6-12-2017

Engineering Synthetic Antibody by Expanded Genetic Code

Elizabeth Batiuk
Santa Clara University, ebatiuk@scu.edu

Tracy Nguyen
Santa Clara University, t1nguyen@scu.edu

Casey Kiyohara
Santa Clara University, ckiyohara@scu.edu

SANTA CLARA UNIVERSITY

Department of Bioengineering

I HEREBY RECOMMEND THAT THE THESIS PREPARED
UNDER MY SUPERVISION BY

Tracy Nguyen, Casey Kiyohara, Elizabeth Batiuk

ENTITLED

ENGINEERING SYNTHETIC ANTIBODY BY EXPANDED GENETIC CODE

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

BACHELOR OF SCIENCE
IN
BIOENGINEERING

Zhiwen Zhang — June 12, 2017

Thesis Advisor — date

[signature] — 6/12/17

Department Chair — date

Engineering Synthetic Antibody by Expanded Genetic Code

By:

Tracy Nguyen, Casey Kiyohara, Elizabeth Batiuk

Senior Design Project Report

Submitted to

The Department of Bioengineering

Of

SANTA CLARA UNIVERSITY

In Partial Fulfillment of the Requirements for the degree of

Bachelor of Science in Bioengineering

Santa Clara, California

Spring 2017

Engineering Synthetic Antibody by Expanded Genetic Code

Elizabeth Batiuk, Tracy Nguyen, Casey Kiyohara

Department of Bioengineering

Santa Clara University

2017

Abstract

Antibodies are extensively used in research for diagnostic and therapeutic purposes because of their unrivaled specificity and biomarker binding strengths.[1] Currently, monoclonal antibodies are most commonly used because of their production consistency and purity.[1] However, there are significant ethical and economic challenges associated with producing monoclonal antibodies.[1] Synthetic antibodies provide a promising alternative to monoclonal antibodies in both clinical and research applications.[2]

Our proposed synthetic antibody system incorporates 3,4-dihydroxy-l-phenylalanine (L-DOPA), an unnatural amino acid used to increase binding affinity, into a peptide sequence specific for the prostate specific antigen (PSA), a biomarker for prostate cancer. This addition is predicted to give the synthetic antibody binding affinity and PSA specificity comparable to existing monoclonal antibodies while avoiding their drawbacks.[3] If successful, our system would replace monoclonal antibodies for PSA detection as well as be a promising model for developing countless other synthetic antibodies.

Acknowledgments

Dr. Zhiwen (Jonathan) Zhang

Santa Clara University, School of Engineering

Santa Clara University, Department of Bioengineering

Santa Clara University, Undergraduate Programs Senior Design Project Fund

Table of Contents

List of Figures

Abbreviations

CIP: Calf Intestinal Phosphatase

E. coli: *Escherichia coli*

ELISA: Enzyme-linked Immunosorbent Assays

FPLC: Fast Protein Liquid Chromatography

GFP: Green Fluorescent Protein

6-His: Poly-histidine Amino Acid Motif

IPTG: Isopropyl β-D-1-thiogalactopyranoside

ITC: Isothermal Titration Calorimetry

kDA: kilo Dalton

Kan: Kanamycin

K_d: Dissociation Equilibrium Constant

LB: Lysogeny Broth

L-DOPA: l-3, 4-dihydroxyphenylalanine

MW: Molecular Weight

NEB: New England Biolabs

OD_{600}: Optical Density at 600nm

PBS: Phosphate Buffered Saline

PMSF: Phenylmethane Sulfonyl Fluoride

PSA: Prostate Specific Antigen

P1: Peptide One

Rpm: Revolutions Per Minute

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

TAG: Amber Stop codon

Tet: Tetracycline

tRNA: Transfer Ribonucleic Acid

UV: Ultraviolet

Introduction

Background, Significance, & Motivation

Antibodies are naturally produced by the immune system to fight off foreign invaders.[1] Because they are specific to a single antigen on a pathogen or infected cell and bind strongly to this antigen, immune cells are able to locate and destroy these target cells effectively (See Figure 1).[1] In order to take advantage of the naturally high strength and specificity of antibody-antigen interactions, scientific researchers have commercialized antibodies for diagnostic and therapeutic purposes.[1] Diagnostic applications include Western blots and ELISA, which detect the presence of specific proteins, antibodies, or antigens in a given sample.[1] Transducer-based assays are other diagnostic tools that use antibodies for a wide range of applications, including cancer and diabetes screening.[4] Therapeutic antibodies are used to treat autoimmune disorders, cardiovascular diseases, and types of cancer.[5]

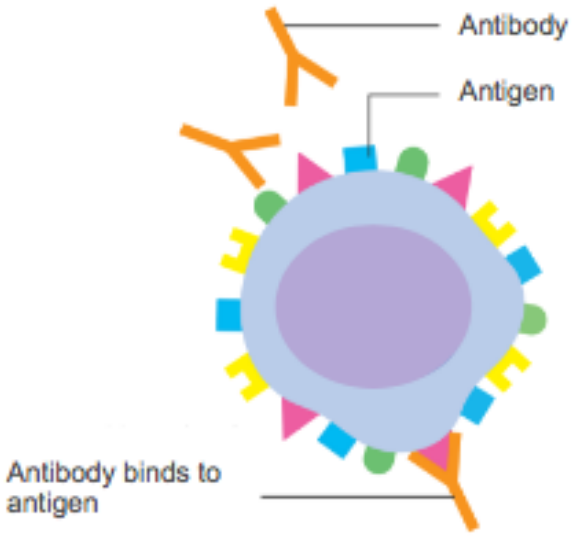


Figure 1: Specific Antibody-Antigen Binding

One application of these commercial antibodies is prostate cancer detection. Prostate cancer is the second most common cancer among men. One in seven men will be diagnosed with prostate cancer in his lifetime.[6] Like most cancer, it is crucial to detect the cancer early, before it

metastasizes.[6] If detected early enough, 96% of diagnosed men will live another 15 years. There are two main methods of detecting prostate cancer.[6] One indication is an enlarged or abnormal prostate discovered during a digital rectal exam.[6] Another is the level of PSA in the blood.[6] As shown in Figure 2, a healthy prostate secretes a small amount of PSA.[6] However, a higher level of PSA in the blood is seen in those with prostate cancer.[6] During routine blood work, the amount of PSA can be detected using a monoclonal antibody.[7]

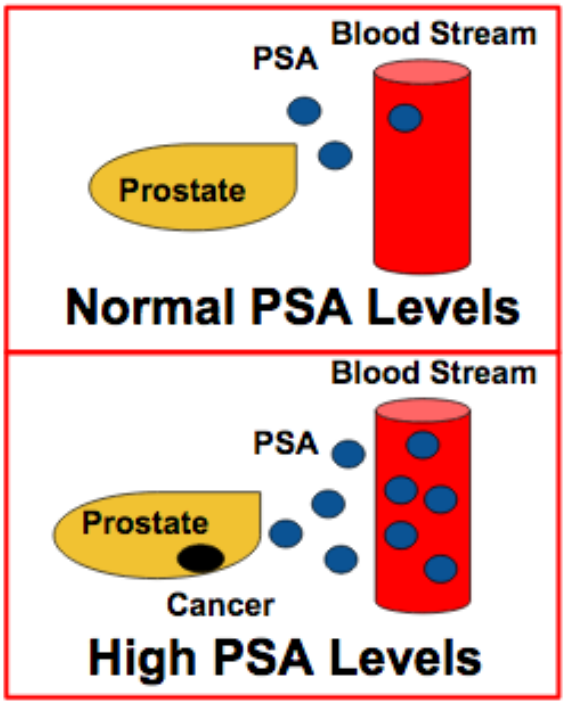


Figure 2. Increased levels of PSA indicate prostate cancer.

Monoclonal antibodies are the gold standard in research and medical applications because of their inherent consistency and purity.[1] However, disadvantages of using monoclonal antibodies include batch-to-batch variability, ethical concerns, and an expensive and lengthy production process.[1] Synthetic antibodies provide a promising alternative to monoclonal antibodies because they overcome many of the significant drawbacks of monoclonal antibodies while maintaining comparable function.[8]

This project focuses on the development of a synthetic antibody, produced in *E. coli*, using L-DOPA, an unnatural amino acid, incorporation for detection of prostate cancer. The antibody is specific for PSA, which can indicate prostate cancer at elevated levels.[7] This synthetic antibody

design offers comparable specificity and binding strength to the monoclonal antibody while also minimizing the drawbacks involved with production. If successful, this system would replace monoclonal antibodies for PSA detection, be a promising model for developing countless other synthetic antibodies, and contribute to the development of personalized medicine in the future.

Literature Review

Design

The basis of this project is a paper by Umeda et al. which details the incorporation of an unnatural amino acid, L-DOPA, to a peptide (TOP1), GFP reporter protein, and 6-His purification tag.[3] This system, with the exception of the peptide sequence specific for PSA, is the same as the proposed system of this project. The peptide of the synthetic antibody described in the paper is specific to the antigen Abelson tyrosine kinase (AbI).[3] Similarly, Therriault and Evans designed a system with a similar design to that of Umeda et al. that targets PSA.[14] Although these systems differ from the synthetic antibody in structure and target, the methods and protocols for unnatural amino acid incorporation and protein expression used in these papers will be a model and starting point for procedures used throughout this project.

Unnatural Amino Acids

Unnatural amino acids are amino acids not used by ribosomes in cells to make proteins naturally.[8] They can have unique chemistry compared to natural amino acids and have therefore been proposed as a solution to the inherently weak binding strength between small peptides and larger proteins when incorporated into the peptide.[3] L-DOPA is the unnatural amino acid used by Umeda et al. to increase the binding strength of a small peptide to a larger protein.[3] The unnatural amino acid performs a redox reaction with a nearby nucleophile in the presence of sodium periodate, creating a strong covalent bond.[3] L-DOPA is particularly useful for unnatural amino acid incorporation into a recombinant protein because it is orthogonal to natural amino acids, which ensures correct translation of the peptide including incorporation of L-DOPA.[3] L-DOPA will therefore be used in this project to increase the binding strength of the synthetic antibody to PSA using a covalent bond.

Unnatural amino acids such as L-DOPA can be incorporated into the peptide by introducing into the cell a mutant tRNA and tRNA synthetase, which are designed to recognize the amber stop codon (TAG/UAG) on mRNA and incorporate the unnatural amino acid at that location on the growing peptide.[8] The amber stop codon is used because of the three stop codons used naturally, it is the least frequently used in *E. coli* (7%) and is rarely is used on essential proteins.[8] Using this stop codon minimizes the likelihood of unnatural amino acid incorporation disrupting normal cell function and killing the cell.[8] The amber stop codon will therefore be used to incorporate L-DOPA into this synthetic antibody. The pAC-DHPheRS-6TRN plasmid used by Therriault and Evans contains the genetic code for the mutant tRNA and tRNA synthetase, and will be used in this project.[14]

The use of unnatural amino acid incorporation in any protein, peptide, or small molecule system is a cutting edge technology.[9] Researchers and pharmaceutical companies have began using this technology to engineer antibodies.[9] As this process is developed, they experience many obstacles to synthetic antibody expression and manufacturing.[9] These production challenges will be important considerations in this project as it attempts to develop an accessible and viable product for companies.[9] The solutions other researchers discovered to overcome these challenges will be helpful for this project.[9]

One issue faced by researchers incorporating unnatural amino acids at the carboxy terminus of a gene is truncation of the protein.[10] It is very difficult to separate the truncated protein from the fully-translated protein, and a group of researchers overcame this issue by attaching an intein protein.[10] The intein protein is attached with a tag after the unnatural amino acid amber stop codon and promotes full translation of the protein.[10] This method achieves higher purity and function and avoids a commonly encountered problem when purifying unnatural amino acids.[10] This method would be considered for this project if truncated protein expression becomes an issue.

Methods and Procedures

One major benefit of this project is avoiding using animals to produce antibodies.[8] Instead, *E. coli* will be used to produce comparable antibodies.[8] *E. coli* are commonly used for protein

production because introducing exogenous recombinant DNA, such as the plasmid for the synthetic antibody, into them is an established process.[18] In addition, by using *E. coli*, the synthetic antibody production is able to avoid the ethical concerns of using and killing animals.[11] The use of *E. coli* also limits economic cost, as scaling up production in *E. coli* is fast due to the short doubling time of the cells, allowing for high density cultures in bioreactors with simple growing conditions, rather than requiring large animal facilities.[18] These characteristics of *E. coli* as a production platform therefore make it ideal for use in this project.

