

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Designing Recombinant Adhesion Molecule Constructs with Fluorescent Proteins and Springy Linkers for Targeting Therapeutic Nanoparticles

Permalink

<https://escholarship.org/uc/item/3zq361b7>

Author

Kumar, Shreya Ramesh

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Designing Recombinant Adhesion Molecule Constructs with Fluorescent Proteins and
Springy Linkers for Targeting Therapeutic Nanoparticles

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Biomedical Engineering

by

Shreya Ramesh Kumar

Thesis Committee:
Assistant Professor Jered Haun, Chair
Associate Professor James Brody
Assistant Professor Chang Liu

2018

DEDICATION

To

my family and friends

in recognition of their love and support

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vii
ACKNOWLEDGMENTS	viii
ABSTRACT OF THE THESIS	ix
INTRODUCTION	1
CHAPTER 1: Fluorescent protein expression of mCherry and eGFP in yeast	6
1.1 Background	6
1.2 Materials and Methods	8
1.2.1 Chemical reagents	8
1.2.2 Strains and Media	8
1.2.3 Transformation	8
1.2.4 Protein Purification and Characterization	8
CHAPTER 2: Construction of recombinant protein with mOx GFP	11
2.1 Background	11
2.2 Materials and Methods	12
2.2.1 Reagents, Materials and Kits	12
2.2.2 Strains and Media	12
2.2.3 Plasmid Design	12
2.2.4 Digestion, Ligation, and Transformation	13
2.2.5 Sequencing Results	16
2.3 Codon optimization of mOx GFP for yeast cells	17
2.3.1 Background	17
2.3.2 Codon Optimization of mOx GFP	18
2.3.3 Plasmid Design	20
CHAPTER 3: Construction of recombinant protein with Flagelliform and GGS Linkers	22
3.1 Background	22
3.2 Materials and Methods	23
3.2.1 Reagents, Materials and Kits	23
3.2.2 Strains and Media	23
3.2.3 Plasmid Design	24
3.2.4 Digestion, Ligation, and Transformation	25
3.2.5 Sequencing Results	28
FUTURE DIRECTION	30

REFERENCES	31
APPENDIX A: Gene Cloning and Propagation	33
APPENDIX B: Yeast Culture and Recombinant Protein Production	39

LIST OF FIGURES

	Page
Figure 1 Medical Application of Nanotechnology	1
Figure 2 Nanoparticles with numerous targeting ligands can bind to the surface cells multivalency with high receptor density	2
Figure 3 Passive and active targeting of nanocarriers	3
Figure 4 Model of targeted drug delivery system	4
Figure 5 Construct of scFv with peptide linkers in pRS expression system	4
Figure 6 Structure of green fluorescent protein	7
Figure 7 Structure of mCherry	7
Figure 8 Fluorescence microscopy image of eGFP in fitc at 20x	9
Figure 9 Fluorescence microscopy of mCherry TRITC at 20X	9
Figure 10 SDS PAGE results of prs-4420-Avitag	10
Figure 11 Western blot result of prs-4420-Avitag	10
Figure 12 Tangled misfolded dark EGFP and moxGFP solution	11
Figure 13 Original pRS vector sequence, with 4420-mCherry-S6-His6-cymc with Mlu1 and Not1 highlighted in red	13
Figure 14 Digested pRS-4420-tags1 (lane 1 & 2) pRS-4420-mCherry-S6-His6-cymc (lane 3 & 4) vector with 1kb ladder	14
Figure 15 Transformed mox GFP in S6 and tags1 vectors	16
Figure 16 Genewiz confirmation for Tags1	16
Figure 17 Genewiz confirmation for S6	17
Figure 18 Translate tool at ExPasy	19
Figure 19 Open reading frame at ExPasy	19

Figure 20 Choosing <i>S. cerevisiae</i> for yeast codon optimization	20
Figure 21 Sequence obtained after optimization	20
Figure 22 Molecular models for relaxed and extended flagelliform protein sequences from spider capture silk	22
Figure 23 Original pRS vector sequence, with 4420-mCherry-S6-His6-cmyc with Mlu1 and Bsw1 highlighted in red	24
Figure 24 Digested pRS-4420-mCherry-S6-His6-cmyc vector with 1kb ladder	25
Figure 25 Transformed flagelliform50 and ggs50 in S6 vector	27
Figure 26 Genewiz confirmation of flag50	28
Figure 27 Genewiz Confirmation for GGS50	29

LIST OF TABLES

	Page
Table 1 Concentration of inserts and backbone -digestion (mOx GFP)	15
Table 2 mOx GFP Ligation set up	15
Table 3 Digestion concentrations for Flag50 and GGS50	26
Table 4 Ligation setup for Flag50 and GGS50	26
Table 5 List of reagents	33
Table 6 PCR setup	34
Table 7 PCR stages	35
Table 8 Digestion set up	36
Table 9 Ligation set up	36
Table 10 List of reagents for yeast transformation	39

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my committee chair, Assistant Professor Jered Haun, for his continuous guidance and support via his immense knowledge, time and patience shown in my master's study and related research. His guidance and motivation for the past 1.5 years, has helped me in my time as a graduate researcher and in writing this thesis.

I would like to thank my committee members, Associate Professor James Brody and Assistant Professor Chang Liu, for their support and insight towards my thesis, and for the use of their equipment, without which my thesis would not have been possible.

The team I worked with in lab have a special place in my heart. A big thank you to Yuanqing Xue, Phinix Dominh, Esther Yu-Tin Chen, and Dr. Maha Rahim, for helping me take my first steps into research. I would also like to thank Thinkh Nguyen and Louis Mejia for walking with me by my side.

Finally, a big thank you to all my lab mates for making my master's journey fun and memorable.

ABSTRACT OF THE THESIS

Designing Recombinant Adhesion Molecule Constructs with Fluorescent Proteins and
Springy Linkers for Targeting Therapeutic Nanoparticles

By

Shreya Ramesh Kumar

Master of Biomedical Engineering

University of California, Irvine, 2018

Assistant Professor Jered Haun, Chair

Nanoparticle mediated targeted delivery has been widely studied as it provides nanoparticle-cell surface interaction and helps shift biodistribution to diseased cells, thereby overcoming the disadvantages of conventional drug delivery methods. Fluorescent proteins have been used as markers to detect protein-protein interaction and can be used to track the movements of other proteins. The first part of this study aimed to produce fluorescent proteins at high yield, so we studied mOx GFP and codon optimized it for yeast. The second part of this study dealt with nanospringy linkers, specifically flagelliform, and its construction. Springy linkers are designed to reduce mechanical forces on bonds produced by the nanoparticle, which will stabilize adhesion. In previous studies, springy linkers with fewer amino acids base lengths were already constructed, and we investigated if we could construct them with high amino acid count. Springy linkers will pave the way to experiments to validate our theory.

Introduction

Nanomedicine and the use of nanobiotechnology for therapeutic use has been in cancer therapy. This technology is the need of the hour as cancer is the second leading cause of death in the United States and all over the world. Traditional treatments like chemotherapy and radiation therapy contain high levels of toxicity – which are designed to kill cancer cells. As explained below, it has its drawbacks, so nanoparticles are being used as they can introduce site-specific targeting bearing potency of current therapeutic agents¹.

Conventional drug admission routes include injection or oral doses, which reduce the efficiency of drugs. They can affect healthy tissues which can cause unpleasant side effects and this basically reduces the concentration and amount of drug available for the unhealthy tissues/cells². Targeted drug delivery aims to overcome such disadvantages by increasing the concentration of the medication in some parts of the body, relative to others. This way, the drug has prolonged and targeted interaction with the diseased tissue, and it can be protected against degradation if it is incorporated into a nanoparticle.

