEISCHMAN

SPRING 2010

A thesis
submitted in partial fulfillment
Of the requirements
for a baccalaureate degree
in Science
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Song Tan
Associate Professor of Biochemistry and Molecular Biology
Thesis Supervisor

Ming Tien
Professor of Biochemistry and Molecular Biology
Honors Advisor

Scott Selleck
Department Head
Department of Biochemistry and Molecular Biology

*Signatures are on file in the Schreyer Honors College.

Abstract

Immunological techniques such as western blotting harness the specificity and avidity of antibodies to detect proteins. In western blotting, antibodies specifically detect proteins separated by SDS-PAGE electrophoresis. Molecular weight markers are important components of a western blot because they allow one to verify protein sizes. However, proteins in a traditional molecular weight marker will not appear on the film of a western blot, since the various primary antibodies cannot recognize them. The goal of this thesis research is to develop universal western blot molecular weight markers that bind commonly used secondary antibodies. The IgG Binding Domain of Protein A (PRA), a domain that binds the heavy chain region of antibodies, was fused to proteins of various sizes. Each recombinant construct was expressed and purified on a small scale to determine the optimal expression conditions and to check the solubility of the protein. These optimal conditions were then used to express the protein on a larger scale. Each marker protein was purified by metal affinity chromatography via the N-terminal polyhistidine tag. The purified proteins were combined into a mixture, producing a set of western blotting molecular weight markers that be used independent of the primary antibody employed.

Table of Contents

List of Figures and Tables

Figures

Tables

Acknowledgments

I would like to sincerely thank all of the current and past members of the Tan Lab who have helped to develop my skills in the lab and aided me throughout my thesis project. In particular, I would like to thank Dr. Song Tan, for giving me the opportunity to learn in a challenging, yet productive environment. My experiences from the past three years will have long lasting effects on my education and career. I also would like to thank the graduate students Jiehuan Huang, Matt Jennings, Tom Koerber, Manoj Saxena, and Gurangad Chandok for their assistance for this project, and my former lab mentor Joe England. Finally, I would like to thank Dr. Ming Tien and Dr. Scott Selleck for reviewing this thesis report.

1. Introduction

1.1 Immunological Applications

Immunology is the branch of biomedical science that examines the immune system and the physiology underlying the detection and protection of an organism from foreign substances. The most familiar aspect of the immune system is humoral immunity, where secreted immunoglobulin molecules (also known as antibodies) bind to a target antigen and induce a response to neutralize the intruding molecule (Chapter 4, Kuby Immunology, 6th edition, 2006). The high specificity and affinity of the association between antibody and antigen has provided a powerful tool for detection of proteins and other molecules in molecular biology and medicine.

The diversity of antibodies is unparalleled, as there are about 10 billion generated by the human immune system, each targeting a distinct epitope (Chapter 5, Kuby Immunology, 6th edition, 2006). Remarkably, the human genome contains only around 30,000 genes, and far fewer associated specifically with antibody production. Such diversity can be attributed to combinatorial recombination in the hyper-variable regions of antibody structure as well as somatic hypermutation (Chapter 5, Kuby Immunology, 6th edition, 2006). Nevertheless, antibodies maintain their classical structures in the constant regions that maintain their functional abilities integral to humoral immunity (Chapter 4, Kuby Immunology, 6th edition, 2006).

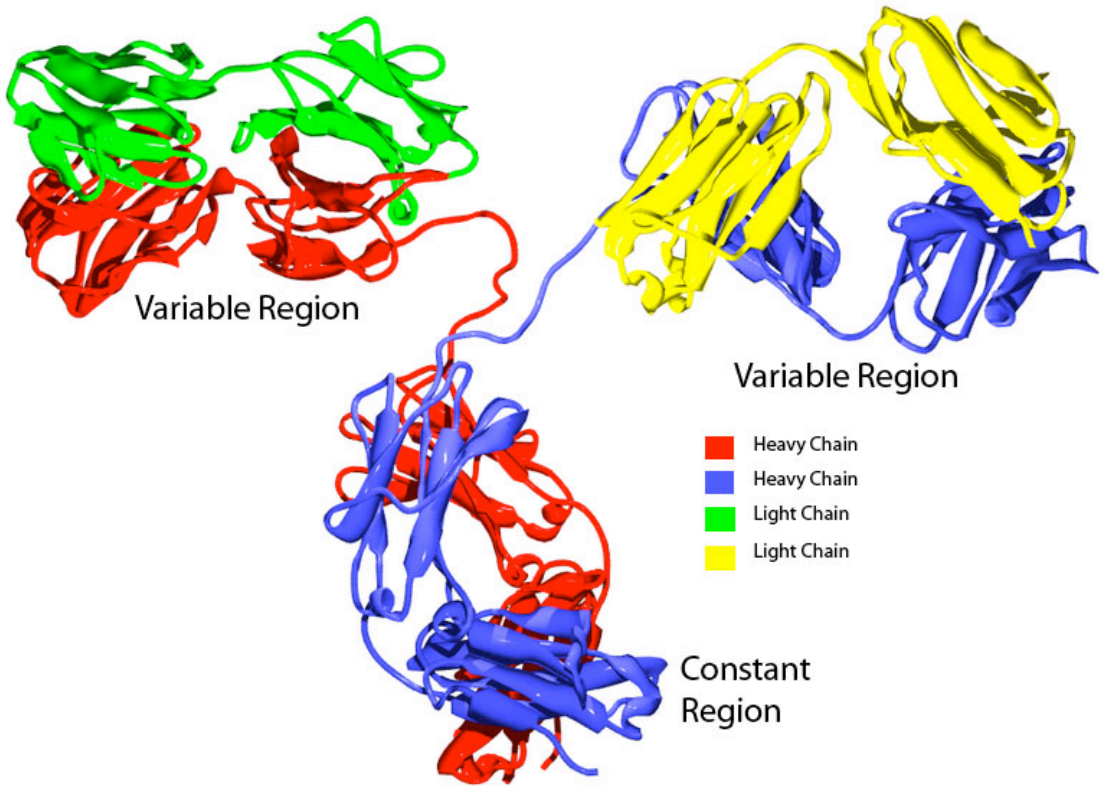


Figure 1. Structure of IgG antibody (modified from The Biochemistry Questions Site) http://biochemistryquestions.wordpress.com/category/proteins

The high specificity and affinity of antibodies is attributed to the hyper-variable region, which utilize an aggregate of weak forces to bind to the associated antigen. The diversity of the hyper-variable regions maximizes the probability an antibody can be produced to target a specific foreign antigen (Chapter 4, Kuby Immunology, 6th edition, 2006). While each of the billions of antibodies initially exists at an extremely low concentration, an antibody will rapidly proliferate in the serum upon contact and binding to its antigen.

Scientists produce specific antibodies by injecting particular antigens into a mammalian host, such as a rabbit, mouse, or goat. These antibodies are considered to be polyclonal, as many different antibodies that bind to different epitopes of a single antigen will proliferate in the serum. In contrast, monoclonal antibodies specific for a single epitope

are propagated by fusing isolated lymphocytes to cancer cells, and then purified from the resultant antiserum (Chapter 6, Kuby Immunology, 6th edition, 2006).

The specificity of antibodies has been exploited to identify and locate target molecules and to detect molecules at very low concentrations. The antibody for the target molecule is typically attached to a radiolabel, fluorescent label, or detectable enzyme. These conjugated reporters can provide either quantitative or qualitative measures of the antibody after binding to the antigen. Well-known techniques such as immunoblotting, ELISA, flow cytometry, and immunofluorescence use antibodies for detection (Chapter 6, Kuby Immunology, 6th edition, 2006).

This thesis will focus on the technique of immunoblotting, also known as western blotting, which utilizes antibodies to analyze proteins separated by electrophoresis (Renart, 1979). Scientists most commonly use electrophoresis gels containing sodium dodecyl sulfate (SDS), which denatures the molecules and introduces a negative charge for every two amino acids that is proportional to mass, ~1.4g/ gram of protein (Laemmli, 1970 and Mullins, 2002). Proteins separate according to size (smaller molecules migrate faster) and are then transferred to a nitrocellulose membrane. Antibodies for thousands of different proteins are commercially available and are used to probe the proteins on the membrane. Usually, the primary antibody use is specific for the desired antigen and a reporter-linked secondary antibody binds to the primary antibody (Chapter 6, Kuby Immunology, 6th edition, 2006). Since the secondary antibody often recognizes entire classes of primary antibodies, such as anti- rabbit IgG or anti-mouse IgG, the same secondary antibody can often be paired with many different primary antibodies with

significant costs benefits. The bound secondary antibody is then detected on the immunoblot by methods such as chemiluminescence, with the emitted light captured on photosensitive film.

Molecular weight markers are important components of gel electrophoresis and western blotting experiments. Such markers, a mixture of proteins of known molecular sizes, allow scientists to infer the molecular weight of the protein of interest by comparing its gel mobility with those of the proteins in the marker set. However, the proteins in a generic molecular marker set cannot be detected on a western blot since the primary antibody will not necessarily recognize the molecular weight marker proteins (the primary antibody is chosen to detect the protein of interest, not the molecular weight marker proteins). Colored or prestained proteins are available that can transfer to the membrane, but these do not appear in the developed film. Consequently, one typically has to overlay the western blot film with the blot itself to mark the blot with the positions of the molecular weight markers, an unsatisfactory and inaccurate process. While protein markers specific for western blots are available, these are cost prohibitive. Therefore, the goal of this thesis project was to produce a set of western blot marker on a large-scale as a low cost alternative.

1.2 Protein Binding Domain

I wanted to generate protein markers that could be visualized on any western blot. This required that the reporter-linked secondary antibody be able to bind all of the molecular marker proteins. The challenge was to find a set of proteins that would bind to any of the most commonly used secondary antibodies.

Protein A, originally found in bacterial cell walls of Stapholococcus aureus, has been used throughout the field of molecular biology due to its inherent ability to bind antibodies. A particular domain of Protein A possesses a high affinity for the heavy chain region of antibodies, known as the F_c or fragment crystallizable region (Moks, 2005). The F_c region of antibodies is highly conserved and constant, as this region facilitates antibody binding to receptors to mediate its effect and activate the complement cascade (Chapter 4 and 7, Kuby Immunology, 6th edition, 2006). This is in contrast to the hyper-variable region of the antibody, which mediates the high affinity for a specific target molecule.

By fusing the immunoglobulin (IgG) binding domain of Protein A as a fusion tag adjacent to a protein, the resulting fusion protein should bind to any commonly used secondary antibodies. When used in a western blotting experiment, a set of appropriate fusion molecular marker proteins will be visible on the exposed film. Thus the experimenter will be able to visualize the antigen in question as well as the molecular weight markers without modifying the western blotting procedure. Methods to use the IgG binding capability of Protein A in a molecular weight marker were first developed by Kihira and Aiba (U.S. Patent #5580788, 1996).

2. Materials and Methods

2.1 IgG Binding Domain Insert Preparation

In order to incorporate the IgG Binding Domain of Protein A as a fusion tag, the recombinant DNA sequence was obtained as a plasmid and prepared for insertion into an expression vector.

2.1.1 PCR Amplification

The polymerase chain reaction (PCR) was used to amplify the IgG binding domain of Protein A (PRA) from Protein A and to incorporate restriction sites to be used for subcloning this coding region into an expression vector (Mullis, 1986). This method allowed for the flexibility to add specific restriction sites on the ends of the product by designing custom primers. Before preparing the PCR mixture, the template plasmid DNA was diluted to 10 ng/µl with TE (10, 0.1) [10 mM Tris, 0.1 mM EDTA]. The PCR mixture was prepared with 68.5 µl of water, 10 µl of 10x Thermo Pol buffer, 10 µl of 2.5 mM dNTP, 1 µl of 10 ng/µl template plasmid, 5 µl of 10 µM forward primer, 5 µl of 10 µm reverse primer, and 0.5 µl of 2 units/µl Pfu polymerase (see Appendix A for reaction mixture). The PCR mix was placed in the thermocycler:

2 min, 95˚C → 5 x [30 sec, 95˚C → 30 sec, T_m – 5˚C → Ext, 75˚C]

→ 25 x [30 sec, 95˚C → 30 sec, 60˚C → Ext, 75˚C] → 3 min, 75˚C

→ 15˚C.

(T_m is the melting temperature of the primers and Ext is the extension time of 1 minute per kilobase of the expected fragment). The PCR product was analyzed on an agarose gel to see that the desired size fragment was amplified in the reaction (see section 2.4). In one case, the PCR product was directly ligated into a “blunt-end” vector (see section 2.1.2).

In order to ligate the product into an expression vector, the PCR product was digested with the appropriate restriction enzymes and agarose gel purified (see sections 2.6 and 2.5).