However, since there are big differences between mammalian and *E. coli* protein production, it is important to understand the drawbacks to using *E. coli*.[8] These include susceptibility to genetic mutations as well as the lack of ability to perform post-translational modifications.[8] Since the synthetic antibody used in this project will only be 45 nucleotides, sequencing will be used to verify the DNA sequence is correct. The synthetic antibody also does not require any post-translational modifications because of its length, making *E. coli* a simple and effective expression platform.

There are three published methods of protein expression with L-DOPA incorporation into peptides with similar designs to this project's synthetic antibody.[3,16,17] They vary by inducer, incorporation time, and induction and growth times. To optimize expression and yield of the synthetic antibody, this project will perform separate but simultaneous experiments using each of these three methods to identify and develop the most effective expression protocol.

Prostate Cancer Detection

After successfully producing the synthetic antibody, one important aspect of this project will be assessing whether it is truly comparable to monoclonal antibodies for both diagnostic and therapeutic applications. Researchers have found that the K_d of PSA to immobilized monoclonal antibody on a diagnostic chip is 1.1 ± 0.2 nM, giving us a metric by which to assess the binding affinity of our system.[7] In addition, researchers also identified a concentration of PSA (2.6 ng/mL to 4.0 ng/mL) which detects the presence of small prostate cancer without over-diagnosing patients.[7] This range is a design specification necessary for the synthetic antibody to

achieve target specificity and sensitivity to PSA. Together, these results will determine whether the synthetic antibody is a viable alternative to monoclonal antibodies.

Critiques of Current Literatures and Technologies

Monoclonal antibodies are extensively used in research.[1] However, there are many disadvantages of using monoclonal antibodies, including the use of animals, a long manufacturing process and high production costs.[1] Antibody production is stimulated in animal immune systems by injecting them with the target antigen.[5] Once the animal's B-cells start producing antibodies specific to the antigen, the B-cells are harvested.[5] In order to scale up antibody production, the B-cells are combined with an immortal cell line.[5] The resulting mass production of antibodies is homogenous and pure.[5] This entire process can take months to accomplish and is associated with significant production costs.[5] As a result, this leads to increased medical expenses for consumers and limits accessibility to these technologies.[11] In addition, the use of animals raises many ethical dilemmas and therefore is highly debated in scientific research.[11]

The primary alternative to monoclonal antibodies is polyclonal antibodies, which have the same structure and specificity as monoclonal antibodies while requiring less time, skill, and money to produce.[1] However, because they still are produced in animals, they have the same ethical concerns as monoclonal antibodies.[1] In addition, polyclonal antibodies derive from several different animals and B cell lines at different time points, resulting in batch-to-batch variation not present in monoclonal antibodies, which stem from a single cell line and are therefore more consistent.[1]

Among the other alternatives to monoclonal antibodies scientists are currently developing is protein scaffolds.[12] Protein scaffolds have a general protein framework with specific peptides or amino acids, which give them specificity to a target, incorporated in them.[12] These scaffolds can be diverse in size, folding, and method of interaction with their targets, giving them more flexibility than antibodies, which have a constant structure.[12] However, because there is very little data on their immunogenicity and degradability in biological fluids, their potential applications are currently limited.[12]

Another alternative to monoclonal antibodies currently being explored is aptamers.[13] Aptamers are strands of DNA or RNA which are selected from large libraries to be specific to a certain target.[13] Because they involve no animal use, they avoid the ethical issues of monoclonal antibodies.[13] In addition, because they are made of DNA or RNA, they are more stable than antibodies when being stored.[13] However, aptamers are known to degrade quickly when exposed to nucleases in biological fluid, limiting their potential applications, particularly for therapeutics.[13]

Project Objectives

Based on the strengths and drawbacks of monoclonal antibodies, the overall goal of this project was established to be designing an alternative that maintained the strengths of monoclonal antibodies while avoiding their drawbacks as much as possible. The final project objective was to develop a cost-effective, ethical alternative to monoclonal antibodies that is easier and faster to manufacture with comparable quality: specificity, sensitivity, and binding strength.

Overall Design

To meet the project objectives, a synthetic antibody produced in *E. coli* was designed. By being produced in *E. coli*, the synthetic antibody has a lower cost, shorter, and more ethical production process than that of monoclonal antibodies.[5,11] The physical design of the synthetic antibody contains four components: peptide, L-DOPA unnatural amino acid, GFP reporter protein, and 6-His purification tag (See Figure 3).

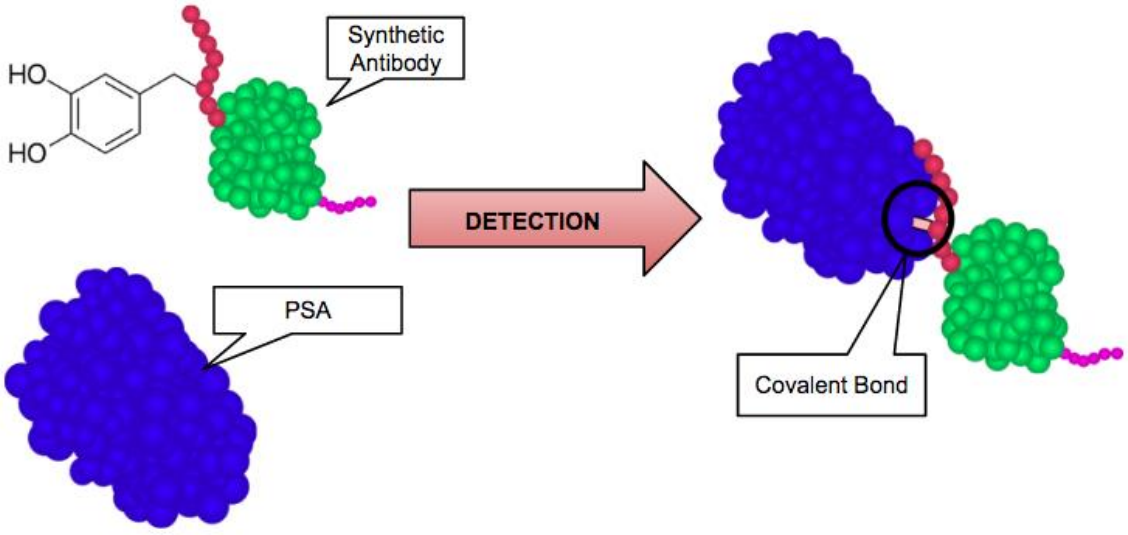


Figure 3. **Synthetic Antibody Detection and Binding of PSA.**

Components of the synthetic antibody include peptide (red), L-DOPA unnatural amino acid (black/red), green fluorescent protein (green), 6-His tag (purple). Exposure of synthetic antibody-PSA complex to sodium periodate (detection) allows for formation of a covalent bond (pink) between them.

The 15-amino-acid experimental peptide was designed by a previous Senior Design group, Therriault and Evans.[14] Using open source software developed by University of California, San Francisco, called Chimera, a library of peptides was screened using several parameters.[14] These parameters include electrostatic interactions and predicted binding affinity that were determined using the known active site where a commercially available monoclonal PSA antibody binds to PSA.[14] The top peptide candidate that was screened was selected as the experimental peptide for the design.[14] The peptide therefore is predicted to give the synthetic antibody high specificity to PSA.[14]

The unnatural amino acid, L-DOPA, was added to the end of the peptide sequence. As discussed in the literature review, this unnatural amino acid is capable of performing a redox reaction that creates a covalent bond when exposed to sodium periodate.[3] When incorporated into the synthetic antibody, this covalent bond is expected to form between the synthetic antibody and PSA once the peptide has specifically targeted the active site of PSA and sodium periodate is added.[3] This covalent bond greatly increases the binding strength of the synthetic antibody to PSA.[3] In order to incorporate this unnatural amino acid, a plasmid designed by Wang et al.,

pAC-DHPheRS-6TRN, will be transformed into the *E. coli* with the synthetic antibody plasmid.[17] This plasmid contains the genetic code for the mutant tRNA and tRNA synthetase necessary for unnatural amino acid incorporation.[17]

GFP was added to the synthetic antibody sequence after the unnatural amino acid. GFP is a commonly used reporter protein in research applications because of its simple visualization process requiring only UV light.[19] It therefore is used to visualize when and how much synthetic antibody is present during detection. GFP's straightforward visualization process was also taken advantage of to verify synthetic antibody expression, as its characteristic green fluorescence is visible in *E. coli* correctly producing the synthetic antibody.

The final component of the synthetic antibody, the 6-His tag, was a sequence of six consecutive histidines at the end of the synthetic antibody sequence. This 6-His sequence is well-established in the scientific community because its unlikelihood of occurring naturally in another *E. coli* protein makes it unique to the protein of interest and therefore ideal for targeting in protein purification.[20] This sequence therefore allows the synthetic antibody to be purified from the *E. coli* using affinity chromatography that targets this 6-His tag.

The synthetic antibody sequence was designed so that the ribosomes in the *E. coli* would first translate the peptide, followed sequentially by L-DOPA, GFP, and the 6-His tag. This design allows the ribosome to incorporate the unnatural amino acid, which stresses the ribosome more than natural amino acid incorporation, before translating the long GFP sequence.[21] It has been hypothesized that translating GFP stresses the ribosome because of its length, and therefore makes the ribosome more prone to translation errors.[21] Because of the expected crucial nature of the unnatural amino acid to the binding strength, and therefore effectiveness, of the synthetic antibody, the antibody was designed to minimize the chance of error in unnatural amino acid incorporation caused by ribosomal stress.

The synthetic antibody sequence was cloned into the pET-28b vector, which then was transformed into TOP10 competent *E. coli* cells. The pET-28b vector includes a lac operon and T7 promoter that allow for regulated expression.[14] It also includes a Kan resistance gene which

allows for selection of *E. coli* successfully transformed with this vector.[14] Co-transformed into the *E. coli* was the second plasmid, pAC-DHPheRS-6TRN (See Appendix A). This plasmid contains a Tet resistance gene used for selection of *E. coli* that have been successfully transformed with this second plasmid.[14] Therefore, cells were grown in the presence of both Kan and Tet in order to select for cells which have been successfully transformed with both plasmids.

Milestones & Expected Results

The first milestone of the project is correct plasmid design. This plasmid provides the *E. coli* with the DNA sequence necessary to successfully produce the synthetic antibody. The final design must incorporate the sequence for all four components of the synthetic antibody adjacent to a promoter in the pET-28b vector to allow the *E. coli* to produce the synthetic antibody (See Appendix B).

The second milestone of the project is successful cloning and transformation. Here, the complete synthetic antibody sequence must be incorporated into the pET-28b vector at the correct location so that the cells transformed with these plasmids will be able to produce the synthetic antibody. The *E. coli* must then successfully be transformed with this plasmid and the pAC-DHPheRS-6TRN plasmid in order to begin producing the synthetic antibody.

Once the *E. coli* have been successfully transformed with the two plasmids, the third milestone of the project is successful expression of the synthetic antibody by *E. coli*. As mentioned in the literature review, there are several previously established methods of expression of proteins similar to that designed in this project. All were performed and tested while developing a successful expression protocol for the synthetic antibody.

After confirming expression, the next milestone is successful purification of the synthetic antibody. In order to reach a high standard of purity that is comparable to that of monoclonal antibodies, the synthetic antibody must be purified from all other proteins and other materials within the *E. coli* which is producing it. The effectiveness of the purification must be tested, optimized, and ultimately confirmed to ensure this high level of purity.

The final milestone of the project is testing of the synthetic antibody. In order to be comparable to monoclonal antibodies, and therefore a viable alternative, the synthetic antibody must either meet or exceed the standards of binding strength, specificity, and sensitivity set by monoclonal antibodies. As described in the literature review, these standards for the anti-PSA monoclonal antibody are well established, and therefore can be used to test and confirm the efficacy of the synthetic antibody.

Team Management

To maintain accountability and efficiency throughout the project, various responsibilities were delegated among the team. Tracy is the liaison with the School of Engineering, handles team's finances, and acts as scribe during team and advisory meetings. She is also responsible for ensuring any deadlines are met. Elizabeth coordinates both meeting and laboratory scheduling for team and facilitates communications with the advisor, Dr. Zhang. Casey supervises all writing and editing for the final report and communicates with industry contacts. The laboratory work, report writing, and presentation development was divided equally among team members.

Chapter 1: Vector Design & Cloning

Introduction

In designing the cloning of the synthetic antibody sequence into the vector, the synthetic antibody sequence must incorporate restriction sites that match ones at specific sites in the vector in order to localize the synthetic antibody sequence near the promoter.