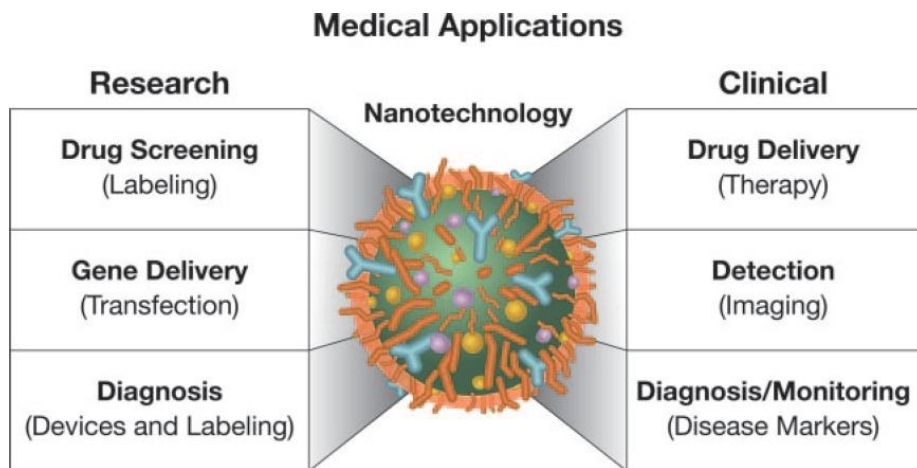


Figure 1: Medical Application of Nanotechnology²

For anticancer drugs to act as efficient drugs in cancer treatment, they need to be able to penetrate the barriers in the body with minimal loss in volume or activity and reach the

specific tumor tissues. Once they reach the site, they need to have the ability to kill tumorous cells only, without affecting normal cells with controlled release of the drug in the active form³. Improvements in patient survival and quality of life can be achieved with increasing intracellular drug concentration and reducing dose-limiting toxicities at the same time, in addition to the strategies explained above.

Small-molecule drugs, peptides, proteins and nucleic acids, all comprise therapeutic entities, which are considered as nanoparticle therapeutics. If nanoparticles contain moieties like small molecules, peptides or proteins, they can target cancer cells by binding with the cancer cell surface receptor proteins, like transferrin receptors (which are increased in number on cancer cells)⁴. There are active and passive targeting techniques by which intracellular concentration of drugs in the cancer cells can be increased while keeping the healthy cells at bay⁵.

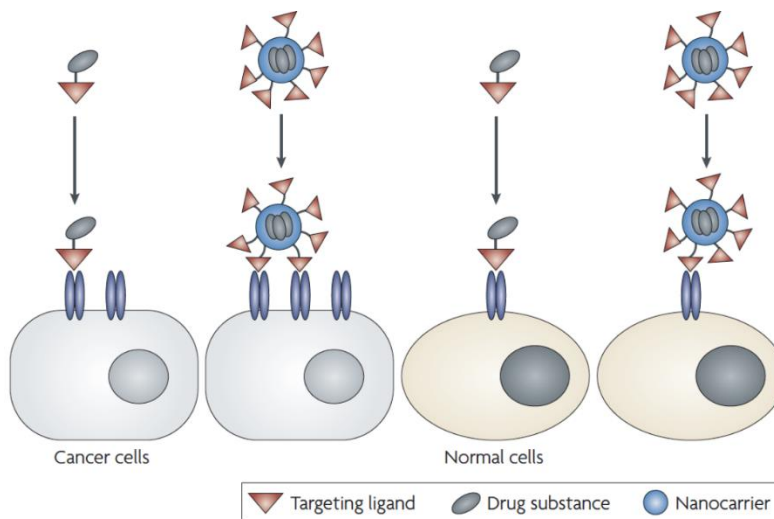


Figure 2: Nanoparticles with numerous targeting ligands can bind to the surface cells multivalency with high receptor density. When the surface density of the receptor is low on normal cells, the molecular conjugation of a single targeting ligand and a targeted nanoparticle can compete equally since only one ligand-receptor may occur⁵.

Targeted drug delivery can be done in one of two ways. Passive targeting transports the nanocarriers through leaky tumor capillary fenestrations into the tumor interstitium and

cells by passive diffusion. But this has a drawback – passive targeting is dependent on the degree of tumor vascularization and angiogenesis. It gives rise to challenges in drug delivery with regard to tumor size and anatomical sites. Active targeting takes place when the receptors at the target site can bind to the surface of nanocarriers which have the appropriate targeting ligands. With this, there isn't any variation of the size of nanocarriers. Targeting ligands can be monoclonal antibodies (mAbs) and antibody fragments, or any other antigens associates with cancer⁶.

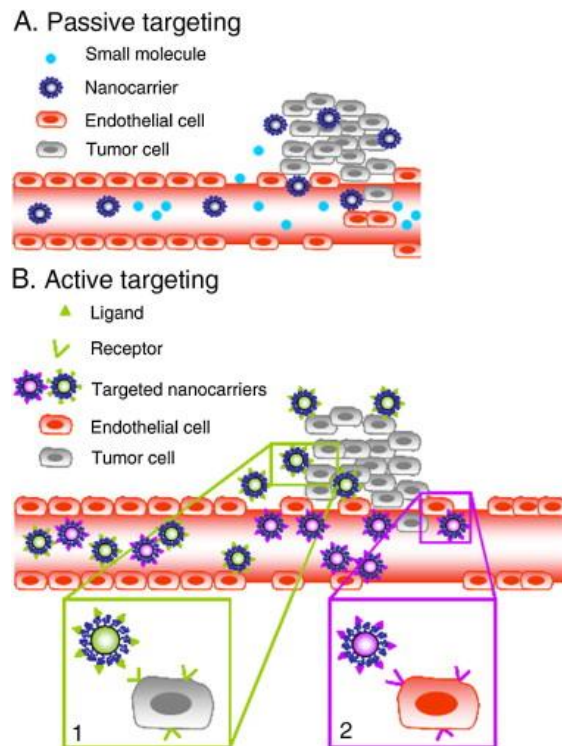


Figure 3: A. Passive targeting of nanocarriers. (1) Nanocarriers reach tumors selectively through the leaky vasculature surrounding the tumors. (2) Schematic representation of the influence of the size for retention in the tumor tissue. Drugs alone diffuse freely in and out the tumor blood vessels because of their small size and thus their effective concentrations in the tumor decrease rapidly. By contrast, drug-loaded nanocarriers cannot diffuse back into the blood stream because of their large size, resulting in progressive accumulation: the EPR effect. B. Active targeting strategies. Ligands grafted at the surface of nanocarriers bind to receptors (over)expressed by (1) cancer cells or (2) angiogenic endothelial cells⁶.

Our lab has established a platform of 4420 scFv fused fluorescent protein (mCherry and eGFP), biotin acceptor protein (AviTag), six-histidine tag (6His), c-myc epitope in expression.

The recombinant antibodies as targeting receptors were biotinylated and then were functionalized directly onto avidin-coated nanoparticles. Molecular binding interactions are measured in flow chamber under fluid flow⁷.

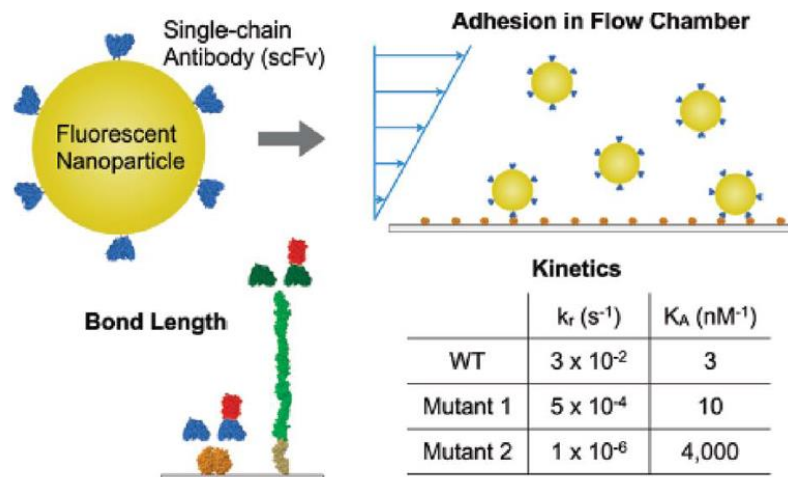


Figure 4: Model of targeted drug delivery system⁷

In this study, we first investigated the production of proteins (mCherry and eGFP) in pRS expression vectors produced in previous studies. We then developed targeting receptor by introducing mOx GFP and yeast codon optimized mOx GFP into pRS314 expression vector. We also investigated the use of spider silk springy linkers and developed another model for targeting receptor into pRS314 expression vector, which can be used for future studies.