2.1.2 Blunt Ligation of PCR Product

My initial attempts to subclone the PRA coding region using the engineered flanking restriction sites were unsuccessful. I therefore used an alternative blunt-end cloning approach to subclone the PRA coding region. The plasmid vector was digested by SmaI restriction endonuclease, which produces blunt DNA ends. These blunt ends are compatible with the blunt ends of the PCR product. Prior to ligation, the vector and PCR product were prepared as follows. First, the vector was digested with a blunt-end restriction endonuclease and alkaline phosphatase (see section 2.7). The PRA PCR product was extracted with phenol/CIA and ethanol precipitated (see section 2.2), before being phosphorylated by a kinase enzyme (see section 2.3). Both the PCR product and vector were purified by agarose gel purification (see section 2.5). After the blunt-ended insert and vector DNA were isolated, the ligation mixture was prepared, with both vector and insert, and a control containing only vector DNA. The ligation mixture contained 4 µl of water, 1 µl of 10x T4 DNA ligase buffer, 0.5 µl of 100 mM DTT, 2 µl of gel-purified vector DNA, 1.5 µl of gel purified insert DNA, and 1 µl of 750 units/µl T4 DNA ligase (see appendix A for reaction mixtures). The ligation mixture was incubated for 15 minutes at room temperature before proceeding to the transformation and plasmid preparation (see section 2.9 and 2.11). At the same time, a second ligation mixture was prepared as a control, which contained only vector DNA and no insert fragment.

2.1.3 Polymerase Chain Reaction Site-Directed Mutagenesis

I employed PCR site-directed mutagenesis, based on the QuickChange method by Stratagene, to remove an undesirable BglII restriction site within the PRA region by silent mutagenesis. The forward primer, and its reverse complement as the reverse primer, was designed complementary to the template sequence, but with a single base change that eliminated the BglII site without changing the translated sequence. Before preparing the PCR reaction mixture, the template construct was diluted to 10 ng/µl in TE(10, 0.1) [10 mM Tris, 0.1 mM EDTA]. The PCR reaction was prepared with 17.6 µl water, 2.5 µl of 10x Thermo Pol buffer, 2.5 µl of 2.5 mM dNTP, 0.5 µl of 10 ng/µl DNA construct, 0.7 µl of 10 uM forward primer, 0.7 µl of 10 µM reverse primer, and 0.5 µl of 2 units/µl Pfu Turbo DNA polymerase (see Appendix A for reaction mixtures). The mixture was incubated in a thermocycler (Techne TC-312): 2 min, 95˚C → 12x [30 sec, 95˚C → 1 min, 55˚C → Ext, 68˚C] → 15˚C (Ext is the extension time of 1 minute per kilobase of the expected fragment). While 2 ul of PCR reaction mixture was saved in a sterile Eppendorf tube (mix A), 0.5 ul of 10 units/ul DpnI enzyme was added to the remainder of the PCR mix and incubated at 37˚C for one hour (mix B). After the incubation, both the sample from mix A and mix B were transformed in TG1 competent cells (see section 2.9). Only the colonies from mix B were screened and prepared as a plasmid. The transformation plate from mix A was used as a negative control to gauge the success of the mutagenesis procedure.

2.2 Phenol/CIA Extraction and Ethanol Precipitation

Phenol/CIA was used to extract protein from the PCR product. The CIA solution is a mixture of 24 volumes of chloroform and one volume of isoamyl alcohol. The CIA is

added to TE-equilibrated phenol in equal volumes in an Eppendorf tube and spun for one minute at 13000 rpm in a microcentrifuge (Sorvall Biofuge). The extraction was accomplished by adding equal volumes of phenol/CIA to the sample and vortexing for 15 seconds. Then the mixture was spun in a microcentrifuge at 13000 rpm for one minute (Sorvall Biofuge). The clear aqueous phase (above the white film interface) at the top of the Eppendorf was transferred to a new tube. After the phenol/CIA extraction, the DNA was ethanol precipitated by adding 0.1 volumes of 3 M Sodium Acetate (pH 5.2) and 2.5 volumes of pure ethanol. The mixture was vortexed for five seconds and spun in a microcentrifuge for 10 minutes at 13000 rpm (Sorvall Biofuge). After the spin, the supernatant was aspirated away and the small pellet was allowed to dry for 3 minutes before resuspending with 30 µl of TE (10, 0.1) [10 mM Tris, 0.1 mM EDTA].

2.3 Kinase PCR Product

After the phenol/CIA extraction and ethanol precipitation (see section 2.3), the PCR product was phosphorylated by a kinase to produce a compatible end for the dephosphorylated vector DNA. The kinase mixture contained 9 µl of water, 3 µl of 10x PNK buffer, 1.5 µl of 10 mM rATP, 1 µl of 100 mM DTT, 15 µl of PCR product, and 0.5 µl of 10 units/µl T4 PNK (see Appendix A for reaction mixture). The kinase reaction mixture was incubated in a 37˚C water bath for 15 minutes. After the DNA ends were kinased, the PCR product was agarose gel purified (see section 2.5).

2.4 Agarose Gel Electrophoresis

Agarose gel electrophoresis is a common method for separating DNA, since smaller fragments migrate more quickly and larger fragments remain closer to the top of the gel.

The 1% (w/v) agarose gel was created by combining 30 ml of 0.5x TBE [45 mM Tris, 45 mM boric acid, 1mM EDTA] with 0.30 g of (HGT) agarose in a 125 ml Erlenmeyer flask. Smaller fragments may require a gel of about 1.5% (w/v) agarose. About 5 ml of deionized water was also added to compensate for water lost while boiling in the microwave. The agarose was dissolved in the solution by microwaving the flask for 80 seconds. Meanwhile, the gel-casting tray was set up in a block and leveled. After microwaving, 1.5 µl of 10mg/ml ethidium bromide was mixed into the agarose solution and poured into the gel casting setup. A 10 well comb, accommodating 30 µl samples was installed for the preparative purposes of excising and purifying the electrophoresis product. A 15 well comb, accommodating only 10 µl samples was used for analytical purposes, such as PCR screening and restriction mapping. The liquid agarose solution was left to cool and solidify for 30 minutes (shorter in cold temperatures). The gel-casting tray was transferred to the electrophoresis box with the comb facing the negative electrode. The comb was removed and the box filled with sufficient 0.5xTBE gel buffer to cover the entire gel. The samples were each mixed with 6x gel loading buffer (GLB) [0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol, 60 mM EDTA] and loaded in the wells alongside an additional molecular weight marker. The gel was run at 125 V for about 40 minutes or until the bromophenol blue dye reached the bottom of the gel. Finally, the gel was photographed in a UV transilluminator to reveal the DNA bands in each well.

2.5 Agarose Gel Purification

The DNA product of the PCR reaction is contained in a mixture of enzymes, salts, and possibly contaminating DNA fragments. A simple method to purify the PCR product is to

electrophorese the sample on an agarose gel and excise the visualized fragment. This method was also used to isolate digested fragments for subcloning. The first step was to electrophorese the DNA on a preparative HGT agarose gel (10 well comb) as described above. A longer electrophoresis time was required to obtain good resolution between digestion fragments of similar size. While the gel was running, a filter assembly was prepared to isolate the DNA fragment. A 0.5 ml Eppendorf tube was pierced on the bottom with a flamed 25-gauge needle and stuffed with a small amount of siliconized glass wool. After the gel ran to completion, the gel was photographed in the transilluminator and the desired DNA band excised with a razor blade. The gel slice was placed inside the filter assembly, which in turn was placed inside a 1.5 ml Eppendorf tube. The assembly was centrifuged in a microcentrifuge for 5 minutes at 7000 rpm (Sorvall Biofuge). The gel-purified fragment was then collected in the 1.5 ml Eppendorf tube and was ready for ligation.

2.6 Restriction Endonuclease Digestion

Restriction endonucleases cleave double-stranded DNA at specific recognition sites. Digestions with two such enzymes were required to prepare the insert and vector DNA prior to subcloning into an expression vector. Similarly, a single enzyme digestion was also used to cut the blunt-end vector to insert a PCR fragment. The digestion mixture was prepared with the appropriate enzyme by adding water, the plasmid or PCR fragment, restriction enzyme buffer (New England Biolabs; Ipswich, MA), 0.1 mg/ml BSA, 3.3 mM DTT, and a sufficient amount of restriction enzyme (see Appendix A for reaction mixtures). The digestion mixture was incubated in a 37˚C water bath for 2 hours. When a single restriction enzyme buffer was not compatible with both restriction enzymes,

digestion with two different enzymes required two steps, first digesting with one enzyme in its appropriate buffer, and then changing the buffer conditions for the second digestion by adding 5 M NaCl (see Appendix A). The insert and vector DNA were then agarose gel purified and ligated (see section 2.5).

2.7 Alkaline Phosphatase Digestion

Alkaline phosphatase was used to dephosphorylate the blunt ends of the vector. Without digesting the ends with a phosophatase, the vector would self-ligate and be unavailable to ligate the insert. In the single-cut digestion mixture, 0.1 units of calf-intestinal phosphatase (CIP) was added to the vector DNA and incubated in a 50˚C water bath for 45 minutes. The vector was then gel purified on an agarose gel for blunt-end ligation.

2.8 Sticky-end Ligation

The most common ligation employs restriction endonucleases that leave an "overhang" after cleaving the double-stranded DNA. Such ligations were employed to subclone insert DNA into expression vectors. Two ligation mixtures were prepared, one with both vector and insert and a control containing only vector DNA. The ligation mixtures contained 4 µl of water, 1 µl of 10x T4 DNA ligase buffer, 0.5 µl of 100 mM DTT, 2 µl of gel-purified vector DNA, 1.5 µl of gel purified insert DNA (except in the control), and 1 µl of 40 units/µl T4 DNA ligase (see appendix A for reaction mixtures). The ligation mixtures were incubated for 15 minutes before proceeding to the transformation (see section 2.9).

2.9 Transformation

In order to grow and replicate the DNA construct in a bacterial host, the plasmid was

transformed and incorporated in competent cells. Such transformation steps were required after each ligation or before protein expression to grow the plasmid in *E. coli*. The competent cells varied for protocols requiring cloning processes and protein processes, as TG1 and BL21(DE3)pLysS *E. coli* cells were used, respectively. Each 100 µl aliquot of competent cells was thawed on ice after storage at -80˚ C. The aliquots of competent cells were mixed with 2 µl of the ligation mixtures (1 µl for protein expression) and put in ice for 15 to 45 minutes. After the incubation on ice, the cells were heat shocked at 42˚ C and placed back on ice for 10 to 20 seconds. Each aliquot then received 0.5 ml of 2xTY media [1.6% (w/v) bactotryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl], and was placed in a 37˚ C shaking incubator for 15 to 45 minutes. After the incubation period, 0.3 ml of cells was plated onto an appropriate TYE media plate [1.0% (w/v) bactotryptone, 0.5% (w/v) yeast extract, 0.8% (w/v) NaCl, 1.5% (w/v) agar] and 100 µg/ml of ampicillin. TYE plates used for protein expression also contained 25 ug/ml chloramphenicol. Transformation plates were grown in a 37˚ C incubator for 10 to 18 hours.

2.10 PCR Screening of Colonies

After the incubation period, the transformation plates contained bacterial colonies that could contain the desired plasmid. Before investing the time and resources into growing and purifying a DNA plasmid, a PCR screen was used to demonstrate that the desired plasmid was successfully incorporated. This was accomplished by choosing primers that would amplify a unique region of the desired plasmid, most likely flanking an insert or manipulated region. Ultimately, the amplified region for the desired fragment was a predicted size distinguishable from that of the original template plasmid. For site-directed

mutagenesis that removed a restriction site, the amplified region was digested with that specific enzyme to select the desired colony. Each single colony was prepared for screening (usually 4 to 8 colonies) by transferring the colony to 100 µl aliquots of water with a sterile loop and re-streaking the loop onto an additional TYE media plate. The re-streaked media plate was incubated in a 37˚C incubator for at least 8 hours. The cell suspensions in each aliquot were vortexed for three seconds before being added to the PCR reaction. The PCR reaction mixture was prepared with water, Thermo Pol buffer, 0.25 mM dNTP, 0.5 µM forward primer, 0.5 µM reverse primer, and 0.01 units/µl Pfu polymerase (see Appendix A for reaction mixtures). The reaction mixtures were scaled up depending on the amount of colonies prepared for screening (see Appendix A for reaction mixtures). For each colony, 1 µl of cell suspension was added to 19 µl of PCR reaction mixture for thermocycling: 2 min, 95˚C → 25x [30 sec, 95˚C → 30 sec, T_m – 10˚C → Ext, 75˚C] → 15˚C (T_m is the melting temperature of the primers and Ext is the extension time of 1 minute per kilobase of the expected fragment). After the PCR reactions, 3 µl of 6x GLB was added to the tubes to a final concentration of 1x GLB and analyzed on an appropriate analytical agarose gel (see section 2.4). If present, two positive clones displaying the predicted PCR product based on size were selected for plasmid preparation (see section 2.11).