Because there is a Kan resistance gene included in the vector, it is expected that cells which have been successfully transformed with the pET-28b vector will be resistant to Kan. Of the cells which are resistant to Kan, some will be selected for DNA sequencing, which will confirm that the cloning of the synthetic antibody sequence into the vector was successful.

Back-up Plan

If cloning is not accomplished in the allotted time, the oligonucleotide sequences will be sent to a company to be cloned into pET28b.

Design Logic and Reasoning

Peptide Design

Two sets of forward and reverse single-stranded oligonucleotides of the peptide fragments coding for the synthetic antibody and negative control were purchased from BioBasic. One set of oligonucleotides has the amber stop codon (TAG) sequence which would allow for successful unnatural amino acid incorporation while the other set has a codon for alanine (GCG) in place of the TAG sequence, serving as the negative control (See Figure 4). This negative control, which besides this one amino acid is identical to the experimental synthetic antibody, will be used during the testing phase to confirm that L-DOPA has been incorporated and is increasing the binding strength of the synthetic antibody as expected. The forward and reverse oligos were constructed to have 17 overlapping nucleotides to optimize annealing efficiency and specificity. Each set of oligonucleotides (experimental (TAG) and control (GCG)) were annealed together to form two double-stranded fragments. Then the double-stranded fragments were extended into

complete inserts through a klenow reaction. The klenow reaction was selected over a PCR reaction due to its high efficiency of extension for short fragments such as this peptide, which is less than 150 base pairs.

```
Experimental        1 ATGTGCGTGGCGTATTGCATTGAACATCATTGCTGGACCTGCTAG     45
                      |||||||||||||||||||||||||||||||||||||||||..|
Control             1 ATGTGCGTGGCGTATTGCATTGAACATCATTGCTGGACCTGCGCG     45
```

Figure 4. Experimental and control nucleotide design sequence alignment. Experimental sequence is specific for PSA and contains 3' TAG codon for L-DOPA incorporation. Control peptide serves as a negative control for L-DOPA incorporation and contains a codon for alanine.

Expression Vector Selection

pET28b-GFP was selected as the expression vector because it contains an enhanced GFP. It also contains a 6-His that will be used during protein purification by affinity chromatography. The oligonucleotides were cloned on the 3' end of the T7 promoter and 5' of the 6-His and GFP (See Figure 5).

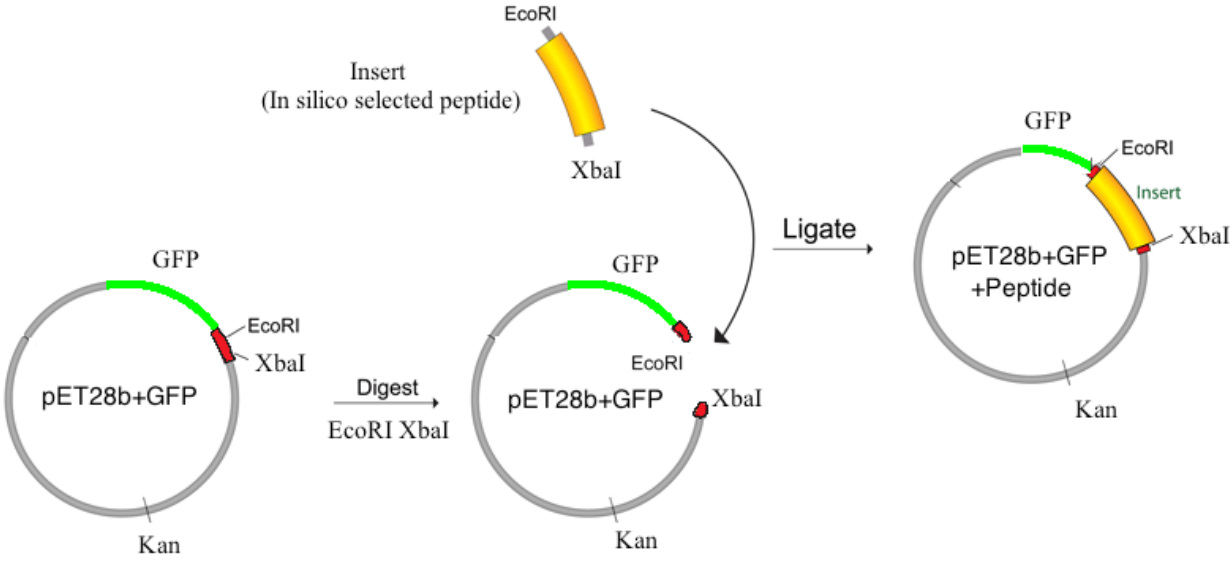


Figure 5. Step-by-step of double digestion of vector to insert peptide.

Materials and Methods

pET28b-PI-TAG-GFP Vector Construction

The peptide predicted to be specific for PSA was selected through an *in silico* screen as described by Therriault and Evans.[1] Oligonucleotides containing experimental or control peptides, restriction sites, and flanking sequences were ordered as two overlapping fragments (BioBasic), annealed, and extended using the large polymerase I, large (klenow) fragment for 15 minutes at 25℃ (NEB, See Appendix E Table 1). Products were run on a 2% agarose gel for verification. Oligonucleotides were purified by QIAquick PCR Purification Kit (Qiagen) and double digested with restriction endonucleases EcoRI-HF and XbaI (NEB) for 10 hours at 37℃ (See Appendix E Table 2). Products were purified (Qiagen) prior to ligation.

pET28b-GFP contains a 6-His and genes for Kan resistance and GFP. 5 mL cultures were grown in LB (Teknova) with 50 μg/mL Kan (Teknova) and plasmids were purified by QIAprep Spin Miniprep Kit (Qiagen). Plasmids were double digested for four hours at 37℃ (See Appendix E Table 3) and run on 1% agarose gel, gel extracted, and purified using the QIAquick Gel Extraction Kit (Qiagen). 5' DNA ends were dephosphorylated using CIP (NEB, See Appendix E Table 4) and subsequently purified (Qiagen) prior to ligation in order to minimize self-ligation of the vector without incorporation of the insert.

Ligations (one set for each the control and experimental oligonucleotides) were performed using T4 DNA Ligase (Promega) at 1:1, 1:2 and 1:10 ratios of pET28b-GFP:oligonucleotide (See Appendix E Table 5) and transformed into TSS TOP10 chemically competent cells (as described in Chung et al.) before being plated onto LB-agar (MP Biomedicals LLC) plates containing 50 μg/mL Kan.[15] Colonies were picked and grown in 5 mL LB + 50 μg/mL Kan cultures. DNA was purified (Qiagen) and sequenced (Sequetech) using the T7 primer.

Results

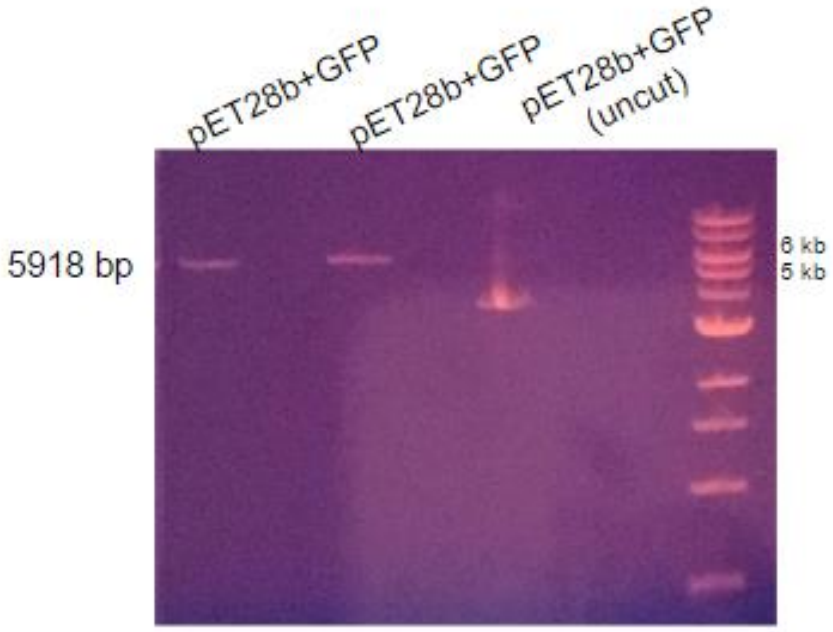


Figure 6. 1% agarose gel for pET28b-GFP vector cut (left two bands) and uncut (right band) by EcoRI-HF and Xba1 restriction enzymes. Visualized using Fisher Scientific Transilluminator and 0.5 µg/mL EtBr. Ladder (far right) was Quick-load 1kb DNA ladder (NEB).

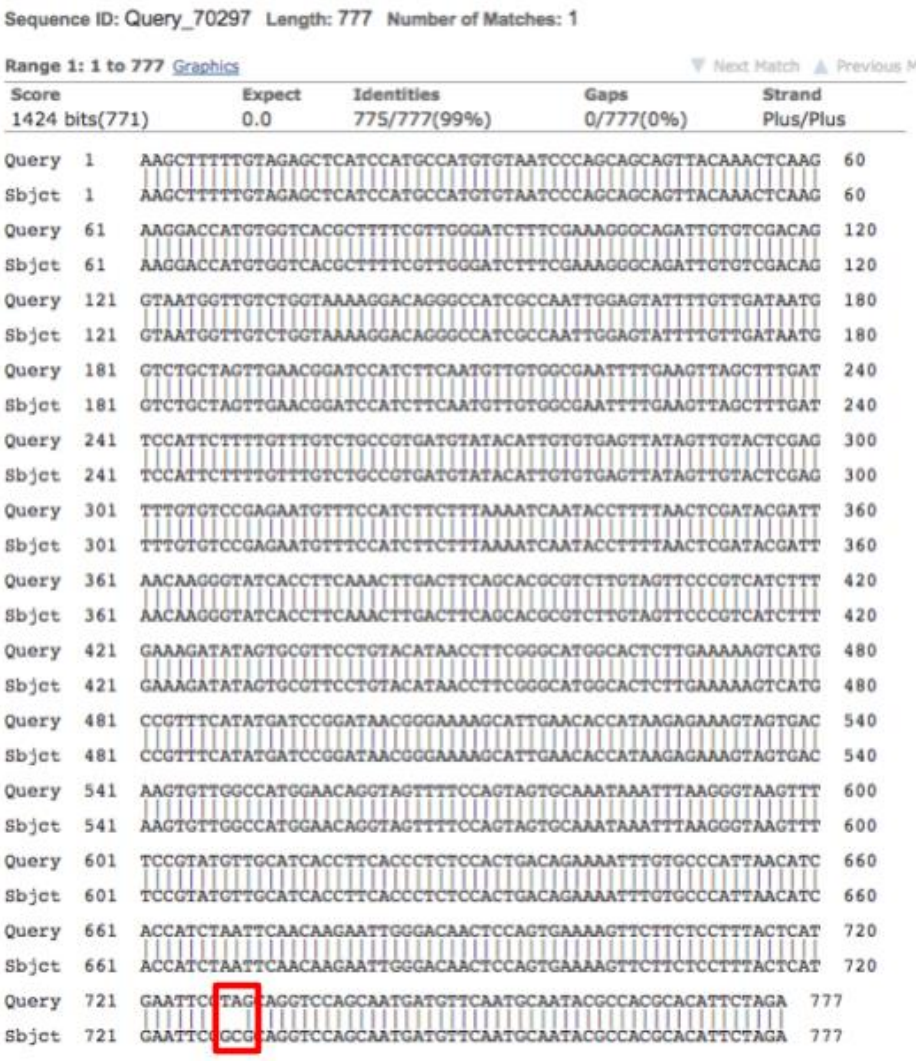


Figure 7. Sequence Alignment of Experimental (Query) and Control (Sbjct) Vectors. Highlighted in red is the one amino acid difference between the two sequences.

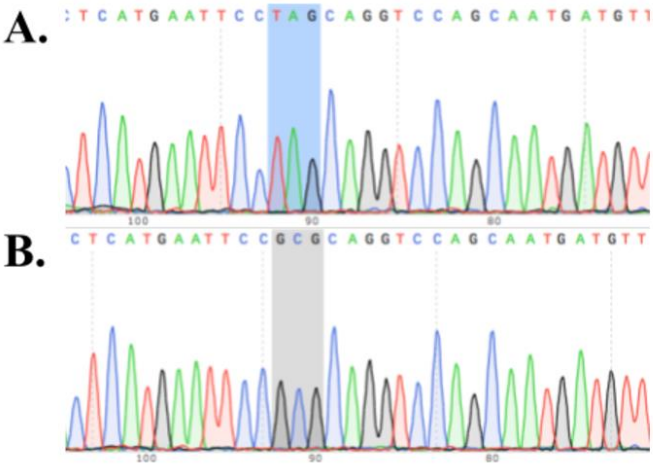


Figure 8. Chromatograms of Experimental (A) and Control (B) sequences at location of the amino acid difference.