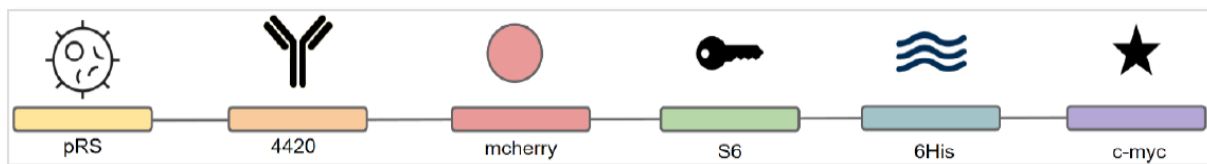


Figure 5: Construct of scFv with peptide linkers in pRS expression system

Our construct includes a GAL1-10 as promoter, synthetic prepro leader as sequence to direct the protein, pRS314 vector, 4420 as antibody region, mCherry as fluorophore, S6 tag as the orthogonal binding site for attachment to a nanoparticle via an enzymatic reaction, 6His for protein purification and c-myc as epitope tag. mCherry was replaced by eGFP in a previous study, along with replacing S6 tag with a polytag linker. The first part of our study replaced both fluorescent proteins with mOx GFP and yeast codon optimized mOx GFP. The second part of our study introduced Flag50 and GGS50 as linkers between 4420 and mCherry.

Chapter 1: Fluorescent protein expression of mCherry and EGFP in yeast

1.1 Background

To better understand cellular pathways, protein-protein interaction within biological membranes, molecular biologists have turned to the use of fluorescent proteins. To produce novel targeted therapeutic agents, fluorescent tags are added into proteins, which can be placed at well-defined positions. Fluorescent proteins can be genetically encoded with the targeting protein (or fusion protein) and can label proteins in live cells. Apart from this, they are very stable in nature, and because of the protein's barrel structure, fluorescent proteins can track the localization and movement of fusion proteins. This can give scientists great insights about proteins in their native environment⁸.

Due to structural features and good biocompatibility, mesoporous silica nanoparticles are being utilized as a multifunctional carrier in the drug delivery systems. To aid in treatment for early stage cancer, the combination of mesoporous silica nanoparticles and ⁶⁴Cu might be a promising theranostic device⁹. After the process of PEGylation, where polyethylene glycol (covalent and non-covalent) attaches to macrostructures like a drug, it may be difficult to understand surface properties of silica nanoparticle. This is especially important when we need to evaluate protein-protein interaction on the surface of silica nanoparticles⁹.

The discovery of green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* was one of the biggest breakthroughs in molecular biology¹⁰. It could be fused with almost any protein of interest to study protein movement and interaction in living cells. In addition to that, it has proved to be a good minimally invasive marker to track protein species,

understanding protein function, and as probes to analyze interactions with other proteins. Lastly, it has also been used as a biosensor to read biological events and signals.

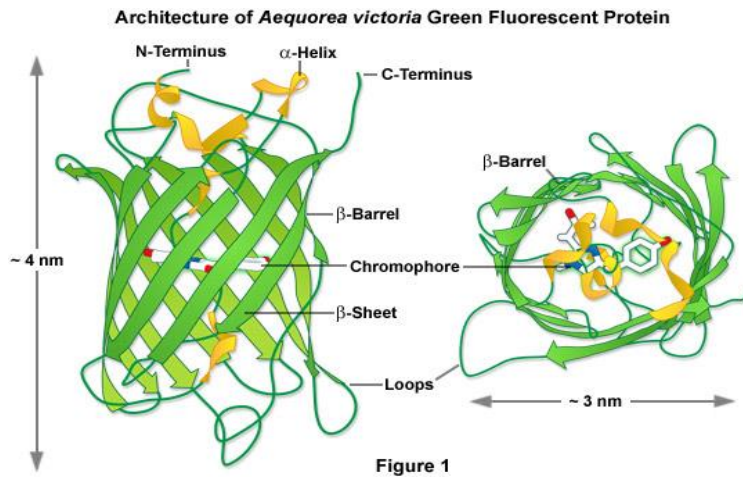


Figure 6: Structure of green fluorescent protein¹¹

mCherry was another fluorescent protein chosen for our study. Showing higher tissue penetration than GFP and better photostability, this protein, derived from the jellyfish *Discosoma striata* is being employed as a tool for *in vitro* and *in vivo* assays in addition to drug screening¹⁰. After activating transcription, results can be seen soon as mCherry matures extremely rapidly. It is also photobleaching resistant.

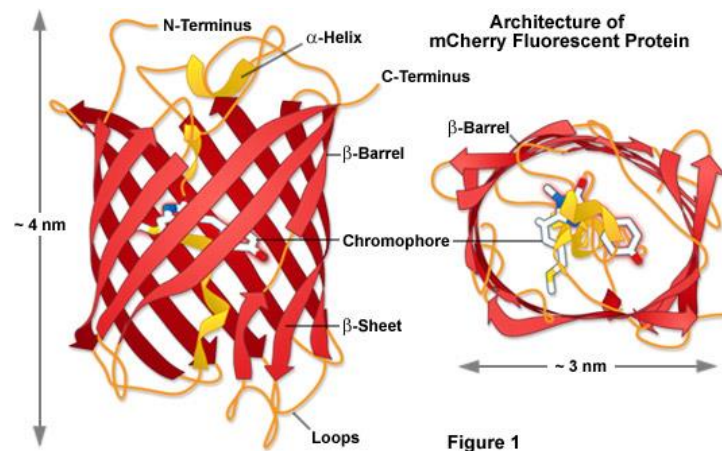


Figure 7: Structure of mCherry¹¹

Hence, using vectors produced by previous studies, we delved into producing these proteins for our targeted drug delivery system.

1.2 Materials and Methods

1.2.1 Chemical reagents

Lithium acetate (1.0 M) – Dissolved 5.1 g of lithium acetate dehydrate in 50 mL water, autoclaved it for 15 min (and sterilizing), and stored at room temperature.

Polyethylene glycol 3350 (50% w/v) – Dissolved 50 g of PEG 3350 in 30 mL of distilled deionized water.

Single-stranded carrier DNA (2 mg/mL) – Dissolved 200 mg of salmon sperm DNA in 100 mL TE buffer, stirring at 4°C for 1-2 hours. Aliquoted samples and stored at -20°C.

1.2.2 Strains and Media

Saccharomyces cerevisiae strain EBY100 containing an integrated copy of Aga1p under GAL1-10 promoter control was used for protein expression and production.

All media used for protein production have been mentioned in the appendix.

1.2.3 Transformation

Transformation of yeast cells with exogenous DNA can be done by a variety of methods like electroporation, agitation with glass beads, DNA-coated microprojectile bombardment or lithium acetate method. Since it has been the most widely used, we decided to pursue lithium acetate method with single-stranded carrier DNA and polyethylene glycol.

We transformed the yeast cells with 3 plasmids, pRS-4420-mCherry-S6-His6-cymc, pRS-4420-eGFP-Tag1, and pRS-4420-Avitag (a non-fluorescent protein) as control.

The transformation and protein induction protocol are mentioned in the appendix.

1.2.4 Protein Purification and Characterization

Once the samples had grown in the SGCAA+Ura media for 3 days, we performed buffer exchange using Amicon Ultra Centrifugal filters and introduced the samples to binding

buffer. We performed HPLC (AKTA Pure) and expected to see peaks at around 587 nm for mCherry and around 488 nm for eGFP. Unfortunately, we were unable to obtain peaks from either of them. We had reserved part of the sample for fluorescence microscopy to detect the expression of the fluorescent proteins.