2.11 Plasmid Preparation

Two positive clones from the PCR screening were chosen for plasmid preparation. For each positive clone, a single colony from the re-streak plate was used to inoculate a 500 ml Erlenmeyer flask with 100 ml of 2xTY media [1.6% (w/v) bactotryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl] with a sterile loop. Ampicillin (100 µg/ml) was also

added to the media before growing the cells in a 37˚C shaking incubator for at least 10 hours. The incubated cultures were poured into 250 ml centrifuge bottles and spun down in a tabletop centrifuge at 4000 rpm for five minutes at room temperature (Heraeus #7570 G). The supernatant was poured off and then the cell pellet was resuspended in 5 ml of a lysis buffer [50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM NaEDTA]. The resuspended cells were transferred to a 50 ml polypropylene Falcon tube into which 10 ml of NaOH/SDS [0.2 M NaOH, 1% (w/v) SDS] solution was added and shaken rigorously until the mixture became clear (about five shakes). The mixture was incubated on ice for five minutes before 10 ml of a cold Kac/HAc [5 M KAc, 2.5 M HAc] solution was added and shaken five times to precipitate the chromosomal DNA. This mixture was incubated on ice for five minutes and then spun down at 4000 for three minutes at room temperature. The clear supernatant was filtered in a sinctured glass funnel and transferred into a 50 ml round-bottomed polypropylene centrifuge tube without a lip. Isopropanol (12.5 ml) was added to the supernatant and the centrifuge was capped, mixed, and incubated at room temperature for five minutes. After the incubation period, the tube was spun down at 13000 rpm for five minutes at 20˚C (Sorvall RC5C Plus; SS-34 rotor). The supernatant was discarded and the pellet was transferred to a 1.5 ml Eppendorf tube by resuspending in 0.5 ml of 70% ethanol, and spun in a microcentrifuge at 13000 rpm for one minute at room temperature (Sorvall Biofuge). The supernatant was aspirated away and the pellet was spun for an additional 30 seconds, as the residual supernatant was removed. The pellet was resuspended in 0.15 ml (TE (10, 50) [10 mM Tris, 50 mM EDTA] and 1.5 ul of 10 mg/ml RNase (DNase free) was added and incubated in a 37˚C water bath for 15 minutes (vortexing to avoid DNA clumps). Then, the sample was

extracted twice with 0.3 ml phenol/CIA (see section 2.2) and once with 0.5 ml CIA, and the top aqueous layer was added to a homemade spun column. A Sephacryl S400 size exclusion spun column was prepared by stuffing a Gilson blue pipette tip with siliconized glass wool and placing the tip in a 5 ml polypropylene tube with an adaptor (top half of a 1.5 ml Eppendorf tube). The column was filled with Sephacryl S400 resin equilibrated in TE (10, 0.1) and spun in a tabletop centrifuge at 2000 rpm for five minutes at 20˚C (Heraeus #7570 G). The collected at the bottom of the column was discarded before the top aqueous layer was loaded and again spun down at 2000 rpm for five minutes at 20˚C. The eluted plasmid was transferred to a labeled 1.5 ml Eppendorf tube and then analyzed by restriction mapping and, if necessary, sequencing (see section 2.12).

2.12 Restriction Mapping and Sequencing

The two purified preparations of plasmid DNA were restriction mapped to confirm that the desired ligation was achieved, or that site-directed mutagenesis was successful. Two different restriction maps were designed each using a set of two restriction enzymes to cut the plasmid to a predicted size that was distinguishable from the original template construct. Commonly, one enzyme site was chosen that flanked the insert DNA and another site was within the insert. For site-directed mutagenesis, an introduced or removed restriction enzyme site was used to determine whether the enzyme could still cleave the plasmid. The digestion mixes were prepared with 1 µl of 10x NEBuffer, 1 µl of 1 mg/ml BSA, 0.5 µl of 100 mM DTT, 0.5 µl of each appropriate restriction enzyme, and water up to a volume of 9 µl. A single µl of plasmid DNA was added to the digestion mixture and incubated in a 37˚C water bath for 1.5 hours. The samples were mixed with 2 µl of 6x GLB to a final concentration of 1x GLB and analyzed on an appropriate

analytical agarose gel (see section 2.4). Those plasmids that produce the correct size fragments had the desired sequence and, if necessary, were sent for sequencing.

Those plasmid constructs created with PCR methods required sequencing analysis due to chance that the procedure introduced unwanted mutations that would affect the translated protein. The plasmid sample was sent to the Nucleic Acid Facility at the Huck Institutes of the Life Sciences at the Pennsylvania State University for sequencing analysis.

2.13 Subcloning of DNA Inserts

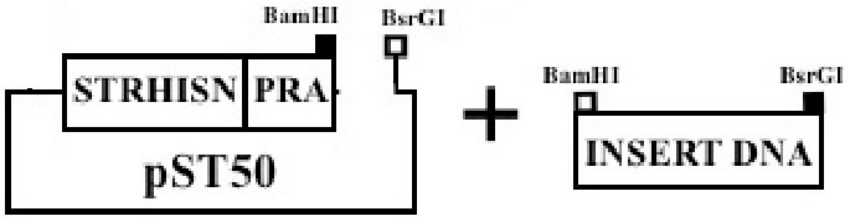


Figure 2. pST50Trc1 Expression Vector with PRA Fusion Tag

Subcloning was completed by ligating the recombinant proteins into cassette 1 on the expression vector adjacent to the N-terminal tags, which included the PRA fusion tag (see section 2.8). The same two restriction enzyme sites were used to ligate each protein to the vector. The insert DNA for each protein was digested and agarose gel purified before ligation (see sections 2.6 and 2.5).

2.14 Small-scale Expression

Before expressing the proteins on a large-scale, the expression levels for each plasmid were tested with small-scale experiments. Since all of the proteins had previously been expressed without the PRA fusion tag, the small-scale expressions were completed only at the temperature producing the highest level in the original protein. The appropriate

plasmid was transformed into BL21(DE3)pLysS competent cells (see section 2.9.) A 500 ml Erlenmeyer flask with 100 ml 2xTY media, 50 µg/ml ampicillin, and 25 µg/ml chloramphenicol was inoculated with three colonies from the transformation plate and placed in a 37˚C shaking incubator at 220 rpm. For expressions at 37˚C, the optical density at 600 nm (OD_{600}), blanked against 2xTY media, was checked continually after about three hours until the OD_{600} was between 0.5 and 0.9 (Doubling time at this temperature was around 20 to 30 minutes.) After the optical density reached that point, the culture was induced by adding 100 µl of 0.2 M Isopropyl β-D-thiogalactopyranoside (IPTG). Prior to induction, a 500 µl sample was transferred to a 1.5 ml Eppendorf tube and spun down in a microcentrifuge for one minute at 13000 rpm (Sorvall Biofuge). The supernatant was aspirated off and the cells were resuspended in 100 µl PGLB [0.5 M Bis-Tris pH 6.8, 20% (v/v) glycerol, 10% (w/v) SDS, 5 M 2-mercaptoethanol, 0.4 mg/ml bromophenol blue]. Samples of 250 µl were also taken, as described above, at time points 1, 2, and 3 hours after induction (and longer if necessary.) Three hours after induction, 50 ml of the culture was harvested by transferring the cells to two separate 50 ml Falcon tubes and centrifuging at 4000 rpm for 10 minutes at room temperature (Heraeus #7570 G). The supernatant was poured off and the each cell pellet was resuspended in 10 ml P300 – EDTA [50 mM sodium phosphate pH 7.0, 300 mM sodium chloride, 1 mM benzamidine, 5 mM 2-mercaptoethanol]. The resuspended cells were flash frozen with liquid nitrogen and stored at -20˚C. For expressions at 28˚C and 18˚C, the expression began the same as that at 37˚C until the OD_{600} of the flask reached between 0.05 and 0.15. At this point, the flask was moved to a shaking incubator at the desired temperature (28˚C or 18˚C) until the OD_{600} reached between 0.5 and 0.9, and the cultures were

induced and samples taken as above. The time points for a 28˚C expression were taken for between 4 and 6 hours (depending on the protein) and the cultures were spun down and resuspended in the same way. For 18˚C, the culture was incubated overnight for about 20 hours before harvesting the cells in the same way. Samples were only collected for the uninduced culture and directly before harvesting. Finally, the samples from the expression were analyzed on an SDS-PAGE gel.

2.15 Small-scale Batch Purification

The solubility of the small-scale expression proteins was tested and the soluble fraction was purified with Talon resin to see whether the method of purification would be successful on a large-scale. Samples for an SDS-PAGE gel were prepared at each step by mixing 25 µl with equal volume of PGLB. First, the frozen 10ml protein sample was thawed in a 30˚C water bath and sonicated twice (10 seconds, 40% of max, 50% cycle) (Branson Digital Sonifier #450) in a 50 ml beaker. The samples were stored on ice for at least 30 seconds between sonication steps and the 1.3 ml aliquots of extract were placed into four Eppendorf tubes and centrifuged for 3 minutes at 13000 rpm (soluble protein in supernatant). At the same time, 1 ml of resuspended Talon Superflow resin (~50% suspension), which binds the N-terminal polyhistadine tag, was transferred to a 15 ml Falcon tube and washed with water by inversion before centrifuging at 1800 rpm for two minutes at room temperature (Heraeus #7570 G). The supernatant was discarded, and then the resin was washed with P300 – EDTA by inversion and centrifuged at 1800 rpm for two minutes at room temperature (Heraeus #7570 G). After that supernatant was discarded, the supernatant of soluble protein from was transferred to the Falcon tube containing the resin and incubated at room temperature for 20 minutes on a rotator. After

the incubation period, the Falcon tube was spun down at 1800 rpm for 5 minutes at room temperature to sediment the resin and the supernatant was transferred to a separate Falcon tube as the flow through. The resin was then washed twice with P300 – EDTA by adding 10 ml of the buffer to the Falcon tube, inverting the tube several times, and spinning the tube at 1800 rpm for 5 minutes at room temperature (discard supernatant). The final wash with 3 ml of P300 – EDTA is transferred to and allowed to flow through a disposable 1ml Biorad BioSpin column (Catalog #727-6008) clamped to a retort stand. The elution buffer was prepared with P300 – EDTA + 100 mM imidazole and four separate 0.5 ml fractions were collected in 1.5 ml Eppendorf tubes as the sample eluted from the BioSpin column, frozen in liquid nitrogen and stored at -20˚C. The samples including the whole cell extract, pellet, supernatant, flow through, and the four eluted fractions were analyzed for solubility and successful purification on an 18% SDS-PAGE gel.

2.16 Large-scale Expression

The large-scale expression of the plasmids began similar to that of the small-scale, as the plasmid was transformed into the competent BL21(DE3)pLysS cells and incubated at 37°C. Then, a 500 ml flask containing 100 ml of 2xTY + 50 μg/ml ampicillin + 25 μg/ml chloramphenicol was inoculated with three of the colonies from the transformation plate and incubated overnight at about 20.5°C for about 16 hours. In the morning, the OD_{600} was between 0.1 and 1.0, and three flasks of 500 ml 2xTY + 0.5 ml of 50 mg/ml ampicillin + 0.5 ml of 25 mg/ml chloramphenicol (1.5 L of media total) were inoculated with three ml each from the original flask. For large-scale expressions at 37°C, the three large flasks were placed in a shaking incubator at that temperature and 220 rpm until the OD_{600} reached between 0.6 and 1.0. At this point, adding 0.5 ml of 0.2 M IPTG to each flask

induced the cultures and the flasks were incubated for the appropriate time determined from the small-scale expression. For large-scale expressions at 28°C and 18°C, the three large flasks were placed in the 37°C shaking incubator until the OD_{600} was between 0.05 and 0.15, when the temperature was lowered to the desired level. Just as with the 37°C expressions, the cultures were induced at an OD_{600} between 0.6 and 1.0, and incubated for the appropriate time. A 250 µl sample of pre-induced culture was spun down for one minute at 13000 rpm and resuspended in 50 µl PGLB. Samples of 125 µl were taken in the same way from the three flasks before the harvesting the cells. These were analyzed on an agarose gel to check the expression of the protein. After the appropriate incubation time, the cells were harvested by centrifuging in 500 ml bottles at 7000 rpm for 5 minutes at 20°C. The three large flasks of cells were resuspended together in 37.5 ml of P300 – EDTA, frozen in liquid nitrogen, and stored at -20°C.

2.17 Talon Column Affinity Purification

Before loading the column, the crude cellular extract was prepared for purification. The harvested cells were thawed in a 30°C water bath and sonicated four times (10 seconds, 70% of max, 50% cycle) (Branson Digital Sonifier #450) in a 100 ml beaker. The samples were placed on ice for 30 seconds in between sonication cycles. The sonicated extract was centrifuged at 18000 rpm for 20 minutes at 4°C (Sorval RC5C Plus; SS-34 rotor) and the supernatant was collected for purification. Samples of the whole cell extract, pellet, and supernatant were prepared for SDS-PAGE by mixing 25 µl of sample with 25 µl of PGLB.

The night before purification, the Talon resin (20 ml) was re-packed in the column (only necessary for every three purifications). While preparing the crude extract, the Talon

column was equilibrated with the appropriate buffers, which were diverted to waste (BioRad BioLogic LP pump). The buffers used for the purification were: P300 – EDTA (Buffer A), P300 – EDTA + 100 mM imidazole (Buffer B), and water (Buffer C). Prior to running the equilibration program, lines A and E were purged for about one minute each.