Discussion

As seen in Figure 6, the vector was successfully double digested, however after sending the vectors to sequencing after cloning, the vectors were confirmed to have self-ligated instead of incorporating the peptide and L-DOPA. Varied time and order of digestion, and time of CIP reaction were performed. Unfortunately, due to time constraints, the designed vectors (one for control and one for experimental) were constructed at Epoch Life Science, Inc (Missouri City, TX). The vectors were sequenced and results are shown in Appendix C and Appendix D. Figure 7 shows the nucleotide alignment between the experimental and control vector sequence to verify they are correct and differ at the last codon of the peptide. Figure 8 shows the chromatogram of both the experiment and control peptides. The little background noise and clean peaks show the clones are correct and pure. The confidence that the vectors are pure and accurate is necessary before moving on to the next step of the project, synthetic and control antibody expression.

Chapter 2: Protein Expression

Introduction

In order to maximize efficiency of transformation, a stepwise transformation was performed, with cells successfully transformed with one plasmid then being made competent before being transformed with the second plasmid (See Figure 9). The pAC-DHPheRS-6TRN plasmid was transformed into E. coli first to increase transformation efficiency and minimize toxicity concerns caused by the mutant tRNA and tRNA synthetase. In addition, transformation of the pAC-DHPheRS-6TRN plasmid first allows for the potential future creation of a commercialized kit, containing competent cells into which a plasmid for any synthetic antibody containing L-DOPA can be transformed and immediately expressed.

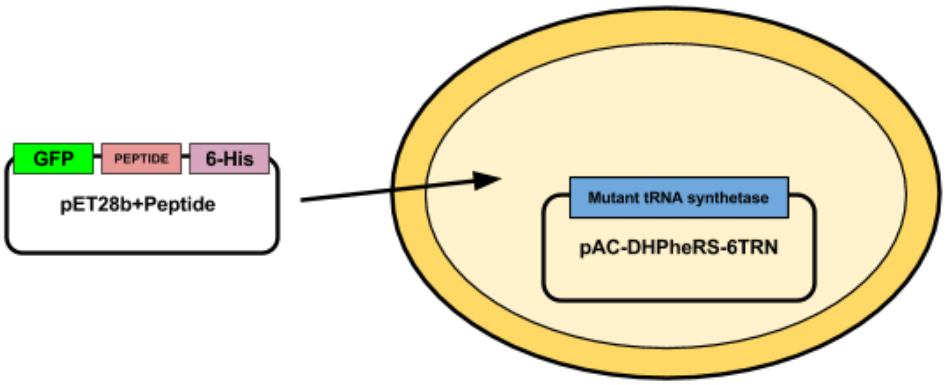


Figure 9. Sequential co-transformation of pET28b+P1+GFP and pAC-DHPheRS-6TRN

In this step, confirmation of successful expression is performed by taking advantage of the design of the synthetic antibody. Because the design includes GFP, cells which are successfully producing synthetic antibody are expected to have a visual green fluorescence in the culture. Therefore, during expression of the synthetic antibody, the *E. coli* cultures were observed under UV light to confirm that there is this indicative green fluorescence.

Back-up Plan

Although there are a few different aspects of expression where issues may arise, the methods of this project were designed to minimize those risks. In order to ensure expression, three different established L-DOPA protein expression methods were separately used to optimize yield and expression.[3,16,17] Since L-DOPA is easily oxidized, the pH was carefully controlled during expression. As discussed previously, expression of proteins with unnatural amino acids can result in truncated protein. Since GFP is at the carboxy terminus of the synthetic antibody, a truncated synthetic antibody would not fluoresce during expression. An intein protein could be attached to minimize truncation.[10] If these issues or others are experienced and not resolved in a timely manner, the experimental and control synthetic vectors will be sent to a company to be expressed.

Materials and Methods

pAC-DHPheRS-6TRN Competent Cell Preparation

pAC-DHPheRS-6TRN contains genes for the mutant tRNA and mutant tRNA synthetase as well as the Tet resistance gene. pAC-DHPheRS-6TRN was transformed into calcium chloride competent TOP10 *E. coli* cells using a high efficiency chemical transformation protocol (NEB). Cells were plated on LB-agar + 10 µg/mL Tet (Teknova). A single colony was picked and grown in LB broth + 10 µg/mL Tet and a glycerol stock was prepared. Chemically competent cells from this glycerol stock were made using the TSS protocol.[15]

Co-Transformation of pET28b-PI-TAG-GFP

5 mL cultures of pET28b-PI-TAG-GFP and pET28b-PI-GCG-GFP (Epoch Life Science, Inc) were grown overnight before plasmids were purified (Qiagen) and transformed into pAC-DHPherRS-6TRN TOP10 competent *E. coli* cells and plated on LB-agar + 30 µg/mL Kan + 25 µg/mL Tet plates.

Antibody Expression Protocol #1

This method was adapted from Zhang et al.[16] Co-transformed colonies (4 from TAG, 4 from GCG) were picked from LB-agar (Teknova, MP Biomedicals LLC) plates and grown in 2.0 mL

LB + 30 µg/mL Tet and 60 µg/mL Kan, shaking in the dark at 37℃ overnight. 25% glycerol stocks were made for each cell line and stored at -80℃. 500 µL of overnight culture was added to 125 mL of M9 Medium Broth (Amresco) with 30 µg/mL Kan + 25 µg/mL Tet and 1 mM L-DOPA (Sigma-Aldrich) and grown in the dark in a shaking incubator at 37℃ until OD_{600} reached 0.5. The pH of the culture was maintained at pH 6.5 during growth. At OD_{600} 0.5, 0.2% L-Arabinose (Sigma-Aldrich) was added and cultures were grown for 4 hours at 37℃. Cultures were harvested by centrifuge at 5,000 x g for 11 minutes at 4℃, supernatant was decanted and pellets were stored at -80℃.

Antibody Expression Protocol #2

This method was adapted from Wang et al.[17] Co-transformed colonies (4 from TAG, 4 from GCG) were picked from LB-agar plates and grown in 2.0 mL LB + 30 µg/mL Tet and 60 µg/mL Kan, shaking in the dark at 37℃ overnight. 25% glycerol stocks were made for each cell line and stored at -80℃. 500 µL of overnight culture was added to 125 mL of M9 Medium Broth with 30 µg/mL Kan + 25 µg/mL Tet and 1 mM L-DOPA (Sigma-Aldrich) and grown in the dark in a shaking incubator at 37℃ until OD_{600} reached 0.5. The pH of the culture was maintained at pH 6.5 during growth. At OD_{600} 0.5, 1 mM IPTG was added and cultures were grown for 5 hours at 37℃. Cultures were harvested by centrifuge at 5,000 x g for 11 minutes at 4℃, supernatant was decanted and pellets were stored at -80℃.

Antibody Expression Protocol #3

This method was adapted from Umeda et al.[3] Co-transformed colonies (4 from TAG, 4 from GCG) were picked from LB-agar plates and grown in 2.0 mL LB + 30 µg/mL Tet and 60 µg/mL Kan, shaking in the dark at 37℃ overnight. 25% glycerol stocks were made for each cell line and stored at -80℃. 500 uL of overnight culture was added to 125 mL of M9 Medium Broth with 30 µg/mL Kan + 25 µg/mL Tet and grown in the dark in a shaking incubator, at 37℃ until OD_{600} reached 0.6. During this time, the pH of the culture was maintained at pH 6.5. At OD_{600} 0.6, 1 mM IPTG was added to the cultures, which then were incubated at 30℃ for 6 hours. Cultures

were viewed after 4 hours using a Fisher Scientific Transilluminator. Cultures were harvested by centrifuge at 5,000 x g for 11 minutes at 4℃, supernatant was decanted and pellets were stored at -80℃.

Results

Co-Transformation of pET28b+PI+TAG and pAC-DHPheRS-6TRN

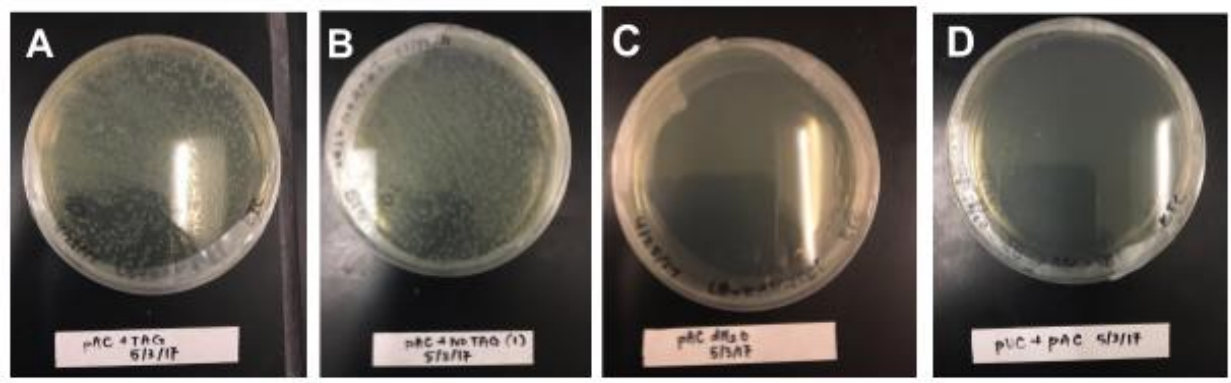


Figure 10. Transformed E. coli + pAC-DHPheRS-6TRN plated on LB + agar + 10 µg/mL Tet+ 50 µg/mL Kan + 0.2% arabinose. TOP10 E. coli + pAC-DHPheRS-6TRN transformed with experimental synthetic antibody plasmid (A), control synthetic antibody plasmid (B), no plasmid (C), plasmid with no synthetic antibody (D).

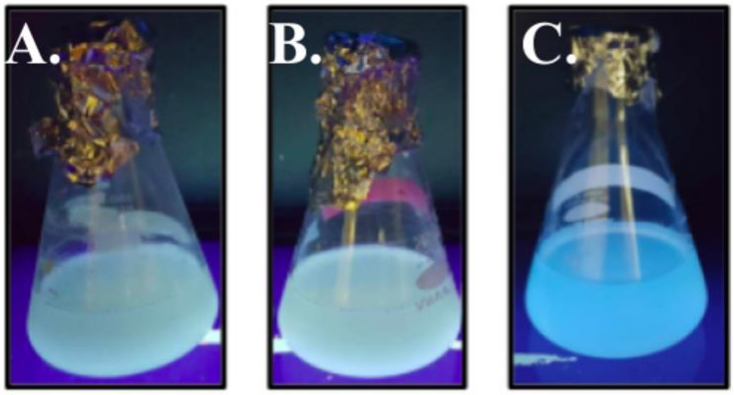


Figure 11. Experimental (A), control (B), and no GFP (C) cultures using antibody expression protocol #3 visualized under UV light after 4 hours of synthetic antibody expression.

Discussion

As shown in Figure 10, the co-transformation of pET28b+PI+TAG and pET28b+PI+GCG (control) into TOP10 competent cells containing pAC-DHPheRS-6TRN was successful. Plates C and D serve as negative controls. Plate C was plated with cells transformed with water and plate

D was plated with cells transformed with a plasmid without Tet resistance, a quality control for the plates.

For the cultures using expression protocol #3, the cultures were placed in UW light 4 hours after induction of synthetic antibody production with IPTG (Figure 11). Compared to the culture with cells not producing the synthetic antibody, which therefore were not producing GFP, the experimental and control cultures had a distinct green color under the UV, indicating that they were producing GFP and therefore the synthetic antibody. These results indicate successful expression of the synthetic antibody within the E. coli cells.

Chapter 3: Protein Purification

Introduction

Before testing the functionality of the synthetic antibodies, they must first be purified. This was done by manual affinity chromatography using 6-His. After lysing the cells, the proteins were separated and passed through a column containing Ni-NTA beads. Only the synthetic antibody will bind to the beads allowing for the other proteins to be removed. The synthetic antibody was then eluted as a pure sample from the beads.

Back-up Plan

Although manual affinity chromatography is a well-defined purification procedure, if any issues arise, FPLC will be used. This automated process produces pure and large quantities of proteins. However, it would be time consuming to troubleshoot and learn how to use this machine. Purification can also be contracted out to a company if necessary.