Figure 8: Fluorescence microscopy image of eGFP in fitc at 20x



Figure 9: Fluorescence microscopy of mCherry TRITC at 20X

We concluded that the proteins might have been expressed in tiny amounts within the cells, but not outside of the cells which was the main goal; so that it could be extracted, concentrated and used for enzymatic assays.

With our non-fluorescent protein control, pRS-4420-Avitag, we decided to run SDS-PAGE and Western Blot to see if our protein was being expressed or not. Our protein was about

25 kD but from the SDS-PAGE results, we ended up having a band at around 60 kD. This meant, there might have been non-specific expression. From western blot, we were unable to find a band.

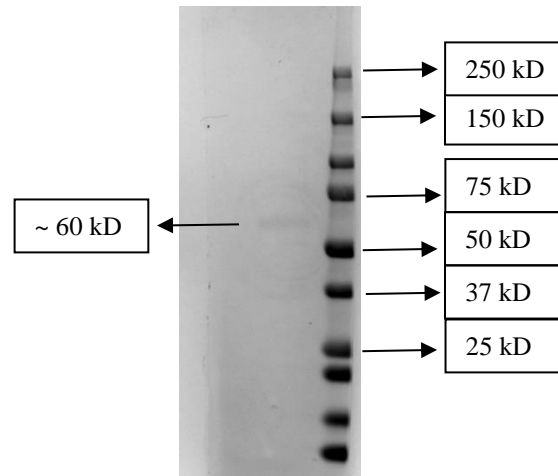


Figure 10: SDS PAGE results of prs-4420-Avitag

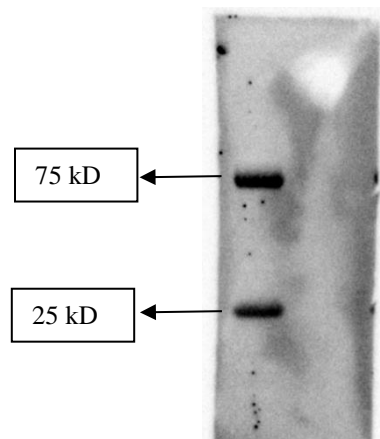


Figure 11: Western blot result of prs-4420-Avitag

Chapter 2: Construction of recombinant protein with mOx GFP

2.1 Background

For targeted drug delivery, we initially depended on fluorescent proteins like mCherry and eGFP to act as biomarkers that would indicate protein localization and levels, at the same time, trying to reduce perturbation in the cell. However, as seen in the previous chapter, we were unable to produce the protein even after purification and concentration. This could be due to a poor secretory pathway where the fluorescent proteins might have misfolded and some non-native disulphide bonds might have been formed which resulted in failure to fluoresce and low (or no) protein production¹².

We turned to using an inert fluorescent protein specifically optimized to fill this gap, which can be used in oxidizing environments and some biological membranes. This monomeric 'ox'GFP or mOx GFP was created (among others) as a super folder protein and was able to clearly label the Golgi complex with high signal/noise. It can also be used to track phenotypic and morphologic changes in biological membranes¹³.

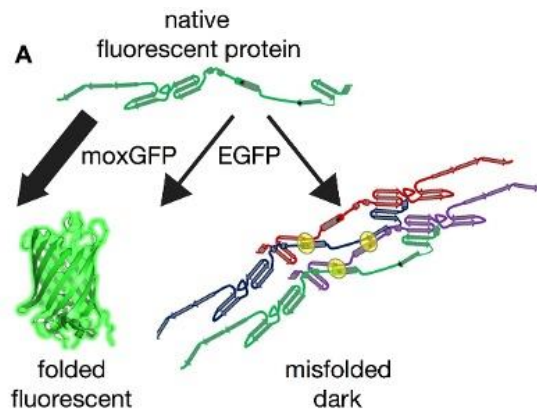


Figure 12: Tangled misfolded dark EGFP and moxGFP solution¹²

In oxidizing environments and some secretory pathways, disruptive interchain disulfide bonds have a high probability of being formed via EGFP, which results in dark fluorescent

proteins. But the advantage that mOx GFP holds, is that it doesn't have cysteines and so it doesn't form the disulfide bonds, which makes it a highly efficient folding protein and gives robust signal. This would perform as a superior biomarker in a targeted drug delivery system¹³.

2.2 Materials and Methods:

2.2.1 Reagents, Materials and Kits

All materials and reagents used in this paper were listed in table 5 in appendix A

2.2.2 Strains and Media

Escherichia coli strain DH5 α (New England Biolabs) was used for recombinant plasmids gene cloning.

2.2.3 Plasmid Design

Plasmid Sequence for mOx GFP is:

```
ATGGTGTCCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA
CGTAAACGGCCACAAGTTCTCCGTGCGGGGCGAGGGCGAGGGCGATGCCACCAACGGCAAGCTGA
CCCTGAAGTTCATCAGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGA
CCTACGGCGTGCAGAGCTTCTCCCGTACCCCGACCACATGAAGCGCCACGACTTCTTCAAGAGC
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTCCTTCAAGGACGACGGCACCTACAAGAC
CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACT
TCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTCAACTCCCACAACGTCTAT
ATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACGTGGAGGA
CGGCTCCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCT
GCCCGACAACCACTACCTGTCCACCCAGTCCAAGCTGTCCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTTCTGGAATTCGTGACCGCCCGGGATCACTCACGGCATGGACGAGCTGTACAAG
```

The pRS backbone vector with 4420-mCherry-S6-His6-cymc is provided below, with the restriction enzymes highlighted in red. The goal is to clone in the insert in that region. We wanted to clone it into another poly-tag backbone, with the same restriction enzymes.

Feature	Direction	Type	Location ↓
▶ 4-4-20\VL\Chain	>>>	misc_difference	3540..3875
▶ Peptide\Linker	>>>	misc_feature	3876..3953
▶ 4-4-20\VH\Chain	>>>	misc_difference	3954..4304
▶ mCherry	>>>	misc_feature	4326..5033
S6 tag	>>>	misc_feature	5049..5078


```

* 10 * 20 * 30 * 40 * 50 * 60 * 70 *
3451 gtattaacagatatataaatcaattc atgaagggttttgattgtcttgttggctatcttcgctgctttgccattgg
3526 ccttggctgctaggc gacgtcggttatgactcaaacaccactatcaacttcctgttagtctaggtgatcaagcctcca
3601 tctcttgcagatctagtcagagccttgtacacagtaatggaacacctatttacgttggtagctgcagaagccag
3676 gccagctctccaaaggtcctgatctacaagtttccaaccgattttctggggtcccagacaggttcagtgaggagtg
3751 gatcagggacagatttcaactcaagatcagcagagtgaggctgaggatctgggagtttatttctgctctcaaa
3826 gtacacatgttccgtggacgttcgggtggaggcaccagcttgaatlaagtcctctgctgatgatgctaagaagg
3901 atgctgctaagaaggatgatgctaagaagatgatgctaagaagatggtagcgtcaaacggatgagactggag
3976 gaggcttgggtcaacctgggagggccatgaaactctcctgtgttgctctggattcaacttttagtgactactgga
4051 tgaactgggtccgcccagctctccagagaaaggactggagtggttagcacaaattagaacaacaccttataatgatg
4126 aaacatattatcagattctgtgaaaggcagattcaccatctcaagagatgattccaaaagttagtctacctgc
4201 aaatgaacaacttaagagttgaagacatgggtatctattactgtacgggttcttactatggttaggactactggg
4276 gtcaaggaacctcagtcaccgtctcctcaacggcgggtggcggctgctaccg atggtgagcaagggcgaggagata
4351 acatggccatcatcaaggagttcatggcctcaaggtgacatggagggtcctgtaacggccacaggttgcaga
4426 tgcagggcgagggcgagggcccccctacaggggcaaccagaccgcaagctgaaagtgaaacaggggtggcccc
4501 tggccttcgcctgggacatcctgtccctcagttcatgtacggctccaaggcctacgtgaaagcaccgccgaca
4576 tccccgactacttgaagctgtccttccccgagggcttcaagtgaggcggctgatgaacttcgaggacggggcg
4651 tggtagcogtgaaccaggactcctcctgcaggacggcgagttcatctacaaggtgaagctgcggggcaccact
4726 tccccctccgacggcccgtaatgcagaagaagaccatgggctgggagggcctcctccgagcggatgtaccocgagg
4801 acggcgccctgaaggcgagatcaagcagaggctgaagctgaaggacggggccactacgacgctgaggtcaaga
4876 ccacctacaaggccaaagaagccogtgcagctgcccggcgctacaacgtcaacatcaagttggacatcacctccc
4951 acaacgaggactacaccatcgtggaaacagtagcaacgcgcggagggcccccactccaccggcgatggacgagc
5026 tgtacaaggcggccgggggtggcGGCGATAGCCTGAGCTGGCTGCTGCCCTGCTGAACcatcaccatcatcatc
5101 aggaacaaaagcttatttctgaagaggacttgaactcgagatctgataacaacagttgtagatgtaacaaaatcg
5176 actttgttcccactgtacttttagctcgtacaaaatacaatatacttttcatttctccgtaaacacatgttttc

```

Figure 13: Original pRS vector sequence, with 4420-mCherry-S6-His6-cymc with Mlu1 and Not1 highlighted in red.