Equilibration program:	**Buffer**	**Flow Rate**
	20 ml Buffer C through line D	2 ml/min
	20 ml Buffer A through line A	3 ml/min
	20 ml Buffer B through line E	3 ml/min
	50 ml Buffer A through line A	3 ml/min

After the column was equilibrated, lines C and D were purged for one minute each and then the sample (supernatant from above) was loaded manually through line A with a flow rate of 3 ml/min. The waste was diverted to a beaker, which was saved as the flow through sample. After the entire sample was loaded, the column was washed and the sample was eluted with the appropriate buffers. Buffers with 10% and 20% buffer B were used to elute proteins with non-specific binding to the column. The sample was collected as 6 ml fractions in the fraction collector (Gilson FC204).

Wash Program:	**Buffer**	**Flow Rate**
	60 ml Buffer A through line A	3 ml/min
	40 ml 10% Buffer B through line C	3 ml/min
	60 ml 20% Buffer B through line D	3 ml/min
Elution Program:	80 ml Buffer B through line E	3 ml/min

After the program was executed, a cleaning program was run with the appropriate buffers

to remove any material remaining on the column. The important buffer for cleaning the column was 6M Guanidine-HCl (Buffer E) and water (Buffer D).

Cleaning Program:	**Buffer**	**Flow Rate**
	20 ml Buffer D through line D	3 ml/min
	10 ml Buffer E through line E	1 ml/min
	20 ml Buffer D through line D	1 ml/min
	30 ml Buffer D through line D	3 ml/min
	50 ml Buffer D through line D	2 ml/min

The elution samples collected by the fraction collector and the flow through sample were prepared for SDS-PAGE by mixing 25 µl of sample with 25 µl of PGLB. These samples were run on an SDS-PAGE gel alongside the whole cell extract, pellet, and supernatant prepared during the cell extract preparation. The fractions with significant protein material were pooled and dialyzed overnight against 2 L of the appropriate buffer with 50 mM NaCl (based on the associated pI). After dialysis, the material was recovered in a falcon tube, flash frozen with liquid nitrogen, and stored in 20% glycerol (by volume) at -20°C. If the sample contained a precipitate after dialysis, the sample was spun in a 50 ml centrifuge bottle at 15000 rpm for 10 minutes at 4°C (Sorval RC5C Plus, SLA-3000 rotor).

2.18 Marker Preparation

Marker preparation required relatively equal concentrations for each of the various proteins. Serial dilutions of the different proteins were run on a western blot to determine their relative concentrations and to determine the dilution that produced a fine band

without noticeable contamination or proteolysis products. The samples were diluted with a mixture containing an equal volume of PGLB and water before being analyzed by western blot. After the relative concentrations were determined, one ml aliquots of the marker were prepared with the appropriate dilution.

2.19 Western Blotting

First, the samples were run on an SDS-PAGE gel. While the gel was running nitrocellulose membrane was cut to size and equilibrated in transfer buffer with 20% methanol [25 mM Tris, 192 mM Glycine, 20% methanol by volume, water] for 2 to 5 minutes. After the gel had run to completion, the gel was also equilibrated in the transfer buffer and placed in the transfer assembly with the nitrocellulose membrane. The nitrocellulose was blotted in the transfer assembly at 100 V and about 350 mA for 60 minutes. The blotted nitrocellulose membrane was equilibrated in TBS buffer [25mM Tris, 150mM NaCl, HCl (pH 8.0), water] for 2 to 5 minutes. Then, the membrane was incubated with 50 ml of pre-incubation buffer [2% non-fat milk, TBS buffer] for 30 minutes and washed four times with 50 ml of TTBS buffer [TBS buffer, 0.05% Tween 20] for 5 minutes each. The membrane was incubated for an hour in 10 ml of TTBS buffer containing a 1:10,000 dilution of the secondary antibody, Donkey HRP-linked anti-rabbit whole antibody (Pharmacia). Again, the membrane was washed four times in TTBS buffer. Finally, mixing 1 ml of each ECL detection solution (horse radish peroxides) and exposing the membrane with film detected the chemiluminescent reaction on the membrane.

3. Results

3.1 Amplification and Cloning of IgG Binding Domain

The plasmid containing the TAP tag, coding for Calmodulin Binding Peptide and Protein A, was obtained from the lab of Dr. Joseph Reese (Reese, 2009). This particular protein domain is used in the Tandem Affinity Purification technique (Rigaut, 1999), and contained two symmetrical IgG Binding Domain units (referred to as the PRA fusion tag: see figure 3). The PRA fusion tag alone can recognize the F_c region of secondary antibodies. Therefore, a cloning scheme was designed to incorporate the PRA fusion tag alongside recombinant DNA of proteins of varying size in an expression vector (see figure 4).

KTAALAQHDEA
VDNKFNKEQQNAFYEILHLPNLNEEQRNA
FIQSLKDDPSQSANLLAEAKKLNDAQAPK
VDNKFNKEQQNAFYEILHLPNLNEEQRNA
FIQSLKDDPSQSANLLAEAKKLNGAQAPK
VDANSAGKST

Figure 3. Amino Acid Sequence of PRA fusion Tag (Rigaut, 1999)
The amino acid sequence contains two virtually symmetrical sequences highlighted in blue and red.

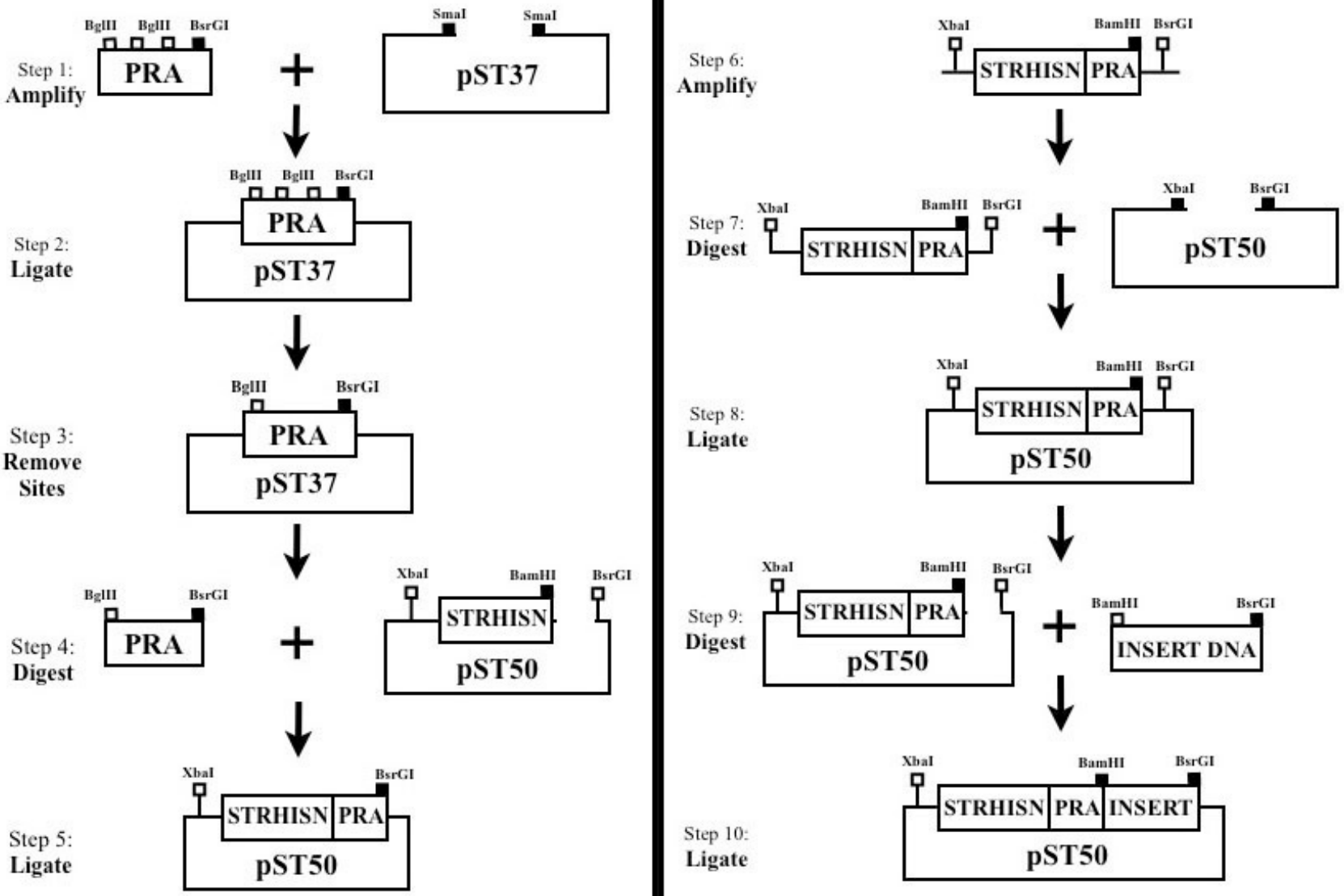


Figure 4. Cloning Method Diagram

All of the digests contain sticky-ends, except for the first vector cut with SmaI.

The PRA fusion tag was successfully amplified with a BglII site at the 5' end and a BsrGI site at the 3' end for insertion within the first cassette of the pST50Tr vector (see figure 5). While the vector contained a BamHI restriction site, the BglII site contained the appropriate “sticky-ends” to ligate with a BamHI site. The PCR fragment was kinased and ligated into a “blunt-end” pST37 vector. Restriction mapping and sequencing were used to confirm the amplification of the PRA fusion tag and the ensuing ligation.

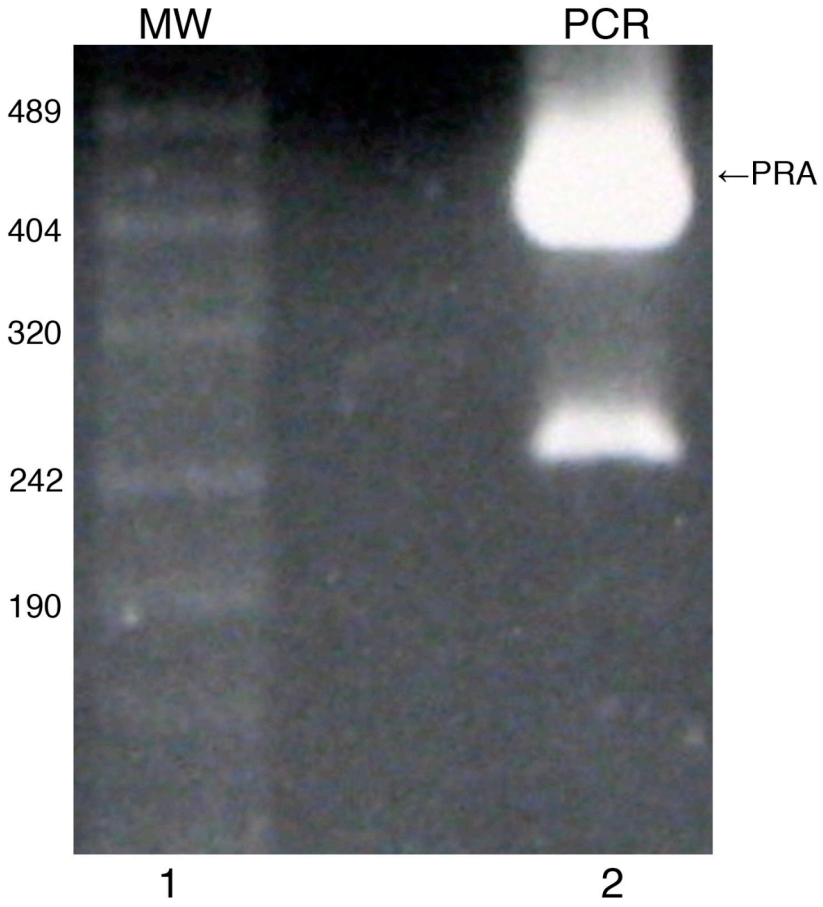


Figure 5. Amplified PRA fragment at 433 base pairs

While in the pST37 vector, two BglII sites internal to the PRA fusion tag were removed by PCR site-directed mutagenesis. The PCR reaction amplified the plasmid by introducing a C to T base change on a wobble base of the two BglII recognition sites to

create pST37-PRAx1. While the restriction enzyme could no longer recognize the site, the translated protein sequence was conserved (see figure 4). The new sequence was confirmed by restriction mapping and sequencing.

Without the internal BglII sites, the PRA fusion tag could be isolated by digesting the pST37-PRAx1 plasmid with BglII and BsrGI restriction endonucleases. Additionally, the expression vector pST50Tr-STRHISNyEpl1Δ3x3 was digested at the BamHI and BsrGI restriction sites, which cleaved open the first cassette releasing the unwanted Epl1Δ3x3 insert from the vector. The sequences upstream of the insert were the hexahistidine affinity tag (HIS) used for purification with Talon resin (see sections 2.16 and 2.18), the Strep II affinity tag (STR) provided for optional purification with Strep-Tactin and the TEV NIa protease site (N). In an alternate application of the western blot, the affinity tags could be the antigenic targets of a primary antibody. The PRA fragment was isolated from the residual pST37 vector and the pST50Tr-STRHISN vector was isolated from the Epl1Δ3x3 DNA fragment by gel purification, while the leftover fragments were discarded. The PRA fusion tag was ligated into the first cassette of the expression vector adjacent to the N-terminal affinity tags to create pST50Tr-STRHISNPRA, and was confirmed by restriction mapping. In ligating the compatible "sticky-ends" of BamHI and BglII, the new sequence no longer contained an enzyme recognition site. This sequence was also used to express the protein of the lowest molecular weight (see figure 4).