Materials and Methods

A separate purification was performed for each individual cell type and growing condition. *E. coli* pellets were thawed in a hot water bath for approximately 10 minutes. Each pellet was then resuspended in 5 mL of lysis buffer (PBS + 1 mg/mL lysozyme + 0.5 mM PMSF + 10 mM imidazole) and left on ice for 20 minutes. The cells were then sonicated 6 times at 30% alternating between 6 seconds on and 6 seconds off. The lysate was then centrifuged at 4℃ and 16,000 rpm for 40 minutes. 1 mL sample of the supernatant was taken and stored in -20℃.

While centrifuging, one set of HisPurTM Ni-NTA beads (Thermo Fisher Scientific) specific to the 6-His tag was prepared for each culture by resuspending 125 µL of the beads (1:1000 ratio of volume of beads:volume of initial culture) in 5 mL of PBS + 25 mM imidazole before being centrifuged for 4 minutes at 2000 rpm. After pouring out the supernatant, this process was repeated 4 times.

The supernatant of the centrifuged lysate solutions was then used to resuspend the Ni-NTA beads (one culture per set of beads). Binding of the proteins to the beads was then done by rocking the resuspended beads at 4℃ for 1 hour before centrifuging them for 4 minutes at 2000 rpm at 4℃. 1 mL of the flow-through after binding was then sampled and the rest was discarded. The beads were then washed 5 times. In these wash steps, the beads were resuspended in PBS + 25 mM imidazole wash buffer, rocked for 20 minutes at 4℃, then centrifuged at 2000 rpm at 4℃ for 4 minutes before the supernatant was sampled and discarded.

For elution, the beads were resuspended in 200 µL of PBS + 250 mM imidazole, rocked for 20 minutes, and centrifuged at 2000 rpm for 4 minutes. The supernatant, which ideally contains the target protein, was then extracted and stored at -20℃. Two elution steps were performed for each culture.

Samples of lysate, flow-through, initial wash step, and elution for each purification were run on a SDS-PAGE (Genscript) at 120 volts for 60 minutes. Gels were visualized and imaged using LabSafe GEL Blue (G-Biosciences).

Results

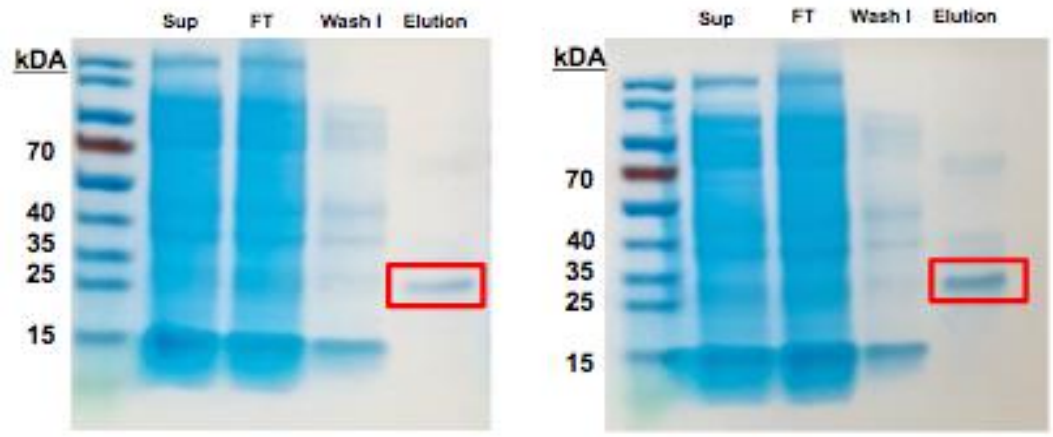


Figure 12. SDS-PAGE of purification Left image shows synthetic antibody and right shows control antibody. Lane one is PageRuler Prestained Protein Ladder (ThermoFisher Scientific). Lane two is sample of supernatant after cells were lysed and centrifuged. Lane three is flow-through sample after binding to column. Lane four is sample after first wash is performed. Lane five is sample of first elution.

Discussion

Both the control and experimental SDS-PAGEs show a large amount of proteins of many sizes in the supernatant lane, which is expected as this was a sample of all of the proteins in the *E. coli*. Many of these proteins did not bind to the beads and were therefore also in the flow-through, which would also be expected because the Ni-NTA HisPur beads which were used were specific to the 6-His unique to the synthetic antibody, preventing other proteins without this tag from binding. The wash lanes indicate that there was some nonspecific binding of non-synthetic antibody proteins to the beads, which were then removed by the wash steps as the higher concentration of imidazole allowed it to competitively bind to the beads.

The elution lanes show a strong band at approximately 27 kDa, which is the predicted size of our synthetic antibody. Therefore, the antibody has essentially been purified. However, there are some remaining bands of various weights shown in the elution lanes, indicating that there is still some non-specific binding occurring, which would need to be eliminated through optimization of the protein purification protocol. This optimization would most likely involve adjustments in the wash steps, such as increasing the concentration of imidazole (increasing its ability to compete with other proteins to bind to the beads) or a greater number of washes.

Chapter 4: Conclusion

Testing & Analysis

The next steps of this project will be to analyze the functionality of the synthetic antibody. To replace the monoclonal antibody used for PSA detection, the synthetic antibody must have comparable functionality. The previous results show the experimental and control synthetic antibodies were correctly expressed. Next, the synthetic antibodies will be tested to ensure proper L-DOPA incorporation in the experimental synthetic antibody. This will be done by measuring the K_d of both synthetic antibodies to PSA using ITC. The K_d value will indicate the binding affinity between the synthetic antibody and PSA. Since the incorporation of L-DOPA allows the experimental synthetic antibody to form a covalent bond to PSA, the binding strength will be much greater than the control synthetic antibody. The values will also be compared to the binding strength of the currently used monoclonal antibody for PSA detection, 1.1 ± 0.2 nM. The experimental synthetic antibody must have, at most, the same K_d value to be comparable to the anti-PSA monoclonal antibody.

Others metrics that will be tested and compared to those of the anti-PSA monoclonal antibody are specificity and sensitivity. It is crucial that the synthetic antibody is specific to PSA so that prostate cancer detection tests are precise and accurate. Both specificity and sensitivity will be tested using a Western blot. PSA samples will be titrated; then the specificity and sensitivity of the experimental synthetic antibody will be measured and compared to the anti-PSA monoclonal antibody. Currently, levels of PSA below 2.5 ng/mL are considered to be within the normal range.[7] Blood concentrations of PSA between 2.5 ng/mL and 4.0 ng/mL are an indication of prostate cancer.[7] The synthetic antibody must be able to clearly indicate the differences between these two ranges.

The design of the synthetic antibody predicts the specificity, sensitivity, and binding strength to be at least as effective, if not better than, the standards set by the currently used anti-PSA monoclonal antibody. However, these specifications must be tested to ensure incorporation of L-DOPA during expression and comparable functionality to currently used technologies.

Summary and Future Applications

This project focuses on creating a synthetic antibody for PSA measurement for prostate cancer detection. If testing follows theoretical values, this antibody can replace the currently used anti-PSA monoclonal antibody. As previously mentioned, antibodies are used extensively in research and medicine and prostate cancer detection is just one of numerous applications. This project evaluates the synthetic antibody design as a possibility to replace other monoclonal antibodies. Synthetic antibodies avoid many of the drawbacks of monoclonal antibodies, including batch-to-batch variability, ethical concerns, and an expensive and long production process.[1]

The results from this project show successful design and cloning of the synthetic antibody. In addition, it was successfully and produced in and purified from *E. coli*. Future testing would verify the incorporation and functionality of the synthetic antibody compared to the currently used monoclonal antibody for prostate cancer detection. This proof of concept is a model for future synthetic antibodies to be produced to essentially any other protein target. The advantages of this modular synthetic antibody design would make it desireable to researchers as well as clinicians and could become the new gold standard antibodies in industry. From the results obtained from this project, synthetic antibodies for diagnostic and therapeutic purposes are a promising technology, however, future testing is still required.

Chapter 5: Ethical Concerns

Monoclonal antibodies are currently widely used in both industry and research settings, and as a result, they are an established and well-known product due to their high specificity and binding strength. However, the production process of the monoclonal antibodies has two main ethical concerns including animal use as well as the expensive nature of the process leading to concerns regarding fair access of patients to medical advancements.

The use of live animals such as mice and rabbits as a hosts of production of monoclonal antibodies is an ethical concern because it requires invoking their immune system prior to the extraction of their blood. This process is harmful to the animals and raises concerns regarding animal abuse. Overall, the use of animals as hosts for this production method is life-threatening to the animals and usually leads to their death. In comparison, synthetic antibodies do not use animals such as mice and rabbits as hosts, instead *E. coli* bacteria are used.

Another ethical concern to monoclonal antibodies is their high cost as a result of their production process. The high cost of this technology can lead to minimal access to those who cannot afford to have access to it. Monoclonal antibodies are extensively used in research for medical advancements that could benefit patients, and the high-cost can deter researchers from using it and can lead to delays in new technology. Similarly, monoclonal antibodies are used in prostate cancer blood test and their high cost increases the cost of the test. In response, health insurance companies are reluctant to pay for it if the patient does not show any signs of prostate cancer. It is necessary for these tests to be done early regardless of symptoms because the lack of early detection contributes to many cancer patient deaths.

Chapter 6: Engineering Standards and Realistic Constraints

For all engineering designs, there are many considerations that must be taken into account prior to creating and implementing the design. The following factors will be discussed in terms of how the design and stakeholders that are involved will be impacted.

Economic

Like any other product, economic factors are necessary to take into consideration in terms of how much resources are needed to fund as well as implement the design and the overall cost of the product for it to be used by its intended audience. For this design, it was necessary to propose a budget for funding that was enough to carry out all of the experiments. Typical antibody production requires an immense amount of resources as well as animal facilities which can drive up costs and increase the amount of labor that is needed. Instead, the synthetic antibody production is performed at a smaller scale for testing and the host that is used for production, *E. coli*, is much cheaper and easier to scale up. Not only for the practicality of this project, but for future implementation, the synthetic antibody design described in this paper is a cheaper alternative than commercial monoclonal antibody production which will benefit researchers and patients who would have increased accessibility due to the lower cost of this technology.

Ethical

There are many ethical concerns with regards to who the stakeholders of this project will be. These stakeholders include who will be producing the synthetic antibody, who will be using the synthetic antibody such as researchers, academic institutions, and industry members, and lastly who would be benefitting or affected by the technology that is developed either from the indirect or direct usage of the synthetic antibody. As elaborated in Chapter 5, there are ethical concerns with regards to animal production as well as high costs that would limit the accessibility to the benefits of this technology. For this design, there are no animals used in the production process and the overall production is much cheaper than the commercial production methods currently used. As a result, this would eliminate ethical conflicts for those who are working to produce the

synthetic antibody, those who are concerned about animal mistreatment, and those who would be deterred from using antibodies due to their high costs.

A big-picture concern of this design could be if it is used a biological weapon due to the ease of the design. The screening method could be used to screen for a peptide that can target and cause disease or a physiological disorder. Then it could be implemented and produced using the production method described in this paper in a small laboratory setting without any regulations. It is necessary to be aware of this ethical concern; however, the screening process as a well as weaponizing of the synthetic antibody requires extensive knowledge that is not as straight-forward as the application of the design in this paper which is to target PSA, a well-known biomarker. However engineering as well as ethical standards should be upheld by those who are producing it to cause no harm to any of the stakeholders.

Social

Social responsibility towards all stakeholders is necessary to be taken into consideration for this project. Stakeholders who have limited access to the benefits of monoclonal antibody technology are those who are financially deterred as a result of high prices that would prevent insurance companies from subsidizing the cost of the application such as prostate cancer blood tests. A method that could allow for equal access to synthetic antibody is by reducing the overall cost of it. For example if the synthetic antibody is used for drug development, it would substantially reduce the cost of research due to failed attempts or multiple iterations that are required during development. This would lead to a reduced cost of the drug, and would allow for increased access to it for those who previously did not have access due to its high cost.