2.2.4 Digestion, Ligation, and Transformation

mOx GFP insert was amplified using PCR. The sense primer (or forward primer) introduces Mlu1 restriction site on the 5' end while the antisense primer (or reverse primer) introduces Not1 on the 3' end to aid in cloning. The insert was then purified via agarose gel electrophoresis to obtain the pure insert. The inserts and backbone vector pRS-4420-mCherry-S6-His6-cymc and pRS-4420-tags1 were double digested using Mlu1 and Not1. Both restriction enzymes are time-saver qualified and require 5-15 min for digestion. To make sure complete digestion takes place, we incubated the samples at 37°C for 2 hours. The backbone length is around 7000 bp and was extracted using QIAquick Gel Extraction Kit from Qiagen, and the insert was obtained using the QIAquick PCR Purification Kit from

the same company. The concentration of both samples was measured using NanoDrop 2000 (Thermo Fisher Scientific). It was then heat inactivated at 80°C for 20 minutes to stop the reaction.

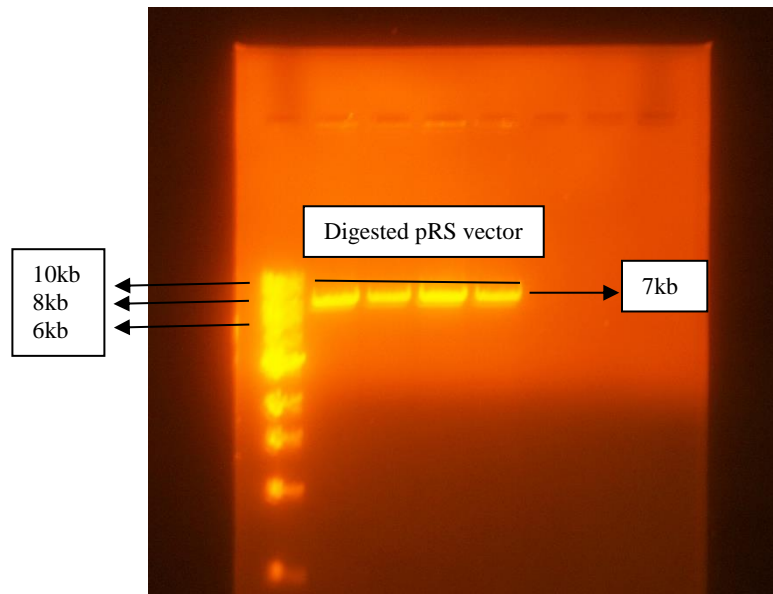


Figure 14: Digested pRS-4420-tags1 (lane 1 & 2) pRS-4420-mCherry-S6-His6-cymc (lane 3 & 4) vector with 1kb ladder

We employed T4 ligase enzyme for ligation of the inserts and backbone. The T4 ligase buffer provided with the enzyme contains ATP, which facilitates the formation of a phosphodiester bond between the juxtaposed 5' phosphate and 3' hydroxyl termini in the duplex DNA. During digestion, the restriction enzymes create sticky ends on the backbone and inserts, and the T4 ligase enzyme brings these two together by ligating the two cut sites on both.

Insert	Concentration (ng/uL)	Backbone Vector	Concentration (ng/uL)
mOx GFP #1	11.5	S6 Tag	18.9
mOx GFP #2	18.8	Tags1	24.5

Table 1: Concentration of inserts and backbone - digestion (mOx GFP)

The ligation mixes were calculated using molar ratio 3:1 for insert and backbone, as shown below:

S6 Tag		Tags1	
T4 Ligase Enzyme	1	T4 Ligase Enzyme	1
T4 Buffer	2	T4 Buffer	2
Insert	0.5	Insert	0.65
Backbone	16.5	Backbone	16.35
Total	20	Total	20

Table 2: mOx GFP Ligation set up

Ligation was performed for 4 hours at room temperature. After ligation, the plasmids were transformed into DH5 α cells (*E. coli*) using heat shock method. Then the cells were grown overnight, on LB Amp plates at 37°C. Most of the colonies grown on the plate were transformed successfully as they contained the gene for resistance that is present in the pRS vector. 10 colonies from each plate were picked and colony PCR was performed to screen for positive colonies. On new LB Amp plates, the colony suspension of the positive colonies were replated. They were miniprep using QIAquick Miniprep Kit from Qiagen. The samples were then sent to Genewiz for DNA sequencing to confirm our sequences.