In order to re-establish the restriction site to subclone recombinant DNA adjacent to the N-terminal tags and the PRA fusion tag in the vector, PCR linear amplification was used to amplify part of the plasmid sequence, while introducing a BamHI restriction site

before the STOP codon and BsrGI restriction site. The forward primer was designed to anneal before the XbaI site of cassette 1 and the reverse primer was designed to introduce the BamHI restriction site (see figure 4). The PCR product was digested at the XbaI site and BsrGI site and the fragment was gel purified from the residual DNA at the ends. The gel purified fragment was ligated into a gel purified pST50Tr vector (previously isolated by Dr. Song Tan) and was also digested at the XbaI and BsrGI site to create pST50Tr-STRHISNPRAi1 (PRAi1 denotes the fusion tag followed by the inserted BamHI site), and then confirmed by restriction mapping and sequencing. This new plasmid contained BamHI and BsrGI restriction sites for subcloning insert DNA adjacent to the PRA fusion tag (see figure 4).

3.2 Subcloning of DNA Inserts

Once the BamHI restriction site was re-introduced in the vector, the pST50Tr-STRHISNPRAi1 plasmid was ready to receive the insert DNA for the various proteins selected. First, the construct was digested with BamHI and BsrGI and gel purified to set up the vector for ligation.

Protein	**DNA Insert Size (bp)**	**Protein Size (kDa)**
STRHISNPRA*	430	19.0
STRHISNPRAyFhl1Δ1	694	29.0
STRHISNPRADHFR	916	37.2
AVTGSTHISNPRA	1151	46.5
STRHISNPRAyReb1Δ1	1675	67.0
STRHISNPRAdRCC1	2074	77.9
STRHISNPRAyBas1+	2815	107.0

Table 1. Characteristics of Proteins Selected for Western Blot Marker
** Created in section 3.1 (not from this subcloning step).*
+DNA construct created, but not expressed on a large-scale for this thesis due to time constraints.

Six constructs were created by subcloning directly into the pST50Tr-STRHISNPRAi1 gel-purified vector. Several of the gel-purified insert DNA fragments (digested with BamHI and BsrGI) were previously prepared for use in other experiments and did not require further manipulation. These included the gel purified DNA for DHFR (created by Dr. Song Tan), dRCC1 (created by Dr. Ravindra Makde), and Reb1Δ1 (created by Dr. Sung-Hoon Jun). The insert DNA for yFhl1Δ1 and yBas1 was isolated from other laboratory constructs by digesting with BamHI and BsrGI restriction enzymes and then gel purifying on an agarose gel. The gel purified insert DNA for each of the six constructs was ligated with the pST50Tr-STRHISNPRAi1 gel-purified vector. All of the newly created plasmids were confirmed by restriction mapping.

The construct for pST50Tr-AVTGSTHISNPRA was created using the original gel purified DNA PRAx1 digested with BglII and BsrGI restriction enzymes (see section 3.1). The gel-purified vector, pST50Tr-AVTGSTHISN digested with BamHI and BsrGI, was previously isolated by Dr. Sun-Hoon Jun. The vector and insert DNA fragments were ligated and also confirmed by restriction mapping. The N-terminal affinity tags of AVTGSTHISNPRA were AviTag (AVT) and glutathione S-transferase (GST), which bind biotin and glutathione respectively.

3.3 Small-scale Expression and Batch Purification

All of the proteins were previously expressed without the PRA fusion tag for use in other experiments in the laboratory (work completed by Dr. Song Tan, Dr. Sung-Hoon Jun, Dr. Ravindra Makde, and Slater Nair). Therefore, the small-scale expression and batch purification were first completed at the optimal temperature from these original attempts.

Protein	Temperature	Expression Level	Solubility	Purification
STRHISNPRA	37°C	High	Soluble	Successful
STRHISNPRAyFhl1Δ1	28°C	Medium	Partially Soluble	Successful
STRHISNPRADHFR	37°C	High	Soluble	Successful
AVTGSTHISNPRA	37°C	High	Mostly Soluble	Successful
STRHISNPRAyReb1Δ1	37°C	Medium	Partially Soluble	Successful
STRHISNPRAdRCC1	28°C	High	Mostly Soluble	Successful
STRHISNPRAyBas1	18°C	Low	Mostly Soluble	Successful

Table 2. Small-scale Expression Level and Solubility
Lower solubility of some proteins may be due to incomplete lysis of whole cell extract.

The four proteins expressed on a small-scale at 37°C were incubated for three hours after induction. The proteins were purified on the Talon resin and the samples were analyzed by SDS-PAGE electrophoresis. The conditions for all four proteins appeared suitable to continue on to expression on a large-scale; the protein that appeared only partially soluble, STRHISNPRAyReb1Δ1, was most likely due to inefficient lysis.

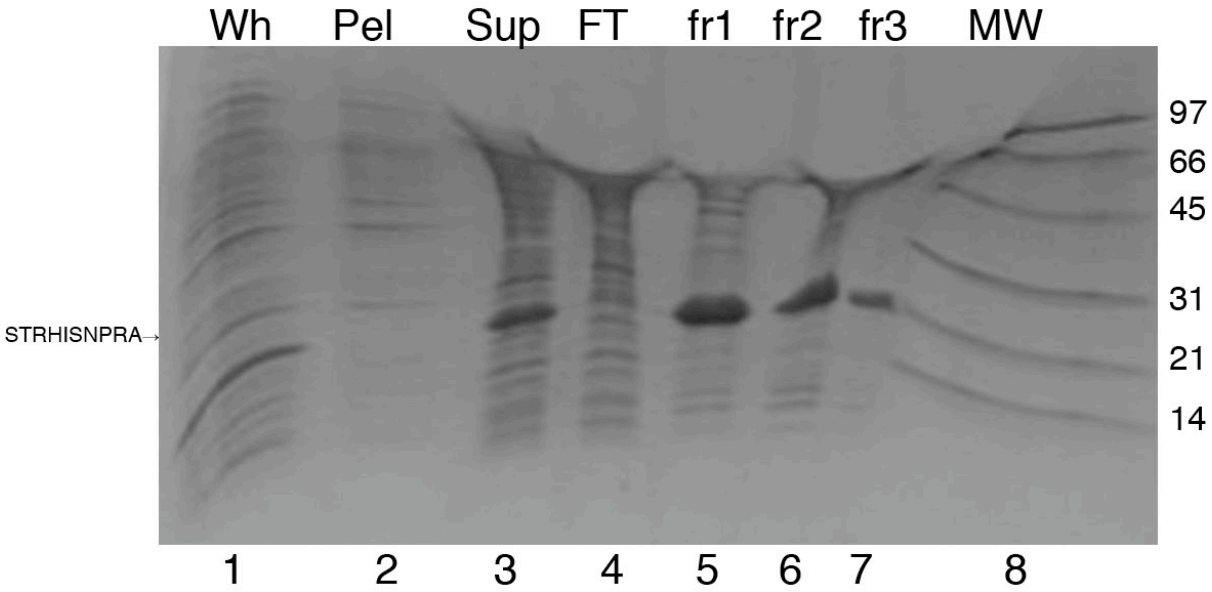


Figure 6. STRHISNPRA Small-Scale Affinity Purification for Cells Grown at 37°C
Mostly the entire crude extract appeared in the supernatant, which suggests the protein was soluble. Also, the elution fractions 1 through 3 showed that the protein was purified on the resin.

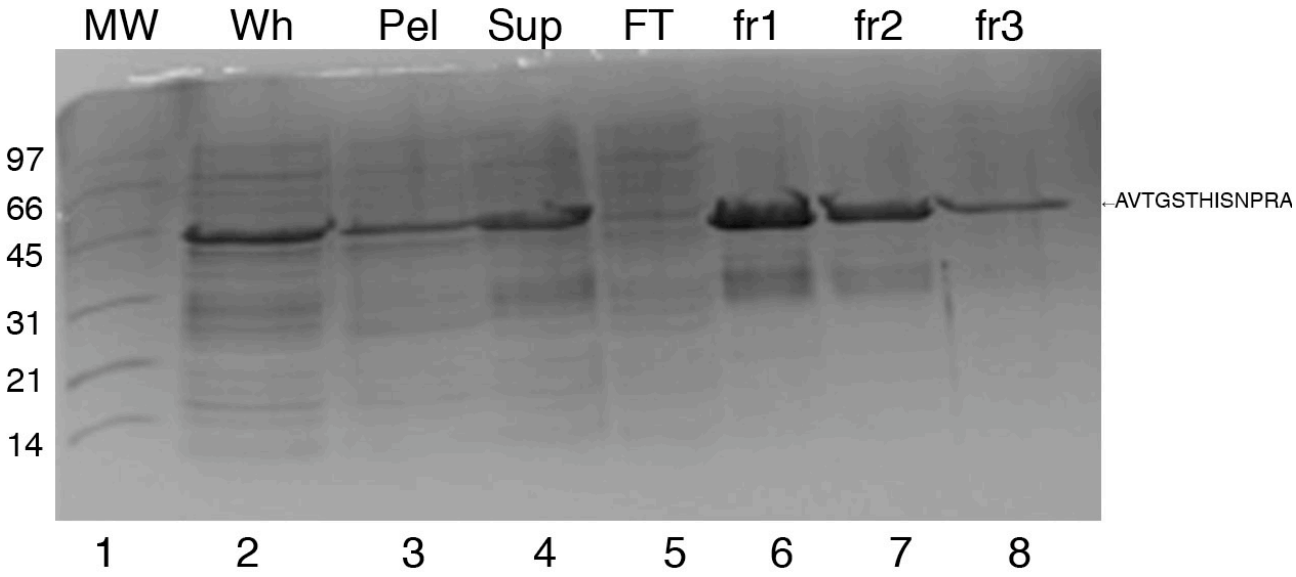


Figure 7. AVTGSTHISNPRA Small-Scale Affinity Purification for Cells Grown at 37°C
A small amount of the crude extract appeared in the pellet, which suggests the protein was only mostly soluble. Also, the elution fractions 1 through 3 showed that the protein was purified on the resin.

The proteins STRHISNPRAdRCC1 and STRHISNPRAyFhl1Δ1 were expressed on a small-scale at 28°C for four hours and six hours after induction respectively. These time points were determined based upon the work done by Dr. Ravindre Makde and Dr. Sung-Hoon Jun. The conditions for both of these proteins were also suitable to continue on to a large-scale expression, as the partial solubility of STRHISNPRAyFhl1Δ1 was most likely due to inefficient lysis. While the expression of STRHISNPRAyBas1 for an incubation period of 20 hours at 18°C produced soluble protein, the expression level was low.

3.4 Large-scale Expression and Purification

All six proteins were expressed on a large-scale in 1.5L of 2xTY media at the optimal conditions determined in the small-scale expression and batch purification. Each protein was purified on the column with 20 ml of Talon resin at 4°C, dialysed overnight in the appropriate buffer with 50mM NaCl, and stored in 20% Glycerol at -20°C (see table 3).