Manufacturability

For almost all designs, it is necessary that the design is validated to work on a smaller-scale before scaling it up in order to conserve resources. However, scaling-up can lead to other challenges and complications. One of the challenges of monoclonal antibody production is that scaling-up requires larger animal facilities, and there is also a potential for variations as a well as functionality in the final product due to it coming from multiple animal hosts.[14] This can lead to inconsistencies in quality control.[14]

This manufacturability concern can be overcome in the synthetic antibody design because the production platform is *E. coli* rather than multiple animal hosts. Scaling-up with E.coli using bioreactors has been documented in literature and has been well-studied so this design has a very real potential for scaling-up to the level of commercial usage with consistent quality.[18]

Health & Safety

It is necessary to ensure the health and safety of not only the members of this project but also other stakeholders who would be using this product. To produce the synthetic antibody, all members of the team were qualified to work in the lab by undergoing an exam to ensure that Santa Clara University Laboratory Safety Protocol and Procedures were understood. All members have also previously worked on protein production and purification and so proper lab techniques as well as handling of hazardous material have been established, and as a result, safety in the lab for all members can be ensured.

For the stakeholders who would be using the product it is necessary that it is appropriately used for its intended purpose. An example of this would be to use the prostate specific synthetic antibody for the purpose of detection of PSA rather than for detecting another biomarker. This will ensure safety and accuracy for the patients who would be receiving the results.

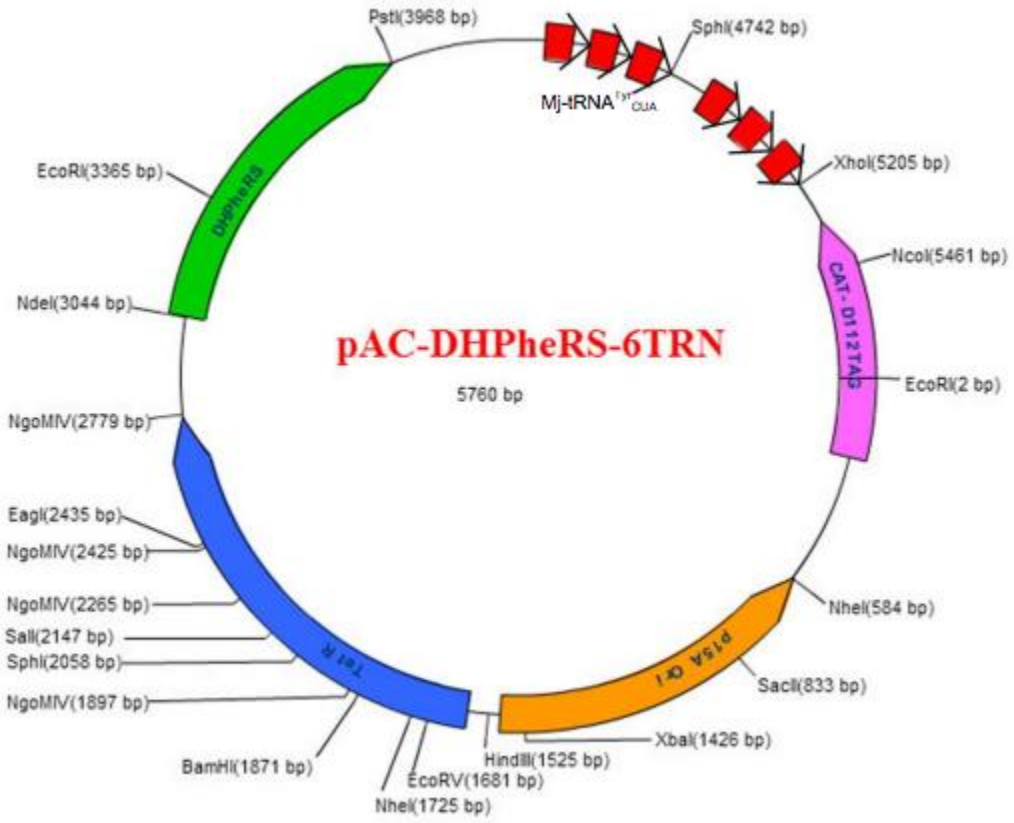
References

[1] Lipman, Neil S., et al. "Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources." *ILAR journal* 46.3 (2005): 258-268.
[2] Miersch, S., and S. S. Sidhu. "Synthetic antibodies: concepts, potential and practical considerations."*Methods* 57.4 (2012): 486-498.
[3] Umeda, Aiko, et al. "A versatile approach to transform low-affinity peptides into protein probes with cotranslationally expressed chemical cross-linker."*Analytical biochemistry* 405.1 (2010): 82-88.
[4] Byrne, Barry, Edwina Stack, Niamh Gilmartin, and Richard Oâ□™Kennedy. "Antibody-Based Sensors: Principles, Problems and Potential for Detection of Pathogens and Associated Toxins." *Sensors* 9.6 (2009): 4407-445.
[5] Leenaars, M., & Hendriksen, C. F. M. (2005). Critical Steps in the Production of Polyclonal a Monoclonal Antibodies: Evaluation and Recommendations. *ILAR Journal*, *46*(3), 269–279.
[6] Key Statistics for Prostate Cancer | Prostate Cancer Facts." *American Cancer Society*. Web. 09 May 2017.
[7] Katsamba, Phinikoula S., et al. "Kinetic analysis of a high-affinity antibody/antigen interaction performed by multiple Biacore users." *Analytical biochemistry* 352.2 (2006): 208-221.
[8] Wals, K., & Ovaa, H. (2014). Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Frontiers in Chemistry*, *2*, 15.
[9] Hallam, Trevor J., et al. "Antibody conjugates with unnatural amino acids." *Molecular pharmaceutics* 12.6 (2015): 1848-1862.
[10] Batjargal, Solongo, Christopher R. Walters, and E. James Petersson. "Inteins as Traceless Purification Tags for Unnatural Amino Acid Proteins." *Journal of the American Chemical Society* 137.5 (2015): 1734- 1737.
[11] Safar, Peter. "Alternatives to Animal Use in Research, Testing, and Education." *Journal of Neuropathology & Experimental Neurology* 48.6 (1989): 714-715.
[12] Binz, H. Kaspar, Patrick Amstutz, and Andreas Plückthun. "Engineering novel binding proteins from nonimmunoglobulin domains." *Nature biotechnology* 23.10 (2005): 1257-1268.
[13] Tombelli, Sara, Maria Minunni, and Marco Mascini. "Analytical applications of aptamers." *Biosensors and Bioelectronics* 20.12 (2005): 2424-2434.
[14] Therriault, Jon Henry, and Thomas Evans. *Computational Design of Synthetic Antibodies for Consumer Diagnostic Tests*. Thesis. Santa Clara University, 2016.
[15] Chung, C. T., Suzanne L. Niemela, and Roger H. Miller. "One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution." *Proceedings of the National Academy of Sciences* 86.7 (1989): 2172-2175.
[16] Zhang, Zhiwen, Brian A. C. Smith, Lei Wang, Ansgar Brock, Charles Cho, and Peter G. Schultz. "A New Strategy for the Site-Specific Modification of Proteinsin Vivoâ€ ."*Biochemistry* 42.22 (2003): 6735-746.
[17] Wang, Lei, Zhiwen Zhang, Ansgar Brock, and Peter G. Shultz. "Addition of the Keto Functional Group to the Genetic Code of Escherichiacoli." *PNAS*. 19 Nov. 2002.
[18] Rosano, Germán L., and Eduardo A. Ceccarelli. "Recombinant protein expression in Escherichia coli: advances and challenges." *Recombinant protein expression in microbial systems* 7 (2014).

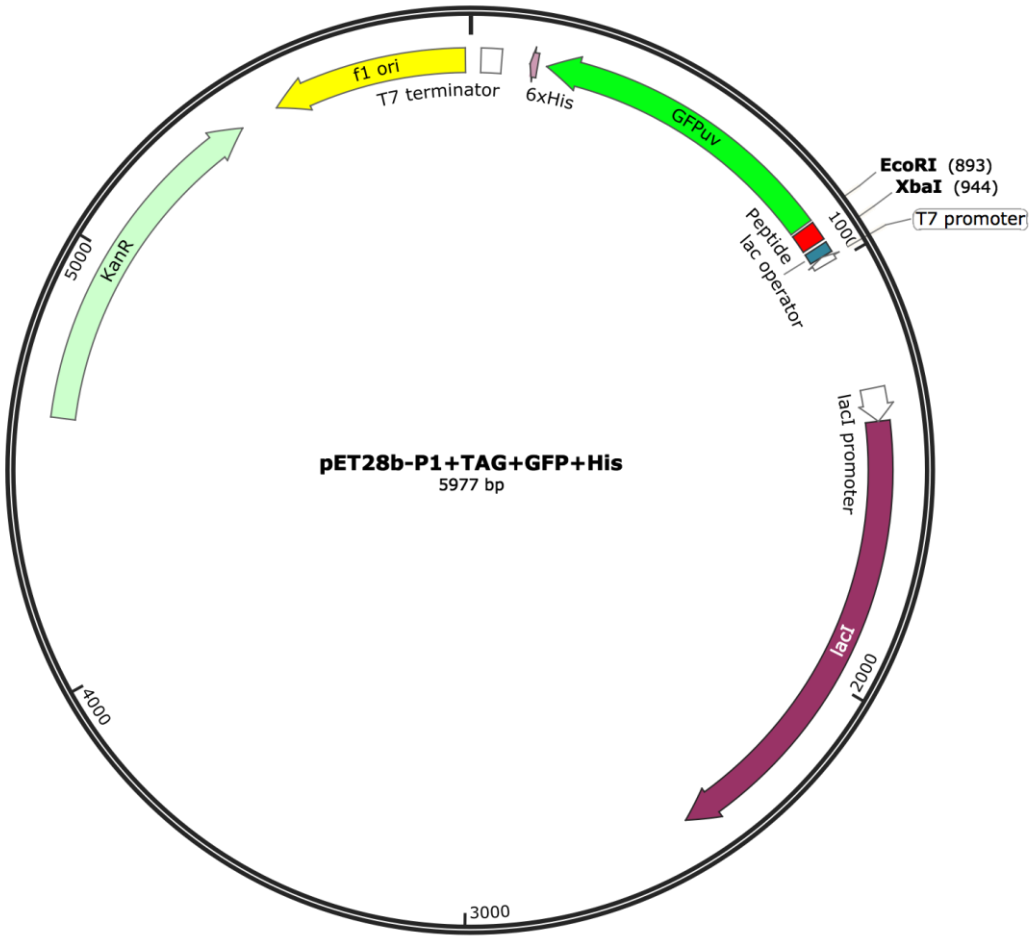
[19] Tsien, Roger Y. "The Green Fluorescent Protein." *Annu. Rev. Biochem. Annual Review of Biochemistry* 67.1 (1998): 509-44.
[20] Bornhorst, Joshua A., and Joseph J. Falke. "[16] Purification of proteins using polyhistidine affinity tags." *Methods in enzymology* 326 (2000): 245-254.
[21] Drummond, D. Allan, and Claus O. Wilke. "The evolutionary consequences of erroneous protein synthesis." *Nature Reviews Genetics* 10.10 (2009): 715-724.

Appendix

Appendix A: pAC DHPheRS-6TRN plasmid map



Appendix B: Experimental Vector



Appendix C. Experimental Vector: pET28-P1+GFP+His Sequencing Information

LOCUS GS61233-1 pET28-P1+GFP+His 5977 bp ds-DNA circular SYN 25-Apr-2017
DEFINITION .
ACCESSION .
VERSION .
KEYWORDS GS61233-1 pET28-P1+GFP+His
SOURCE synthetic DNA construct
ORGANISM synthetic DNA construct
REFERENCE 1 (bases 1 to 5977)
AUTHORS .
TITLE Direct Submission
FEATURES Location/Qualifiers
source 1..5977
/organism="synthetic DNA construct"
/mol_type="other DNA"
terminator 26..73
/note="T7 terminator"
/note="transcription terminator for bacteriophage T7 RNA
polymerase"
/note="color: #ffffff"
CDS complement(140..157)
/codon_start=1
/product="6xHis affinity tag"
/note="6xHis"
/note="color: #cc99b2"
/translation="HHHHHH"
gene 173..949
/note="P1+GFP+His"
/note="color: #a6acb3; direction: RIGHT"
CDS complement(179..892)
/codon_start=1
/product="GFP variant optimized for excitation by UV light"
/label="GFPuv"
/note="GFPuv"
/note="color: #05fd14"
/translation="MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLK

FICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDG
NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKA
NFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYK"

protein_bind 952..976

/bound_moiety="lac repressor encoded by lacI"

/note="lac operator"

/note="The lac repressor binds to the lac operator to
inhibit transcription in E. coli. This inhibition can be
relieved by adding lactose or
isopropyl-beta-D-thiogalactopyranoside (IPTG)."