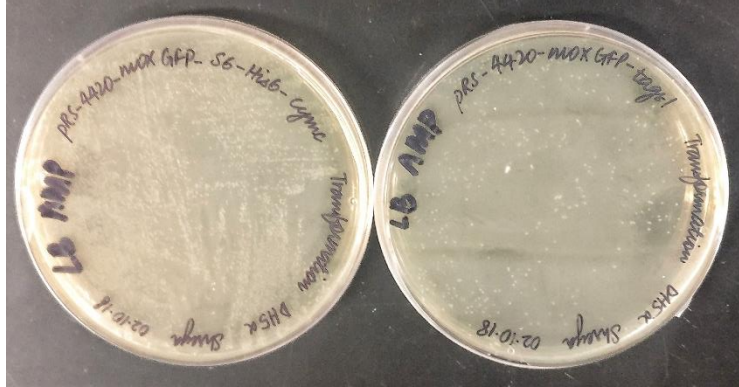


Figure 15: Transformed mox GFP in S6 and tags1 vectors

2.2.5 Sequencing Results

Below are the sequence alignments for mOx GFP in pRS-4420-mCherry-S6-His6-cymc and pRS-4420-Tags. The enzymes are highlighted in red.

```

Mon Apr 23, 2018 13:24 PDT
New DNA from 1 to 717
Alignment to
pra-4420-moxgfp-tags1-R5-sk-prspctbbrev (2).seq from 1 to 1079
Matches (|):717
Mismatches (#):0
Gaps ( ):362
Unattempted (.) :0

1 ----- 1
1 NNNNNNNNNNNNNNNNTCTTNNNGANTNNGCTTTTGTTCGTGATGATGGTGGATGTCGTGCCATTTCGATTTTCTGAGCCTCGAAGATGTCGCTCAG 100

1 -----CTGTACAGCTCCATGCCCTGAGTGTATCCCGCGCCGTCACGAATTCAGAAAGACCATGTGATCCGCTTCTCGTTG 82
101 ACCGCCACCGCCGCCCTTGTACAGCTCCATGCCCTGAGTGTATCCCGCGCCGTCACGAATTCAGAAAGACCATGTGATCCGCTTCTCGTTG 200

83 GGGTCTTTGGACAGCTTGGACTGGGTGGACAGGTAGTGGTTGTCCGGGACGACGACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTGTGTCGCGGA 182
201 GGGTCTTTGGACAGCTTGGACTGGGTGGACAGGTAGTGGTTGTCCGGGACGACGACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTGTGTCGCGGA 300

183 GCTGCCACGGAGCCCTCTCCACCTGTGGCCGATCTTGAAGTTGGCTTGTATGCCCTTCTTCTGCTTGTCCGGCGTATATAGACGTGTGGGAGTTGAA 282
301 GCTGCCACGGAGCCCTCTCCACCTGTGGCCGATCTTGAAGTTGGCTTGTATGCCCTTCTTCTGCTTGTCCGGCGTATATAGACGTGTGGGAGTTGAA 400

283 GTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCTTGAAGTTCGATGCCCTTCAGCTCGATGGCGTTACACAGGGTGTGCCCTCGAACTTACC 382
401 GTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCTTGAAGTTCGATGCCCTTCAGCTCGATGGCGTTACACAGGGTGTGCCCTCGAACTTACC 500

383 TCCGGCCGGGCTCTTGTAGTTCGCCCTCTCTGAAGAGATGTCGCCCTCTGGACGTAGCCTTCCGGCATGGCCCTTGAAGAACTCTGGCCCTTCA 482
501 TCCGGCCGGGCTCTTGTAGTTCGCCCTCTCTGAAGAGATGTCGCCCTCTGGACGTAGCCTTCCGGCATGGCCCTTGAAGAACTCTGGCCCTTCA 600

483 TGTGGTCCGGGTACCGGAGAACTCTGCACGCCGTAGGTACAGGTGGTCCAGAGGTGGGCCAGGGCACGGGCAGCTTGGCCGTGGTGTGATGAACCTT 582
601 TGTGGTCCGGGTACCGGAGAACTCTGCACGCCGTAGGTACAGGTGGTCCAGAGGTGGGCCAGGGCACGGGCAGCTTGGCCGTGGTGTGATGAACCTT 700

583 CAGGTCAGCTTGGCCCTGTGGCATCGCCCTCGCCCTCGCCCGCACGGAGAATCTGTGGCCGTTTACCTCGCCCTCCAGCTCGACCCAGGATGGGCACC 682
701 CAGGTCAGCTTGGCCCTGTGGCATCGCCCTCGCCCTCGCCCGCACGGAGAATCTGTGGCCGTTTACCTCGCCCTCCAGCTCGACCCAGGATGGGCACC 800

683 ACCCCGGTGAACAGCTCCTCGCCCTTGC-----G-AC-A---CCA-----T----- 717
801 ACCCCGGTGAACAGCTCCTCGCCCTTGTACACCATCGTACGACCGCCACCCAGCCTTGAGGAGACGGTACTGANTTCTTGCACCCAGTATGTCACAT 900

717 ----- 717
901 CATAGTAAAGAACCCGTACAGTAAATAGATACCCATGTCTTCAACTCTTAAAGTTGTTTCATTTCGNNTAGACACTACTTTTGAATCATCTTGGAGATGGTC 1000

717 ----- 717
1001 AATCTGCCTTTCACAGATCTGAATATATGTTTCATAAATAAAGGTTGNTTCTAATTTGTGCTACCCACTCCAGTCN 1079
  
```

Figure 16: Genewiz confirmation for Tags1

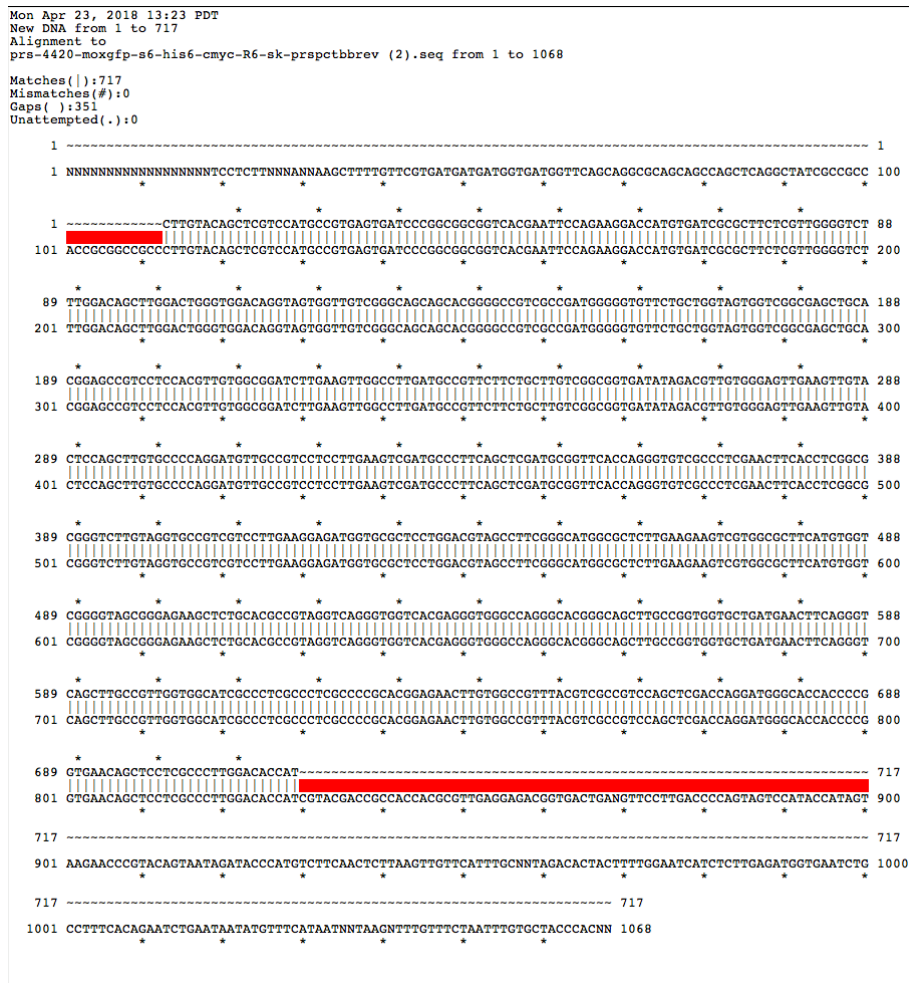


Figure 17: Genewiz confirmation for S6

As it can be seen, both samples paired up correctly. Part of the sample was grown up in LB Amp broth overnight, and stored with glycerol at -80°C, which can be used in the future for protein production and enzymatic assays.

2.3 Codon Optimization of mOx GFP for yeast cells

2.3.1 Background

Translation is the final step of synthesis of protein from DNA via mRNA. In the genetic code, there are 64 codons, which give rise to 20 amino acids which means that 3 nucleotides code for one amino acid. By this math, each amino acid has multiple codons. Since the translation from nucleotide codons to amino acids is not 1:1, “degeneracy” occurs and some parts of

the genetic code are redundant. Every organism prefers particular codons, and those are selected over others that code for the same protein. Codon optimization works towards retrieving the codons preferred by the organism in selection and this is thought to increase protein expression¹⁴.

The amount of protein expressed will depend on the protein and organism, in addition to other factors like protein stability, protein transport, and protein folding kinetics, to name a few. But optimizing the protein for the organism experimented upon could possibly open up better protein expression results.

Our main focus was to produce fluorescent proteins from yeast cells as they have fast growth and can be genetically manipulated easily. They also have good secretory pathways for proper protein processing and superior post-translational modification. Hence, it only made sense to codon optimize mOx GFP to fit our needs with yeast¹⁵.

2.3.2 Codon Optimization of mOx GFP

The original mOx GFP sequence can be viewed in chapter 2. The first step to this was using ExPASy tool to translate the DNA sequence to protein. Once I inputted my sequence, it was translated and I had to choose an open reading frame. I chose 5' to 3' frame 1 and viewed the results.