Protein	Buffer	Recovery	Fractions	Volume
STRHISNPRA	5mM Hepes pH 7.5	~ 30%	30-36	50 ml
STRHISNPRAyFhl1Δ1	5mM Hepes pH 7.5	~50%	26-32	50 ml
STRHISNPRADHFR	20mM Tris-Cl pH 8	~40%	27-33	50 ml
AVTGSTHISNPRA	20mM Tris-Cl pH 8	~85%	26-36	60 ml
STRHISNPRAyReb1Δ1	20mM Tris-Cl pH 8	~20%	28-32	35 ml
STRHISNPRAdRCC1	5mM Hepes pH 7.5	~75%	26-32	50 ml

Table 3. Talon Column Purification and Dialysis Results

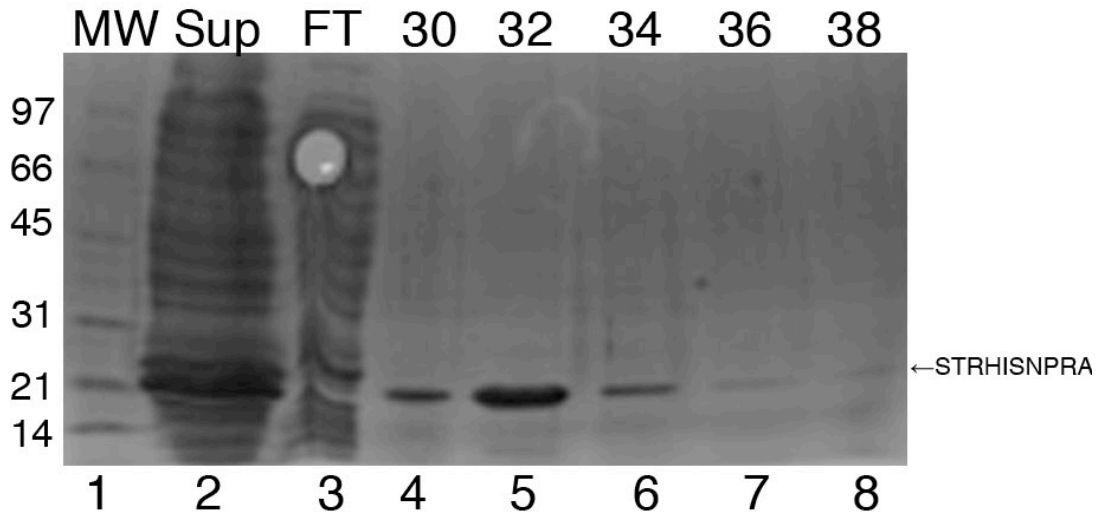


Figure 8. STRHISNPRA Talon Purification Pool with fractions 30-38
**Whole and Pellet of crude extract not shown*

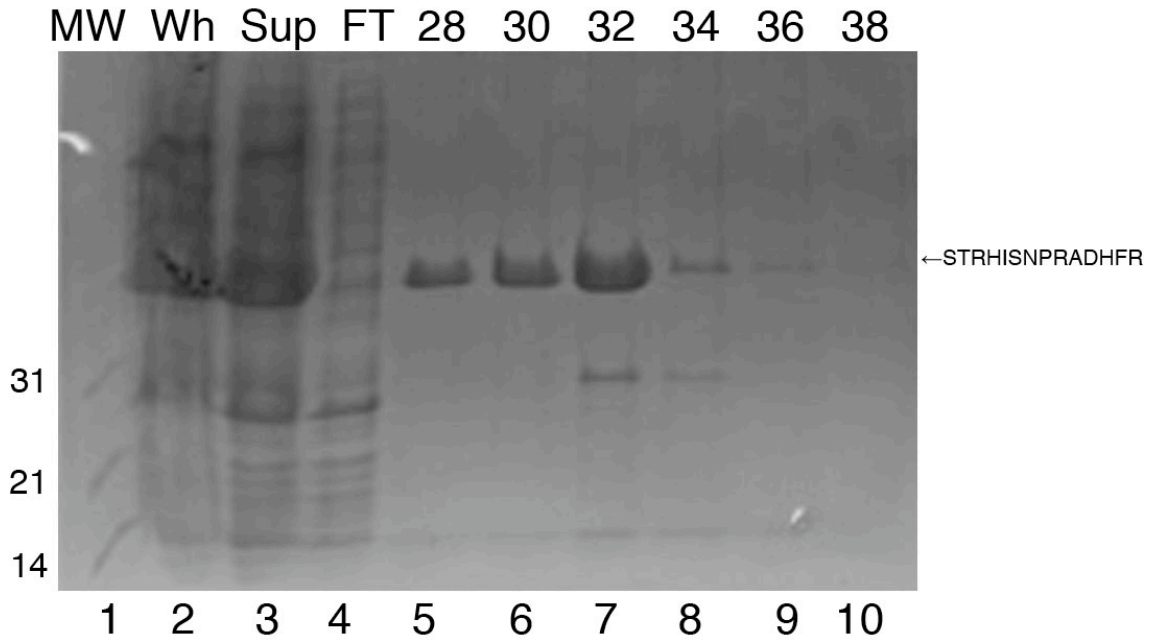
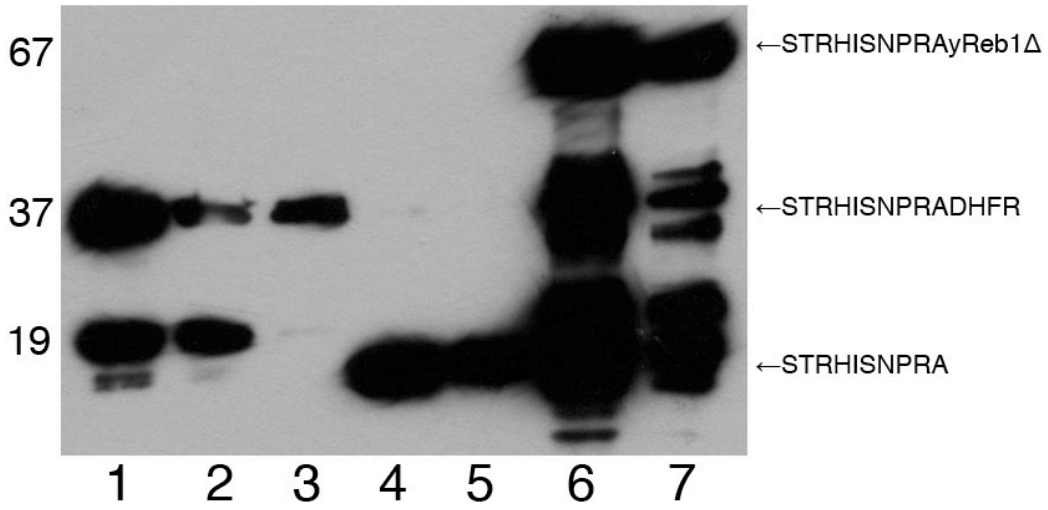


Figure 9. STRHISNPRADHFR Talon Purification Pool with fractions 28-38
**Pellet of crude extract not shown*

After purification, five of the six proteins showed adequate expression and a sufficient yield. However, STRHISNPRAyReb1Δ1 did not express well and appeared to be of too low a concentration. STRHISNPRAyReb1Δ1 also appeared to have several contaminants or proteolytic products that were of a considerable concentration. Additionally, STRHISNPRADHFR appeared to have a significant contaminant at 25kDa (as seen in figure 9).

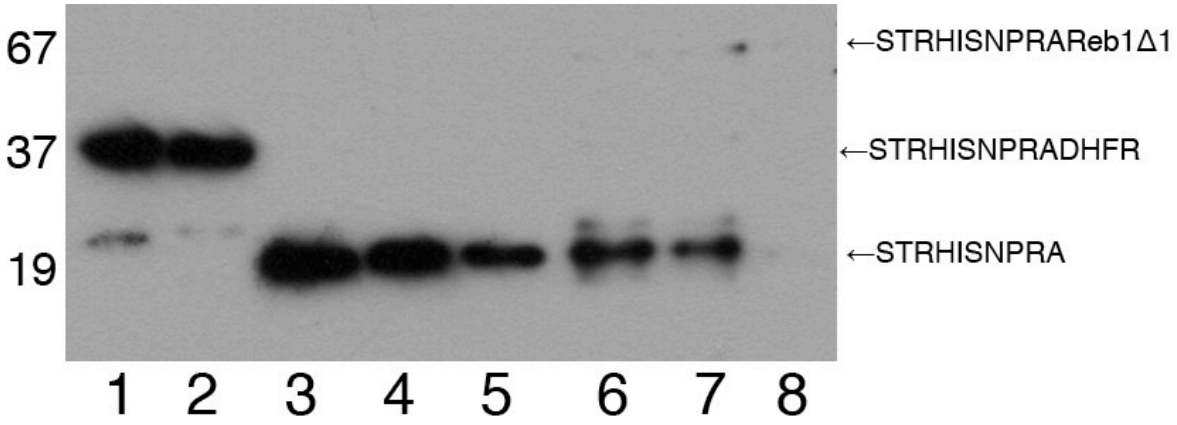
3.5 Marker Preparation

After purification on the Talon column, the proteins were diluted to determine their relative concentrations, as they would appear on the western blot after exposing for 30 seconds. The proteins were diluted with a mixture containing equal volumes of PGLB and water. The first set of three proteins (marker set 1), STRHISNPRA, STRHISNPRADHFR, and STRHISNPRAyReb1Δ1, were analyzed at specified dilutions (see figure 10). While STRHISNPRA and STRHISNyReb1Δ1 appeared saturated at each of the associated dilutions in trial 1, STRHISNPRADHFR appeared as a medium sized band without its contaminant for the 1:100 dilution. Then, the same three proteins were analyzed after further diluting the samples (see figure 11). For the second trial, STRHISNPRA and STRHISNPRADHFR appeared to have about the same relative concentration for the 1:100 and 1:200 dilutions (see figure 11). Also, the contaminant in the STRHISNPRADHFR sample was even less apparent for the 1:200 dilution. For both dilution trials, STRHISNyReb1Δ1 appeared to contain many contaminants.



Lane	Protein	Dilution
1	STRHISNPRADHFR	1:20
2	STRHISNPRADHFR	1:40
3	STRHISNPRADHFR	1:100
4	STRHISNPRA	1:20
5	STRHISNPRA	1:40
6	STRHISNPRAyReb1Δ1	1:1
7	STRHISNPRAyReb1Δ1	1:2

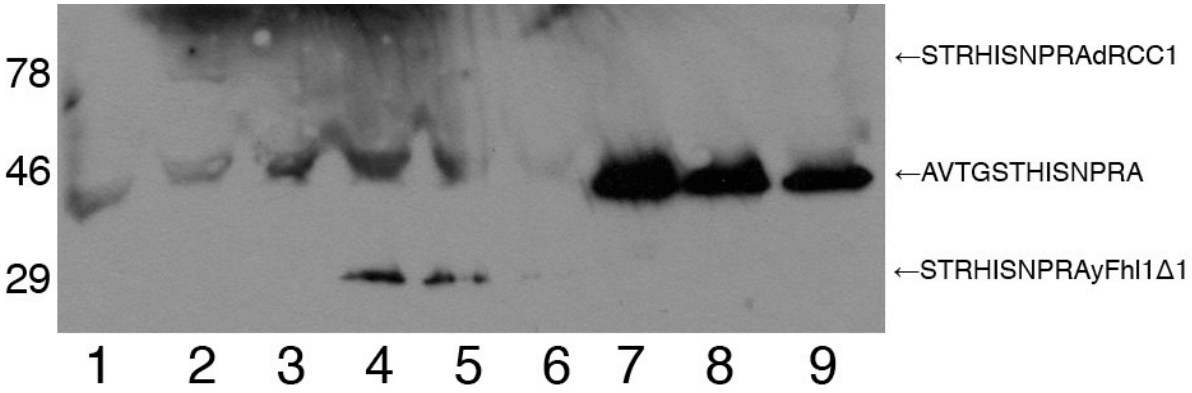
Figure 10. Western Blot Trial 1 Serial Dilutions for Marker Set 1



Lane	Protein	Dilution
1	STRHISNPRADHFR	1:100
2	STRHISNPRADHFR	1:200
3	STRHISNPRA	1:100
4	STRHISNPRA	1:200
5	STRHISNPRA	1:500
6	STRHISNPRAyReb1Δ1	1:20
7	STRHISNPRAyReb1Δ1	1:40
8	STRHISNPRAyReb1Δ1	1:100

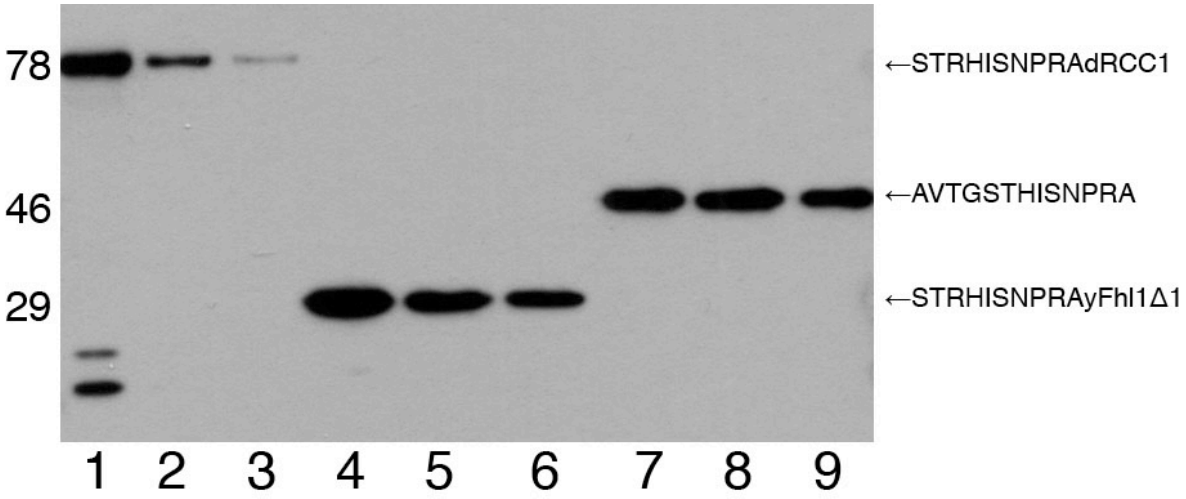
Figure 11. Western Blot Trial 2 Serial Dilutions for Marker Set 1

The second set of three proteins (marker set 2), STRHISNPRAdRCC1, STRHISNPRAyFhl1Δ1, and AVTGSTHISNPRA were analyzed at specified dilutions (see figure 12). Little or no STRHISNPRAdRCC1 or STRHISNPRAyFhl1Δ1 material appeared in the dilutions for trial 1, whereas there was a large amount of AVTGSTHISNPRA material in all the dilutions for trial 1. For trial 2, the three proteins, STRHISNPRAdRCC1, STRHISNPRAyFhl1Δ1, and AVTGSTHISNPRA appeared to be relatively equal in concentration at the 1:20, 1:50, and 1:1500 dilutions respectively.