/note="color: #31849b"

promoter complement(977..995)

/note="T7 promoter"

/note="promoter for bacteriophage T7 RNA polymerase"

/note="color: #ffffff; direction: LEFT"

promoter 1304..1381

/gene="lacI"

/note="lacI promoter"

/note="
"

/note="color: #ffffff; direction: RIGHT"

CDS 1382..2464

/codon_start=1

/gene="lacI"

/product="lac repressor"

/note="lacI"

/note="The lac repressor binds to the lac operator to
inhibit transcription in E. coli. This inhibition can be
relieved by adding lactose or
isopropyl-beta-D-thiogalactopyranoside (IPTG)."

/note="color: #993366"

/translation="MKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAEL
NYIPNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGV
EACKAAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSH
EDGTRLGVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSA
MSGFQQTMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSC

YIPPLTTIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPR
ALADSLMQLARQVSRLESGQ"
CDS 3273..3464
/codon_start=1
/gene="rop"
/product="Rop protein"
/note="rop"
/note="color: #993366"
/translation="MTKQEKTALNMARFIRSQTLTLLEKLNELDADEQADICESLHDHA
DELYRSCLARFGDDGENL"
rep_origin complement(3894..4482)
/direction=LEFT
/note="ori"
/note="high-copy-number colE1/pMB1/pBR322/pUC origin of
replication"
/note="color: #ffff00"
CDS 4604..5419
/codon_start=1
/gene="aph(3')-Ia"
/product="aminoglycoside phosphotransferase"
/note="KanR"
/note="confers resistance to kanamycin"
/note="color: #ccffcc"
/translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYG
KPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGK
TAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDA
SDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVGRVGI
ADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
rep_origin complement(5511..5966)
/direction=LEFT
/note="f1 ori"
/note="f1 bacteriophage origin of replication; arrow
indicates direction of (+) strand synthesis"
/note="color: #ffff00"
ORIGIN
1 atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa
61 ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt

121 tgttagcagc cggatctcag tggtggtggt ggtggtgctc gagtgcggcc gcaagctttt
181 tgtagagctc atccatgcca tgtgtaatcc cagcagcagt tacaaactca agaaggacca
241 tgtggtcacg cttttcgttg ggatctttcg aaagggcaga ttgtgtcgac aggtaatggt
301 tgtctggtaa aaggacaggg ccatcgccaa ttggagtatt ttgttgataa tggtctgcta
361 gttgaacgga tccatcttca atgttgtggc gaattttgaa gttagctttg attccattct
421 tttgtttgtc tgccgtgatg tatacattgt gtgagttata gttgtactcg agtttgtgtc
481 cgagaatgtt tccatcttct ttaaaatcaa tacctttaa ctcgatacga ttaacaaggg
541 tatcaccttc aaacttgact tcagcacgcg tcttgtagtt cccgtcatct ttgaaagata
601 tagtgcgttc ctgtacataa ccttcgggca tggcactctt gaaaaagtca tgccgtttca
661 tatgatccgg ataacgggaa aagcattgaa caccataaga gaaagtagtg acaagtgttg
721 gccatggaac aggtagtttt ccagtagtgc aaataaattt aagggtaagt tttccgtatg
781 ttgcatcacc ttcaccctct ccactgacag aaaatttgtg cccattaaca tcaccatcta
841 attcaacaag aattgggaca actccagtga aaagttcttc tcctttactc atgaattcct
901 agcaggtcca gcaatgatgt tcaatgcaat acgccacgca cattctagag gggaattgtt
961 atccgctcac aattccccta tagtgagtcg tattaatttc gcgggatcga gatctcgatc
1021 ctctacgccg gacgcatcgt ggccggcatc accggcgcca caggtgcggt tgctggcgcc
1081 tatatcgccg acatcaccga tggggaagat cgggctcgcc acttcgggct catgagcgct
1141 tgtttcggcg tgggtatggt ggcaggcccc gtggccgggg gactgttggg cgccatctcc
1201 ttgcatgcac cattccttgc ggcggcggtg ctcaacggcc tcaacctact actgggctgc
1261 ttcctaatgc aggagtcgca taagggagag cgtcgagatc ccggacacca tcgaatggcg
1321 caaaaccttt cgcggtatgg catgatagcg cccggaagag agtcaattca gggtggtgaa
1381 tgtgaaacca gtaacgttat acgatgtcgc agagtatgcc ggtgtctctt atcagaccgt
1441 ttcccgcgtg gtgaaccagg ccagccacgt ttctgcgaaa acgcgggaaa aagtggaagc
1501 ggcgatggcg gagctgaatt acattcccaa ccgcgtggca caacaactgg cgggcaaaca
1561 gtcgttgctg attggcgttg ccacctccag tctggccctg cacgcgccgt cgcaaattgt
1621 cgcggcgatt aaatctcgcg ccgatcaact gggtgccagc gtggtggtgt cgatggtaga
1681 acgaagcggc gtcgaagcct gtaaagcggc ggtgcacaat cttctcgcgc aacgcgtcag
1741 tgggctgatc attaactatc cgctggatga ccaggatgcc attgctgtgg aagctgcctg
1801 cactaatgtt ccggcgttat ttcttgatgt ctctgaccag acacccatca acagtattat
1861 tttctcccat gaagacggta cgcgactggg cgtggagcat ctggtcgcat tgggtcacca
1921 gcaaatcgcg ctgttagcgg gcccattaag ttctgtctcg gcgcgtctgc gtctggctgg
1981 ctggcataaa tatctcactc gcaatcaaat tcagccgata gcggaacggg aaggcgactg
2041 gagtgccatg tccggttttc aacaaaccat gcaaatgctg aatgagggca tcgttcccac
2101 tgcgatgctg gttgccaacg atcagatggc gctgggcgca atgcgcgcca ttaccgagtc
2161 cgggctgcgc gttggtgcgg atatctcggt agtgggatac gacgataccg aagacagctc
2221 atgttatatc ccgccgttaa ccaccatcaa acaggatttt cgcctgctgg ggcaaaccag
2281 cgtggaccgc ttgctgcaac tctctcaggg ccaggcggtg aagggcaatc agctgttgcc

2341 cgtctcactg gtgaaaagaa aaaccaccct ggcgcccaat acgcaaaccg cctctccccg
2401 cgcgttggcc gattcattaa tgcagctggc acgacaggtt tcccgactgg aaagcgggca
2461 gtgagcgcaa cgcaattaat gtaagttagc tcactcatta ggcaccggga tctcgaccga
2521 tgcccttgag agccttcaac ccagtcagct ccttccggtg ggcgcggggc atgactatcg
2581 tcgccgcact tatgactgtc ttctttatca tgcaactcgt aggacaggtg ccggcagcgc
2641 tctgggtcat tttcggcgag gaccgctttc gctggagcgc gacgatgatc ggcctgtcgc
2701 ttgcggtatt cggaatcttg cacgccctcg ctcaagcctt cgtcactggt cccgccacca
2761 aacgtttcgg cgagaagcag gccattatcg ccggcatggc ggccccacgg gtgcgcatga
2821 tcgtgctcct gtcgttgagg acccggctag gctggcgggg ttgccttact ggttagcaga
2881 atgaatcacc gatacgcgag cgaacgtgaa gcgactgctg ctgcaaaacg tctgcgacct
2941 gagcaacaac atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa acgcggaagt
3001 cagcgccctg caccattatg ttccggatct gcatcgcagg atgctgctgg ctaccctgtg
3061 gaacacctac atctgtatta acgaagcgct ggcattgacc ctgagtgatt tttctctggt
3121 cccgccgcat ccataccgcc agttgtttac cctcacaacg ttccagtaac cgggcatgtt
3181 catcatcagt aacccgtatc gtgagcatcc tctctcgttt catcggtatc attaccccca
3241 tgaacagaaa tccccccttac acggaggcat cagtgaccaa acaggaaaaa accgccctta
3301 acatggcccg ctttatcaga agccagacat taacgcttct ggagaaactc aacgagctgg
3361 acgcggatga acaggcagac atctgtgaat cgcttcacga ccacgctgat gagctttacc
3421 gcagctgcct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg
3481 agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt cagggcgcgt
3541 cagcgggtgt tggcgggtgt cggggcgcag ccatgaccca gtcacgtagc gatagcggag
3601 tgtatactgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatatg
3661 cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct
3721 tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac
3781 tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga
3841 gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat
3901 aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac
3961 ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct
4021 gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg
4081 ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg
4141 ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt
4201 cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg
4261 attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac
4321 ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga
4381 aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt
4441 gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt
4501 tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaac

4561 aataaaactg tctgcttaca taaacagtaa tacaaggggt gttatgagcc atattcaacg
4621 ggaaacgtct tgctctaggc cgcgattaaa ttccaacatg gatgctgatt tatatgggta
4681 taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa
4741 gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca atgatgttac
4801 agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca
4861 ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg ggaaaacagc
4921 attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt
4981 gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt
5041 atttcgtctc gctcaggcgc aatcacgaat gaataacggt ttggttgatg cgagtgattt
5101 tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataaacttt
5161 gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt
5221 tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata
5281 ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg
5341 gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat
5401 gctcgatgag tttttctaag aattaattca tgagcggata catatttgaa tgtatttaga
5461 aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct aaattgtaag
5521 cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca
5581 ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttgag
5641 tgttgttcca gtttggaaca agagtccact attaaagaac gtggactcca acgtcaaagg
5701 gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct aatcaagttt
5761 tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag
5821 agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc
5881 gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc
5941 gcttaatgcg ccgctacagg gcgcgtccca ttcgcca
//

Appendix D. Control Vector: pET28-P1+GFP+His Sequencing Information

LOCUS GS61233-2 pET28-control+GFP+His 5977 bp ds-DNA circular SYN 25-Apr-2017
DEFINITION .
ACCESSION .
VERSION .
KEYWORDS GS61233-2 pET28-control+GFP+His
SOURCE synthetic DNA construct
ORGANISM synthetic DNA construct
REFERENCE 1 (bases 1 to 5977)
AUTHORS .
TITLE Direct Submission
FEATURES Location/Qualifiers

source 1..5977
/organism="synthetic DNA construct"
/mol_type="other DNA"
terminator 26..73
/note="T7 terminator"
/note="transcription terminator for bacteriophage T7 RNA
polymerase"
/note="color: #ffffff"
CDS complement(140..157)
/codon_start=1
/product="6xHis affinity tag"
/note="6xHis"
/note="color: #cc99b2"
/translation="HHHHHH"
misc_feature 173..949
/note="control+GFP+His"
/note="color: #a6acb3"
CDS complement(179..892)
/codon_start=1
/product="GFP variant optimized for excitation by UV light"
/label="GFPuv"
/note="GFPuv"
/note="color: #05fd14"
/translation="MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLK
FICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDG
NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKA
NFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYK"
protein_bind 952..976
/bound_moiety="lac repressor encoded by lacI"
/note="lac operator"
/note="The lac repressor binds to the lac operator to
inhibit transcription in E. coli. This inhibition can be
relieved by adding lactose or
isopropyl-beta-D-thiogalactopyranoside (IPTG)."
/note="color: #31849b"
promoter complement(977..995)

/note="T7 promoter"
/note="promoter for bacteriophage T7 RNA polymerase"
/note="color: #ffffff; direction: LEFT"
promoter 1304..1381
/gene="lacI"
/note="lacI promoter"
/note="
"
/note="color: #ffffff; direction: RIGHT"
CDS 1382..2464
/codon_start=1
/gene="lacI"
/product="lac repressor"
/note="lacI"
/note="The lac repressor binds to the lac operator to
inhibit transcription in E. coli. This inhibition can be
relieved by adding lactose or
isopropyl-beta-D-thiogalactopyranoside (IPTG)."
/note="color: #993366"
/translation="MKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAEL
NYIPNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGV
EACKAAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSH
EDGTRLGVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSA
MSGFQQTMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSC
YIPPLTTIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPR
ALADSLMQLARQVSRLESGQ"
CDS 3273..3464
/codon_start=1
/gene="rop"
/product="Rop protein"
/note="rop"
/note="color: #993366"
/translation="MTKQEKTALNMARFIRSQTLTLLEKLNELDADEQADICESLHDHA
DELYRSCLARFGDDGENL"
rep_origin complement(3894..4482)
/direction=LEFT
/note="ori"

/note="high-copy-number colE1/pMB1/pBR322/pUC origin of replication"
/note="color: #ffff00"