Translate

Translate tool

New translate tool, please test!
Use the [ExPASy contact form](#) to submit feedback and bug reports

Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored).

```

ATGGTGTCCAAGGGCGAGGAGCTGTTCAACCGGGTGGTGCCTCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCTCCGTGCGGGGCGAGGGCGAGGGCGATGCCACCAACGGCAAGCTGACCTGAAGTTCA
TCAGCACCACCGGCAAGCTGCCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGACAGGC
TTCTCCCGCTACCCCGACACATGAAGCGCCACGACTTCTTCAAGAGCGCCATGCCCGAAGGCTACGTCCA
GGAGCGCACATCTCCTTCAAGGACGACGGCACCTACAAGACCCGCGCGAGGTGAAAGTTGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG
GAGTACAACCTTCAACTCCCAACAGCTATATCACCGCCGACAAGCAGAAAGAACGGCATCAAGGCCAATT
CAAGATCCGCCACAACGTGGAGGACGGCTCCGTGACGCTCGCCGACCACTACCAGCAGAAACACCCCATCG
GGACGGCCCGGTGCTGCGCCGACAACCACTACCCTGTCACCCAGTCCAAGCTGTCCAAGGACCCCAAC
GAGAAGCGGATCACATGGTCTTCTGGAATTCGTGACCGCCGCGGATCACTACGGCATGGACGAGCT
GTACAAG

```

Output format:

or

Figure 18: Translate tool at ExPasy

Translate

Translate Tool

Open reading frames are highlighted in **red**. Please select one of the "Methionine" or one of the highlighted residues following a **Stop** codon (or the beginning of the sequence). This will create a virtual Swiss-Prot entry, comprising the residues from your chosen start position up to the following **Stop** codon.

```

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

```

Figure 19: Open reading frame at ExPasy

Once I had the translated protein sequence, I switched over to DNA works at High Performance Computing at NIH, to use their tool to get the codon optimized DNA. I selected the organism as *S. cerevisiae* and kept the pre-conditioned parameters as is. Then, I inputted the protein sequence and selected the design oligos option.

Codon Frequency Table:

GCG CodonFrequency format only. For more organisms, check out the [Codon Usage Database](#)

Choice 1 - Choose a standard organism: Choice 2 - Enter custom codon frequency table manually:

Choice 3 - Upload codon frequency file: Choose File No file chosen

M. musculus
P. pastoris
R. norvegicus
S. cerevisiae
X. laevis

Figure 20: Choosing *S. cerevisiae* for yeast codon optimization

The DNA sequence # 1 is:

```
-----  
1 ATGGTTTCTAAAGGCGAAGAATTGTTTACTGGTGTGTACATTTTTGGTTGAACTGGAT  
61 GGTGATGTTAATGGTCATAAGTTTTCTGTCAGAGGTGAAGGGGAGGGTGATGCTACTAAT  
121 GGCAAACTGACTCTGAAATTCATCTCTACTACTGGTAAATTGCCAGTCCCATGGCCAAT  
181 TTGGTAACTACTTTGACTTATGGTGTTCAGTCTTTTTCTAGATATCCAGATCACATGAAA  
241 AGACATGACTTCTTTAAATCTGCAATGCCAGAGGGTTATGTTCCAGGAAAGAACTATATCT  
301 TTTAAGGACGATGGTACATATAAGACTAGAGCTGAAGTTAAGTTCGAAGGTGATACATTG  
361 GTTAATAGAATCGAGTTGAAAGGTATTGATTTCAAAGAGGATGGGAACATTTTGGGACAT  
421 AAAC TTGAATATACTTTAACTCTCATAATGTATATATCACTGCTGATAAACAAAAAAC  
481 GGTATTAAGCTAATTTCAAGATTAGGCACAATGTAGAAGATGGGTCTGTTCAATTGGCT  
541 GATCATTATCAACAGAACACCCCAATCGGTGACGGTCTGTCTTGTGTCAGATAATCAC  
601 TATTTGCTACTCAATCTAAGTTGTCTAAAGATCCAAATGAAAAAGAGATCATATGGTT  
661 TTATTGGAATTTGTGACTGCTGCTGGTATTACTCATGGCATGGACGAGCTATATAAG  
-----
```

Figure 21: Sequence obtained after optimization

2.3.3 Plasmid Design

Now that I was able to obtain my codon optimized DNA sequence, I added in the restriction enzymes I wanted to work with, Mlu1 and Not1 to make the final design:

```
TCAGTGCTAAGACGCGTGGTGGCGGTTCGTACGATGGTGTCAAAGGGTGAAGAATTGTTTACTGG  
TGTTGTTCCCTATTTTGGTTGAATTGGATGGTGACGTTAATGGACATAAGTTCTCTGTTAGAGGG  
GAAGGTGAGGGAGATGCTACTAATGGTAAATTGACCTTGAAATTCATTTCTACTACTGGGAAAT  
TGCCAGTTCATGGCCAACCTTGGT TACTACTTTGACTTATGGTGTTC AATCTTTTTCCAGATAT  
CCAGATCACATGAAAAGACATGATTTTTTCAAATCTGCTATGCCAGAGGGTTATGTTCAAGAAA  
GAACTATTAGTTTCAAAGATGATGGCACCTATAAAACAAGAGCTGAGGTTAAATTCGAAGGAGA  
CACTTTGGTTAACAGAATTGAACTAAAAGGTATTGATTTTAAGGAAGATGGTAATATTCTGGGT  
CATAAATTGGAATACA ACTTCAACTCCCACAACGTCTACATTACTGCTGATAAGCAAAAAAACG
```

GTATTAAGGCTAACTTCAAATTAGACATAACGTTGAAGACGGCAGTGTTTCAGTTGGCAGATCA
TTATCAACAAAATACTCCAATCGGTGATGGTCCAGTCTTGTTGCCAGATAATCACTATTTGTCTA
CTCAATCTAAGCTTCCAAAGATCCAAATGAGAAAAGAGATCATATGGTTTTGCTTGAATTTGT
TACTGCAGCCGGTATCACTCATGGTATGGATGAACTATATAAGGCGGCCGCTACATGCTACTAC

The PCR forward primer (for colony PCR) was designed in accordance with the new
sequence as: ATG GTG TCA AAG GGT GAA GAA TTG TTT ACT GG

Chapter 3: Construction of recombinant protein with Flagelliform and GGS

Linkers

3.1 Background

For thousands of years, spider silk has been highly studied and been in the limelight due to its strong mechanical properties, ductility, and toughness. Flagelliform silk protein from *Nephila clavipes* has been identified to be superior to other spider silks like major ampullate (MA) silk, aciniform silk, and dragline silk¹⁶. Because of its repeating structural motif of GPGGA, it gives itself high elasticity power, and the ability to stretch to 500-1000%¹⁷. Compared to MA silk which has 35% extension, flagelliform silk has over 200% extension in its spring like spirals. Weaving spiders make aerial nets (to entrap flying prey) which have spiral orb webs that are made of flagelliform silk as the core fiber; which results in its high elasticity¹⁸.

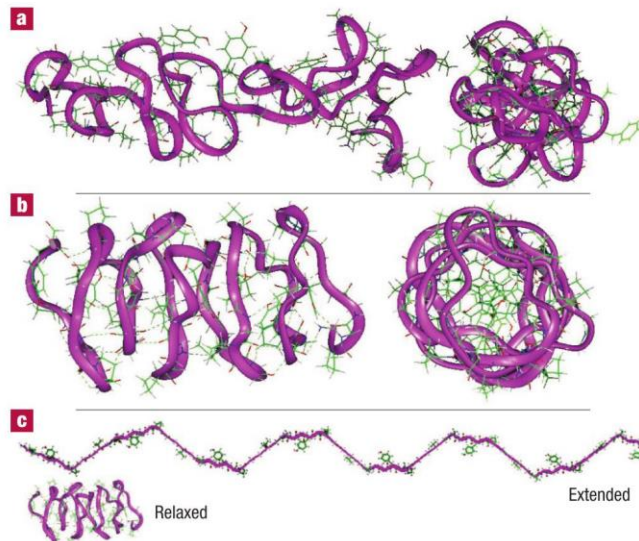


Figure 22: Molecular models for relaxed and extended flagelliform protein sequences from spider capture silk. a, b, Side and end views for possible flagelliform protein conformations of (a) *Araneus gemmoides*, 85 amino acids long (sequence is VGPGGAYGPGGVYGPAGGLSGPGGAGPYGPGGVGPGGAGPYGPGGVGPGGAGPYGPGGVGPGGAGPYGPGGVGPGGAGPYGPGG), and (b) *Nephila clavipes*³³, 75 amino acids long, (GPGGX)¹⁵, where X is Y or V, alternately. c, Scale models for extended and relaxed GPGGX sequences, 75 amino acids long. The extended model is at the maximum extension of the protein, without deforming bond angles.