Lane	Protein	Dilution
1	STRHISNPRAdRCC1	1:100
2	STRHISNPRAdRCC1	1:200
3	STRHISNPRAdRCC1	1:500
4	STRHISNPRAyFhl1Δ1	1:100
5	STRHISNPRAyFhl1Δ1	1:200
6	STRHISNPRAyFhl1Δ1	1:500
7	AVTGSTHISNPRA	1:200
8	AVTGSTHISNPRA	1:500
9	AVTGSTHISNPRA	1:1000

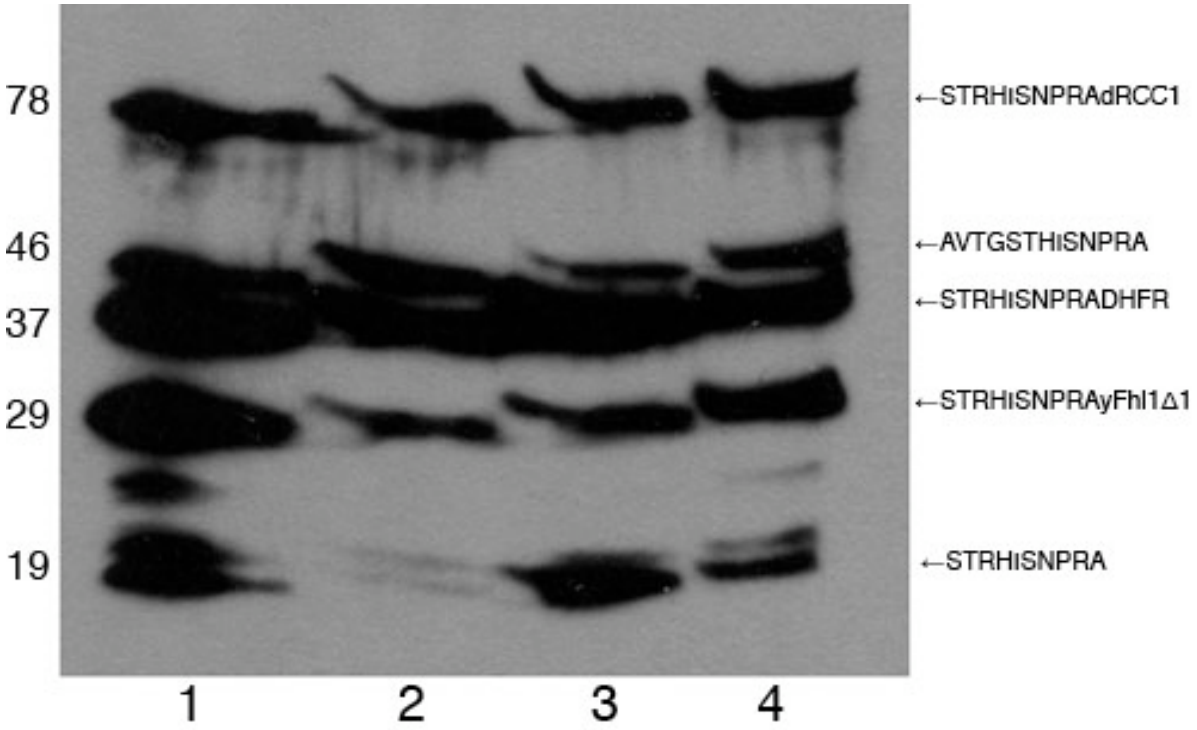
Figure 12. Western Blot Trial 1 Serial Dilutions for Marker Set 2



Lane	Protein	Dilution
1	STRHISNPRAdRCC1	1:20
2	STRHISNPRAdRCC1	1:50
3	STRHISNPRAdRCC1	1:100
4	STRHISNPRAyFhl1Δ1	1:20
5	STRHISNPRAyFhl1Δ1	1:50
6	STRHISNPRAyFhl1Δ1	1:100
7	AVTGSTHISNPRA	1:1000
8	AVTGSTHISNPRA	1:1500
9	AVTGSTHISNPRA	1:2000

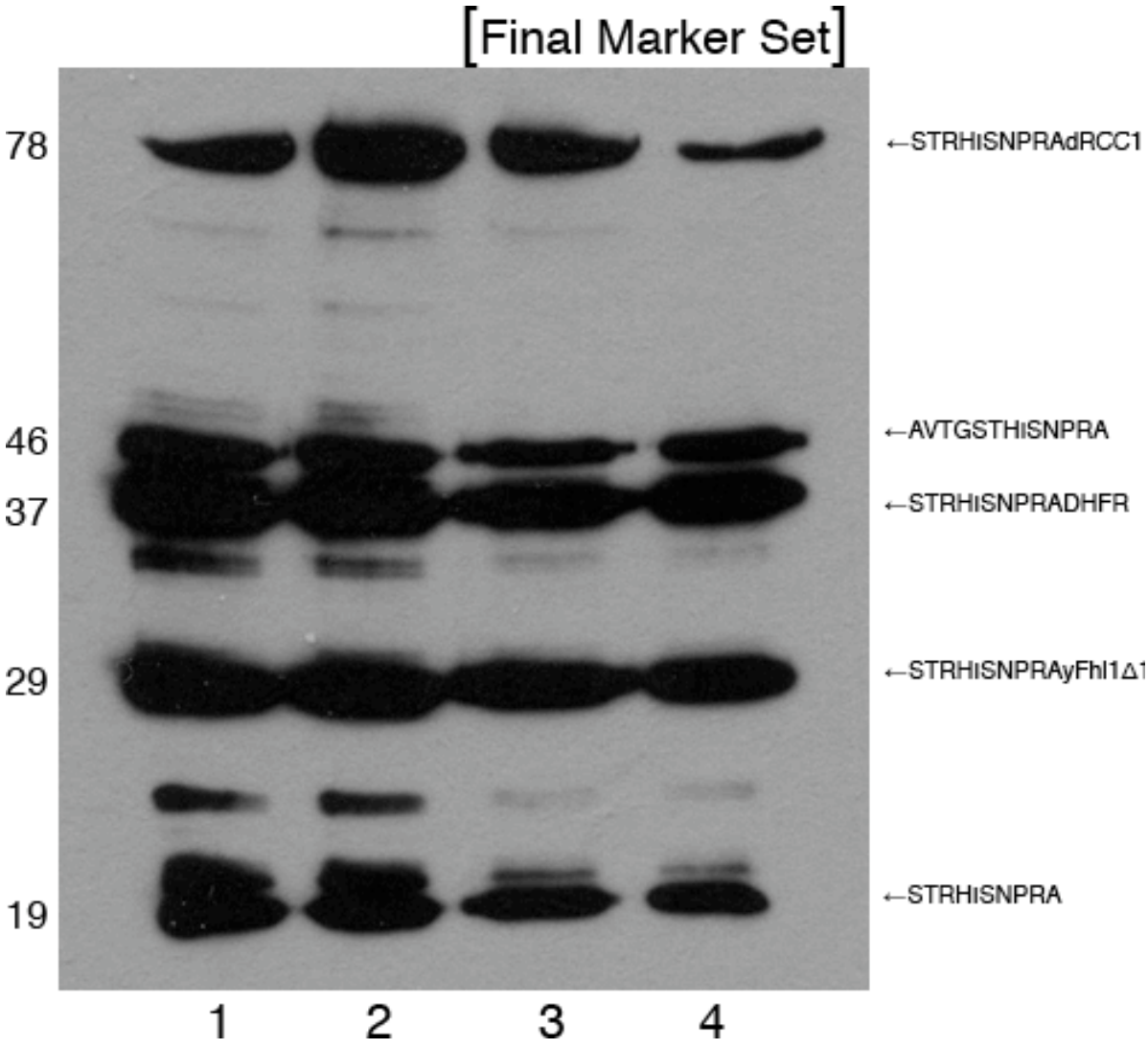
Figure 13. Western Blot Trial 2 Serial Dilutions for Marker Set 2

The proteins from set 1 and set 2 were incorporated into a single marker lane for the final two trials (see figure 14 and 15). For the first trial, four lanes were prepared with varying dilutions. The more dilute samples in lane 4, seen in figure 14, appeared to produce the best molecular weight marker, but STRHISNPRA (~19 kDa) and AVTGSTHISNPRA (~47 kDa) were of a slightly lower concentration than the other three proteins. In the final trial, seen in figure 15, each lane contained the optimal concentration for each sample, but lanes 3 and 4 were diluted by an additional 50% by volume. While each protein appeared relatively equal in concentration, the contaminant at 25 kDa was less apparent in the more dilute sample.



Protein	Lane 1	Lane 2	Lane 3	Lane 4
STRHISNPRADHFR	1:200	1:500	1:200	1:500
STRHISNPRA	1:200	1:500	1:200	1:500
STRHISNPRAdRCC1	1:20	1:20	1:30	1:30
STRHISNPRAyFhl1Δ1	1:50	1:50	1:100	1:100
AVTGSTHISNPRA	1:1500	1:1500	1:2000	1:2000

Figure 14. Combined Western Blot Trial 1



Protein	Lane 1	Lane 2	Lane 3	Lane 4
STRHISNPRADHFR	1:500	1:500	1:750	1:750
STRHISNPRA	1:200	1:200	1:300	1:300
STRHISNPRAdRCC1	1:30	1:30	1:45	1:45
STRHISNPRAyFhl1Δ1	1:100	1:100	1:150	1:150
AVTGSTHISNPRA	1:1500	1:1500	1:2250	1:2250

Figure 15. Combined Western Blot Trial 2

4. Discussion

4.1 Summary

My goal for this thesis project was to develop a set of molecular weight markers that would be helpful in analyzing western blots. The IgG binding capability of Protein A was employed to create a fusion tag (PRA fusion tag) that could be universally detected by the same antibody-conjugate reagents used to visualize the target protein. This required expressing and purifying protein standards with the PRA fusion tag. The protein standards were created in a wide range of sizes to maximize the utility of the molecular weight marker.

There were three major steps required to produce the western blots tagged with the IgG Binding Domain from Protein A (PRA fusion tag). First, the PRA fusion tag was isolated from a plasmid containing the TAP tag and subcloned as a fusion tag in an expression vector. This required amplification of the sequence, purification of the insert with the appropriate restriction enzymes, and ligation of the insert DNA into the expression vector. Challenges to constructing these plasmids included confounding internal restriction sites as well as technical difficulties with ineffective partial digests and "blunt-end" cloning. The second part of the project was to select available genes to subclone into the vector alongside the PRA fusion tag to produce proteins of varying molecular weight. This step required isolating and gel purifying the DNA for each protein and ligating the insert into the prepared expression vector. Once all of the DNA constructs were created, the tagged proteins were expressed and purified for use in the western blot marker. Each protein was first expressed and purified on a small scale to determine the optimal conditions. Then, the proteins were expressed on a large scale at their optimal

temperature and purified by affinity chromatography. The final step was to prepare serial dilutions for each purified protein to determine the suitable concentration in the molecular weight marker mixture.

4.2 Analysis of Isolation and Cloning of IgG Binding Domain

The original design for isolating and cloning the IgG Binding Domain of Protein A (PRA fusion tag) was simple and straightforward. The coding region for the tag was to be amplified using PCR primers that incorporated appropriate restriction sites. In particular, I wanted to subclone the PRA fusion tag as a BglII-BamHI restriction fragment into a BamHI digested vector fragment (BglII and BamHI restriction sites share compatible sticky ends). The resulting plasmid with the insert ligated in the correct orientation would act as the parent plasmid (plasmid 1) into which other coding regions could be subcloned, thus creating the set of expression vectors for different proteins. In addition, I wanted to subclone the PRA fusion tag as a BglII-BsrGI restriction fragment to create an expression plasmid (plasmid 2) that would express only tagged PRA (to provide the smallest molecular weight marker protein).

After receiving the plasmid containing the Protein A sequence and analyzing the Protein A coding sequence, I realized isolating the PRA coding fragment would be complicated by two internal BglII sites, one in each of the two tandem IgG binding domain repeats. I initially attempted to use partial BglII digestion to overcome the problem of the internal BglII sites. However, the desired fragments were too similar in size to the undigested PCR products and I was not able to resolve and to isolate the desired partial digestion products. Therefore, I decided to use “blunt-end” ligation of the undigested PCR products

to create an intermediate plasmid (pST37-PRA). Doing so allowed me to resolve the desired fragment of 430 base pairs from the larger 2833 base pair vector fragment. This approach allowed me to create plasmid 2, but not plasmid 1.

Instead of pursuing the partial digestion approach to handle the internal BglII sites, I decided to use site-directed mutagenesis to remove these restriction sites. I designed mutagenesis primers and then used these primers and the QuikChange PCR-based mutagenesis procedure to remove the internal BglII sites from the PRA fusion tag. The PRA coding region was then amplified by PCR, digested with the appropriate restriction enzymes (BglII and BsrGI) and inserted into an expression vector to create the desired pST50Tr-STRHISNPRA plasmid.