CDS 4604..5419
/codon_start=1
/gene="aph(3')-Ia"
/product="aminoglycoside phosphotransferase"
/note="KanR"
/note="confers resistance to kanamycin"
/note="color: #ccffcc"
/translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLYG
KPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIPGK
TAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGLVDA
SDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVGRVGI
ADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"

rep_origin complement(5511..5966)
/direction=LEFT
/note="f1 ori"
/note="f1 bacteriophage origin of replication; arrow indicates direction of (+) strand synthesis"
/note="color: #ffff00"

ORIGIN

1 atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa
61 ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt
121 tgttagcagc cggatctcag tggtggtggt ggtggtgctc gagtgcggcc gcaagctttt
181 tgtagagctc atccatgcca tgtgtaatcc cagcagcagt tacaaactca agaaggacca
241 tgtggtcacg cttttcgttg ggatctttcg aaagggcaga ttgtgtcgac aggtaatggt
301 tgtctggtaa aaggacaggg ccatcgccaa ttggagtatt ttgttgataa tggtctgcta
361 gttgaacgga tccatcttca atgttgtggc gaattttgaa gttagctttg attccattct
421 tttgtttgtc tgccgtgatg tatacattgt gtgagttata gttgtactcg agtttgtgtc
481 cgagaatgtt tccatcttct ttaaaatcaa taccttttaa ctcgatacga ttaacaaggg
541 tatcaccttc aaacttgact tcagcacgcg tcttgtagtt cccgtcatct ttgaaagata
601 tagtgcgttc ctgtacataa ccttcgggca tggcactctt gaaaaagtca tgccgtttca
661 tatgatccgg ataacgggaa aagcattgaa caccataaga gaaagtagtg acaagtgttg
721 gccatggaac aggtagtttt ccagtagtgc aaataaattt aagggtaagt tttccgtatg
781 ttgcatcacc ttcaccctct ccactgacag aaaatttgtg cccattaaca tcaccatcta
841 attcaacaag aattgggaca actccagtga aaagttcttc tcctttactc atgaattccg

901 cgcaggtcca gcaatgatgt tcaatgcaat acgccacgca cattctagag gggaattgtt
961 atccgctcac aattccccta tagtgagtcg tattaatttc gcgggatcga gatctcgatc
1021 ctctacgccg gacgcatcgt ggccggcatc accggcgcca caggtgcggt tgctggcgcc
1081 tatatcgccg acatcaccga tggggaagat cgggctcgcc acttcgggct catgagcgct
1141 tgtttcggcg tgggtatggt ggcaggcccc gtggccgggg gactgttggg cgccatctcc
1201 ttgcatgcac cattccttgc ggcggcggtg ctcaacggcc tcaacctact actgggctgc
1261 ttcctaatgc aggagtcgca taagggagag cgtcgagatc ccggacacca tcgaatggcg
1321 caaaaccttt cgcggtatgg catgatagcg cccggaagag agtcaattca gggtggtgaa
1381 tgtgaaacca gtaacgttat acgatgtcgc agagtatgcc ggtgtctctt atcagaccgt
1441 ttcccgcgtg gtgaaccagg ccagccacgt ttctgcgaaa acgcgggaaa aagtggaagc
1501 ggcgatggcg gagctgaatt acattcccaa ccgcgtggca caacaactgg cgggcaaaca
1561 gtcgttgctg attggcgttg ccacctccag tctggccctg cacgcgccgt cgcaaattgt
1621 cgcggcgatt aaatctcgcg ccgatcaact gggtgccagc gtggtggtgt cgatggtaga
1681 acgaagcggc gtcgaagcct gtaaagcggc ggtgcacaat cttctcgcgc aacgcgtcag
1741 tgggctgatc attaactatc cgctggatga ccaggatgcc attgctgtgg aagctgcctg
1801 cactaatgtt ccggcgttat ttcttgatgt ctctgaccag acacccatca acagtattat
1861 tttctcccat gaagacggta cgcgactggg cgtggagcat ctggtcgcat tgggtcacca
1921 gcaaatcgcg ctgttagcgg gcccattaag ttctgtctcg gcgcgtctgc gtctggctgg
1981 ctggcataaa tatctcactc gcaatcaaat tcagccgata gcggaacggg aaggcgactg
2041 gagtgccatg tccggttttc aacaaaccat gcaaatgctg aatgagggca tcgttcccac
2101 tgcgatgctg gttgccaacg atcagatggc gctgggcgca atgcgcgcca ttaccgagtc
2161 cgggctgcgc gttggtgcgg atatctcggt agtgggatac gacgataccg aagacagctc
2221 atgttatatc ccgccgttaa ccaccatcaa acaggatttt cgcctgctgg ggcaaaccag
2281 cgtggaccgc ttgctgcaac tctctcaggg ccaggcggtg aagggcaatc agctgttgcc
2341 cgtctcactg gtgaaaagaa aaaccaccct ggcgcccaat acgcaaaccg cctctccccg
2401 cgcgttggcc gattcattaa tgcagctggc acgacaggtt tcccgactgg aaagcgggca
2461 gtgagcgcaa cgcaattaat gtaagttagc tcactcatta ggcaccggga tctcgaccga
2521 tgcccttgag agccttcaac ccagtcagct ccttccggtg ggcgcggggc atgactatcg
2581 tcgccgcact tatgactgtc ttctttatca tgcaactcgt aggacaggtg ccggcagcgc
2641 tctgggtcat tttcggcgag gaccgctttc gctggagcgc gacgatgatc ggcctgtcgc
2701 ttgcggtatt cggaatcttg cacgccctcg ctcaagcctt cgtcactggt cccgccacca
2761 aacgtttcgg cgagaagcag gccattatcg ccggcatggc ggccccacgg gtgcgcatga
2821 tcgtgctcct gtcgttgagg acccggctag gctggcgggg ttgccttact ggttagcaga
2881 atgaatcacc gatacgcgag cgaacgtgaa gcgactgctg ctgcaaaacg tctgcgacct
2941 gagcaacaac atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa acgcggaagt
3001 cagcgccctg caccattatg ttccggatct gcatcgcagg atgctgctgg ctaccctgtg
3061 gaacacctac atctgtatta acgaagcgct ggcattgacc ctgagtgatt tttctctggt

3121 cccgccgcat ccataccgcc agttgtttac cctcacaacg ttccagtaac cgggcatgtt
3181 catcatcagt aacccgtatc gtgagcatcc tctctcgttt catcggtatc attaccccca
3241 tgaacagaaa tccccccttac acggaggcat cagtgaccaa acaggaaaaa accgccctta
3301 acatggcccg ctttatcaga agccagacat taacgcttct ggagaaactc aacgagctgg
3361 acgcggatga acaggcagac atctgtgaat cgcttcacga ccacgctgat gagctttacc
3421 gcagctgcct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg
3481 agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt cagggcgcgt
3541 cagcgggtgt tggcgggtgt cggggcgcag ccatgaccca gtcacgtagc gatagcggag
3601 tgtatactgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatatg
3661 cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct
3721 tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac
3781 tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga
3841 gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat
3901 aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac
3961 ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct
4021 gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg
4081 ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg
4141 ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt
4201 cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg
4261 attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac
4321 ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga
4381 aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt
4441 gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt
4501 tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaac
4561 aataaaactg tctgcttaca taaacagtaa tacaaggggt gttatgagcc atattcaacg
4621 ggaaacgtct tgctctaggc cgcgattaaa ttccaacatg gatgctgatt tatatgggta
4681 taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa
4741 gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca atgatgttac
4801 agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca
4861 ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg ggaaaacagc
4921 attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt
4981 gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt
5041 atttcgtctc gctcaggcgc aatcacgaat gaataacggt ttggttgatg cgagtgattt
5101 tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataaactttt
5161 gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt
5221 tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata
5281 ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg

5341 gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat

5401 gctcgatgag tttttctaag aattaattca tgagcggata catatttgaa tgtatttaga

5461 aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct aaattgtaag

5521 cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca

5581 ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttgag

5641 tgttgttcca gtttggaaca agagtccact attaaagaac gtggactcca acgtcaaagg

5701 gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct aatcaagttt

5761 tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag

5821 agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc

5881 gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc

5941 gcttaatgcg ccgctacagg gcgcgtccca ttcgcca

//

Appendix E. Cloning Reactions

Table 1. Sample Klenow Extension Reaction.

	Reaction 1 (P1+TAG)	**Reaction 2 (P1+GCG)**
Klenow (NEB M0210S 2015)	0.5 µL	0.5 µL
Oligo 1 (F)	1 µL (1.36 µg)	1 µL (1.36 µg)
Oligo 2 (R)	1 µL (1.36 µg)	1 µL (1.36 µg)
10X NEB 2 Buffer (2016)	2.0 µL	2.0 µL
dNTPs (Promega 12/17)	6.6 µL	6.6 µL
ddH_2O	8.85 µL	8.85 µL
Total Reaction Volume	20 µL	20 µL
Add EDTA (0.5M stock)	0.4 µL	0.4 µL

*Incubate reaction for 15 mins at 25℃

**End reaction by adding 0.4 µL of 0.5M EDTA and heating at 75℃ for 20 minutes

Table 2. Sample Insert Double Digestion Reaction.

DNA (insert, 1µg)	10 µL
10X NEB Buffer	5 µL
Restriction Enzyme 1: EcoRI-HF	1 µL
Restriction Enzyme 2: XbaI	1 µL
ddH_2O	33 µL
Total Reaction Volume	50 µL

*Incubate reaction for 10 hours at 37℃.

Table 3. Sample pET28b-GFP Double Digestion Reaction.

DNA (pET28b-GFP, 1 µg)	10 µL
Restriction Enzyme 1: EcoRI-HF	1 µL
Restriction Enzyme 2: XbaI	1 µL
10X NEB Buffer	5 µL
ddH_2O	33 µL
Total Reaction Volume	50 µL

*Incubate reaction for 4 hours at 37℃.

Table 4. Sample Calf Intestinal Alkaline Phosphatase (CIP) Reaction for pET28b-GFP.

DNA (pET28b-GFP, 1 µg)	15 µL
10X Cutsmart Buffer	2 µL
CIP	0.5 µ
ddH_2O	2.5 µL
Total Reaction Volume	20 µL

*Incubate reaction for 50 minutes at 37℃.

Table 5. Sample pET28b-GFP + Oligonucleotide Ligation Reaction.

pET28b-GFP (47.5 ng)	5 µL
Oligonucleotide (4.82 ng)	3.2 µL
T4 DNA Ligase	1 µL
T4 DNA Ligase Buffer	1 µL
Total Reaction Volume	10.2 µL

*Incubate reaction for 4 hours at 23℃

** Reaction shown is 1:10 pET28b-GFP:oligonucleotide. 1:1 and 1:2 reactions were also performed.

Appendix F. Project Expenses

Product	**Cost**
Sequencing	$156
L-arabinose	$61
Oligonucleotides	$118.12
L-DOPA	$51.90
Cloning	$535
Total	**$922.02**

Appendix G. Project Timeline

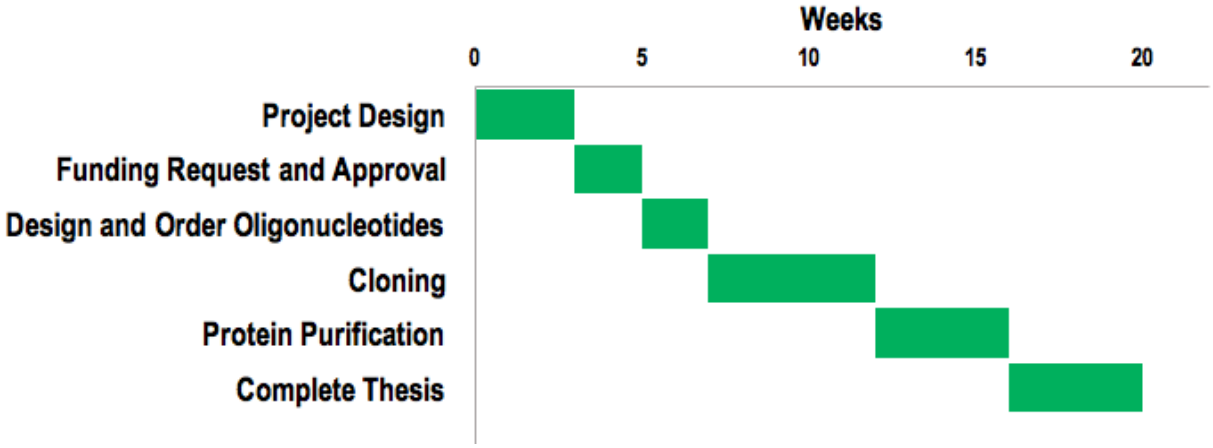


Figure 4. Project Timeline.