This pST50Tr-STRHISNPRA plasmid construct did not contain the requisite 5' BamHI site needed for subsequent subcloning steps. To remedy this, PCR primers were designed to insert the BamHI site before the BsrGI site. This resulting PCR product was used to create pST50Tr-STRHISNPRAi1, where i1 indicates insert construct 1. This plasmid provided 5' BamHI and 3' BsrGI restriction sites for subcloning inserts downstream of the PRA fusion tag. Thus, the desired DNA construct was created, although not in the direct manner as originally planned.

After creating the expression vector required for subcloning, insert DNA fragments were ligated into the vector to create PRA tagged expression plasmids for each marker protein. Insert DNA fragments ranging in size from 700 base pairs to as many as 2800 base pairs were successfully subcloned into the pST50Tr-STRHISNPRAi1 plasmid. I was fortunate that there were a large variety of appropriate gel-purified insert DNA fragments coding

for many proteins available in my laboratory. This was not entirely a coincidence, but instead reflects the modular design of the pST50Tr plasmid component of the pST44 expression system (Tan et al, Protein Expr and Purif, 2005).

4.3 Analysis of Small-scale Expression and Batch Purification

In addition to those proteins with previously prepared insert DNA, the principal considerations for selecting proteins for the marker were expression temperature, expression level, and solubility. As most of the proteins had been used in various experiments within our laboratory, the expression of each protein without the PRA tag was previously tested on a small-scale to determine the optimal conditions. The first proteins considered showed high levels of expression and solubility at 37°C. Proteins requiring lower temperatures and expression levels were used as necessary or if the insert DNA was available. Also, proteins that displayed smearing in previous SDS-PAGE experiments were avoided.

As expected, all of the constructs expressed similarly to their counterparts without the PRA fusion tag. While all of the proteins did not express at high levels and were not completely soluble, sufficient protein was expressed and purified for use in the western blot marker in most cases. However, this was not true for STRHISNPRAyBas1, as the expression level at 18°C was quite low, and other expression temperatures were not attempted. Incomplete lysis of the whole cell extract via sonification may have contributed to the low yields. The soluble material was purified with Talon Metal Affinity resin, which binds tightly to the N-terminal polyhistadine tag (HIS) with six histidine residues.

4.4 Analysis of Large-scale Expression and Purification

The proteins were expressed on a large-scale under the conditions optimized for the small-scale expression studies. However, expression of the 107 kDa STRHISNPRAyBas1 was not pursued on a large-scale due to its low expression level and time constraints for the project. The expression level for STRHISNPRAyReb1Δ1 was somewhat lower than expected, and the protein was not adequately purified from contaminants or proteolysis products. Further purification by high performance liquid chromatography (HPLC) was required to isolate the desired protein. Additionally, the protein STRHISNPRADHFR appeared to have a single contaminant at 25 kDa, although this extra fragment did not appear on the western blot after the material was adequately diluted.

4.5 Analysis of Marker Preparation

The individual proteins were expressed at different levels, and the purified samples were all recovered at different concentrations. In order to mix the samples to the appropriate concentration, the relative concentrations were deduced by diluting the samples and running western blots. Just as in section 4.4, the sample of STRHISNPRAReb1Δ1 was not adequate for use in the marker, and required further purification. Two trials were required for each protein to standardize the protein concentrations at an appropriate medium that did not appear saturated on a western blot (figures 10-13). However, additional trials were designed to incorporate all five of the purified proteins into a single marker lane (see figure 14 and 15). While this produced a fairly “clean” marker lane in trial 1, two of the proteins were of a considerably lower concentration. The last trial, seen in figure 15, confirmed the optimal dilution for each of the five proteins that were

relatively equal in concentration. While my goal of producing a marker with a range above 100 kDa has not yet been achieved, I have created a western blot marker ranging from ~19 kDa to ~78 kDa.

4.6 Future Experiments

While the western blot marker produced in this project will largely satisfy the needs of the laboratory and department, additional modifications could be used to add to the effectiveness of the marker in identifying and resolving proteins on the film. First, the largest marker protein, STRHISNPRAyBas1 was not expressed and purified on a large scale. Incorporating this and even larger proteins into the western blot marker would also widen its effective range. Also, STRHISNPRAyReb1Δ1 (~67 kDa) was not adequately purified and was not used in the final marker set. This protein could be further purified using HPLC and introduced to fill the gap between 46 kDa and 78 kDa.

References

Kihira, Y, and Scott Aiba. "Artificial immunoglobulin G-binding protein mimetic to staphylococcal protein A. Its production and application to affinity purification of immunoglobulin G.." Journal of Chromatography A. 597.1-2 (1992): 277-83.

Kindt, Thomas J., Barbara A. Osborne, and Richard A. Goldsby. Kuby Immunology. Sixth Edition. New York, NY: W.H. Freeman, 2006.

Laemmli, UK. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature*. 227.5259 (1970): 680-5.

Moks, Tomas, Lars Abrahmsen, Bjorn Nilsson, Ulf Hellman, and John Sjoquist. "Staphylococcal protein A consists of five IgG-binding domains." *European Journal of Biochemistry*. 156.3 (2005): 637-43.

Mullins, Dyche. "HOW DOES AN SDS PAGE GEL REALLY WORK?." University of California, San Francisco. N.p., 02/01/2002. 2 Apr 2010. http://mullinslab.ucsf.edu/Protocols%20HTML/SDS_PAGE_protocol.htm>.

Mullis, K, FC Faloona, S Scharf, R Saiki, and G Horn. "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986. ." Cold Spring Harbor Symposia on Quantitative Biology. 51.Pt1 (1986): 263-73.

Renart, J, J Reiser, and GR Stark. "Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure." Proceedings of the National Academy of Sciences of the United States of America. 76.7 (1979): 3116-20.

Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. (1999), Nature Biotechnology. Vol. 17 (10): 1030-1032

Tan, Song, RC Kern, and W Selleck. "The pST44 polycistronic expression system for producing protein complexes in Escherichia coli." Protein Expression and Purification. 40.2 (2005): 385-95.

Tan, Song. "Tan Lab Protocols." The Pennsylvania State University. State College, PA: Department of Biochemistry and Molecular Biology, 2009.

Appendix A

*Adapted from the protocols written by Dr. Song Tan

PCR Amplification
The following reaction mixture was prepared in a 0.2 ml PCR tube for thermocycling.

Water	68.5 µl
10×ThermoPol Buffer	10 µl
2.5 mM dNTP	10 µl
10 ng/µl plasmid	1 µl
10 µM forward primer	5 µl
10 µM reverse primer	5 µl
2.5 units/µl Pfu polymerase	0.5 µl
	100 µl

PCR Primers

Amplify PRA Fusion Tag (introducing BglII and BsrGI sites)
Forward: CCCGAGATCTAAAACCGCGGCTCTT
Reverse: CGGTGTACATTAGGTTGACTTCCCCGC

Insert BamHI Site Before STOP Codon
Forward: TCCCGCGAAATTAATACGAC
Reverse: CGGTGTACATTAGGATCCGGTTGACTTCCCCGC

Site Directed Mutagenesis
The following reaction mixture was prepared in a 0.2 ml PCR tube for thermocycling.

Water	17.7 µl
10×ThermoPol Buffer	2.5 µl
2.5 mM dNTP	2.5 µl
10 ng/µl plasmid	0.5 µl
10 µM forward primer	0.7 µl
10 µM reverse primer	0.7 µl
2.5 units/µl Pfu Turbo polymerase	0.4 µl
	25 µl

PCR Mutagenesis Primers

Remove Internal BglII Sites
Forward: CAAAACGCGTTCTATGAAATCTTACATTTACCT
Reverse: AGGTAAATGTAAGATTTCATAGAACGCGTTTTG

Preparation of Vector DNA
Digest plasmid DNA with appropriate restriction enzymes.

Water	15 µl
0.2 µg/µl plasmid DNA	5 µl
10×NEBuffer	3 µl
1 mg/ml BSA	3 µl
100 mM DTT	1 µl
10-20 units/µl restriction enzyme	1.5 µl
10-20 units/µl restriction enzyme	1.5 µl
	30 µl

Preparation of Insert DNA
Digest DNA with appropriate restriction enzymes to release insert.

Water	5 µl
PCR product in TE	15 µl
10×NEBuffer	3 µl
1 mg/ml BSA	3 µl
100 mM DTT	1 µl
10-20 units/µl restriction enzyme	1.5 µl
10-20 units/µl restriction enzyme	1.5 µl
	30 µl

Ligation
Set up vector only (A) and vector + insert (B) ligations as follows:

	A	B
Water	5.5 µl	4 µl
10×T4 DNA ligase buffer	1 µl	1 µl
100 mM DTT	0.5 µl	0.5 µl
Gel-purified vector DNA (~15 ng/µl)	2 µl	2 µl
Gel-purified insert DNA	None	1.5 µl
40 units/µl T4 DNA ligase	1 µl	1 µl
(*750 units/µl for Blunt-ended)	10 µl	10 µl

PCR Screening of Colonies
Prepare PCR reaction mix: (in 0.5 ml Eppendorf tube)

	4 samples	6 samples	8 samples
Water	58.5	90.3	116.1
10×ThermoPol buffer	9	14	18
2.5 mM dNTP	9	14	18
10 µM forward primer	4.5	7	9
10 µM reverse primer	4.5	7	9
2 units/µl Pfu pol	0.5	0.7	0.9
	86	133	171

Add 19 µl of PCR reaction mix
Add 1 µl of cell suspension

Affinity/Fusion Tag Abbreviations

Peptide	Description	Sequence
HIS	hexahistidine	GSSHHHHHH
STR	Strep II	SWSHPQFEK
AVT	AviTag	GLNDIFEAQKIEWHE
GST	glutathione S-transferase	25 kDa Domain
N	TEV NIa protease recognition site	ENLYFQG
PRA	IgG Binding Domain of Protein A	(see figure 3

PRA DNA Sequence

```
1   AGATCTAAAACCGCGGCTCTTGCGCAACACGATGAAGCCGTGGACAACAAATTCAACAAA 60
61  GAACAACAAAACGCGTTCTATGAGATCTTACATTTACCTAACTTAAACGAAGAACAACGA 120
121 AACGCCTTCATCCAAAGTTTAAAAGATGACCCAAGCCAAAGCGCTAACCTTTTAGCAGAA 180
181 GCTAAAAAGCTAAATGATGCTCAGGCGCCGAAAGTAGACAACAAATTCAACAAAGAACAA 240
241 CAAAACGCGTTCTATGAGATCTTACATTTACCTAACTTAAACGAAGAACAACGAAACGCC 300
301 TTCATCCAAAGTTTAAAAGATGACCCAAGCCAAAGCGCTAACCTTTTAGCAGAAGCTAAA 360
361 AAGCTAAATGGTGCTCAGGCGCCGAAAGTAGACGCGAATTCCGCGGGGAAGTCAACCTAA 420
421 TGTACA 426
```

*BglII site highlighted in blue and BsrGI site highlighted in red.

Academic Vitae

Education

College: **The Pennsylvania State University**
Schreyer Honors College
Major: Bachelor of Science in Life Science
Honors in Biochemistry and Molecular Biology

Research Experience

2 Undergraduate Research/Independent Study Laboratory of Dr. Song Tan The Pennsylvania State University Department of Biochemistry and Molecular Biology		Spring 2007 – Present
2009 Molecular Biology Summer Research Program University of Pennsylvania (and Children's Hospital of Philadelphia) Cystic Fibrosis and Protein Trafficking Research		Summer 2009

Work and Volunteer Experience

3. Personal Racket Stringing and Tennis Apparel Business Over 250 customers and 6 years of experience		Summer 2000 – Spring 2006
2. FOP Holiday Party Overall Chairperson Christmas party for over 1000 disabled children Coordinated over 500 student participants		Fall 2004 and Fall 2005
2. THON Dance Marathon Philanthropy Four Diamonds Fund Raised money for Pediatric Cancer (Over $400,000 with fraternity)		Fall 2006 – Present
• **Peer Leadership in Biology** Tutored freshman students in Introductory Biology The Pennsylvania State University		Fall 2007
Treasurer, Sigma Alpha Mu Fraternity Managed funds and 63 tenant residential property Participated in philanthropic events and blood drives		Spring 2008 – Fall 2008

Clinical Experience

2. Virtua West Jersey Hospital Emergency Room Volunteer		Summer 2007
3. Physician Observation/Shadowing Followed specialists in areas such as pulmonology, ENT, and retina		Summer 2007 - Present

Affiliations/Awards/Activities

• **Cherry Hill East High School Tennis Team** Senior Captain and Four Year Starter All-State and South Jersey Honors		Spring 2003- Spring 2006
• **Dean's List, 3.5 GPA and above** The Pennsylvania State University		Fall 2006 - Present
• **Summer Grant Winner and Research Exhibitionist** Extended Research Internship and Annual Poster Competition The Pennsylvania State University		Summer 2008 – Spring 2009
• **Eberly College of Science "Student Marshall"**		Spring 2010