This elasticity and flexibility encouraged us to use it as a short peptide linker between the 4420 (single chain antibody) and mCherry fluorescent protein, along with S6 and His6 tags. In the bigger picture, this construction will actually be the bridge connecting therapeutic nanoparticles and cancer cells. We constructed another flexible peptide linker (GGG)_n which mainly has copies of Gly and Ser residues. To maintain the necessary inter-domain interactions, or separation of functional domains, we can adjust the copy number or 'n' to suit our GGS linker needs. With this dynamic linker duo, the mechanical and kinetic properties of the targeted drug delivery platform can be improved even further due to increased flexibility and connection of nanoparticles and cancer cells. Previous studies were performed by constructing sequences using Flagelliform and GGS linkers with 25 amino acids in length. We are interested in fusing Flagelliform and GGS linkers with 50 amino acids (as copy number) and perform computer simulation tools to see which would perform better in a drug delivery platform. Experimentally, the dynamic data will be measured in a flow chamber after successfully conjugating with surface modified nanoparticles.

3.2 Materials and Methods:

3.2.1 Reagents, Materials and Kits

All materials and reagents used in this paper were listed in table 5 in appendix A

3.2.2 Strains and Media

Escherichia coli strain DH5α (New England Biolabs) was used for recombinant plasmids gene cloning.

3.2.3 Plasmid Design

The inserts, Flagelliform (Flag 50) and GGS (50) as gBlocks® gene fragments, were purchased from Integrated DNA Technologies (IDT). As mentioned, GGS linker was constructed using repeated Gly-Gly-Ser motifs. We cannot perform PCR for both these inserts as the primers do not have a specific binding region. Below are the Flag50 and GGS50 sequences:

Flag50

TAA TTA TGA AAC ATA TTA TTA CGC GTG GGC CAG GGG GTG CCG GTC CGG GAG GGG CAG
GTC CAG GAG GTG CGG GCC CCG GGG GAG CGG GCC CTG GGG GAG CGG GTC CGG GCG GCG
CAG GAC CCG GAG GCG CGG GTC CTG GTG GAG CAG GCC CTG GAG GAG CCG GCC CGG GTG
GCG CAC GTA CGT AAA CAA ACC TTA TAA TTA T

GG50

GGA TTC ACT TAC GCG TGG CGG CAG TGG AGG CAG CGG CGG TAG TGG TGG ATC AGG AGG
CAG TGG GGG ATC CGG CGG AAG CGG TGG GTC TGG AGG ATC AGG TGG GTC TGG CGG GAG
CGG TGG GAG TGG GGG GAG CGG CTC TGG GGG TTC TGG TGG TTC AGG CGG CCG TAC GGT
TAT GAC TCA AA

PCR primers for each were made by a former graduate student. The pRS backbone vector with 4420-mCherry-S6-His6-cmyc is provided below, with the restriction enzymes highlighted in red. The goal is to clone in the insert in that region.

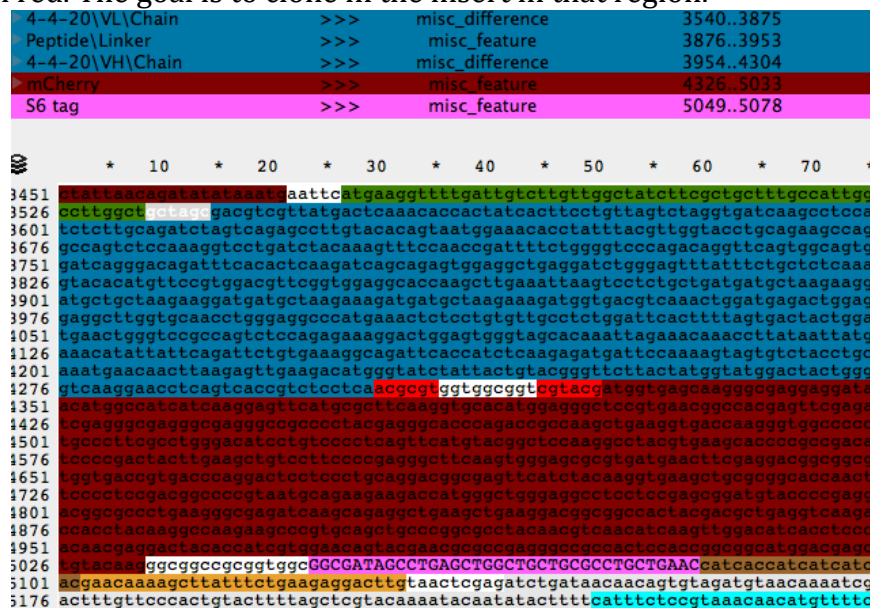


Figure 23: Original pRS vector sequence, with 4420-mCherry-S6-His6-cmyc with Mlu1 and Bsw1 highlighted in red.

3.2.4 Digestion, Ligation, and Transformation

The products were delivered by IDT and hydrated with TE buffer to the required concentration. The inserts and backbone vector pRS-4420-mCherry-S6-His6-cmyc were double digested using Mlu1 and Bsiw1. The backbone length is around 7000 bp and was extracted using QIAquick Gel Extraction Kit from Qiagen, and the insert was obtained using the QIAquick PCR Purification Kit from the same company. The concentration of both samples was measured using NanoDrop 2000 (Thermo Fisher Scientific). It was then heat inactivated to stop the reaction.

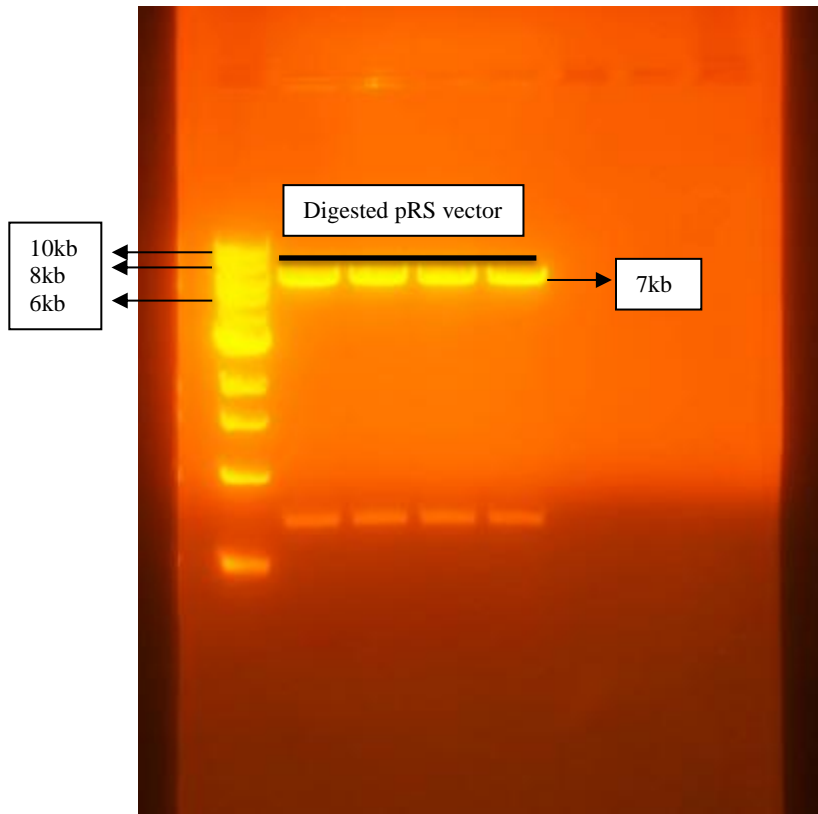


Figure 24: Digested pRS-4420-mCherry-S6-His6-cmyc vector with 1kb ladder

We employed T4 ligase enzyme for ligation of the inserts and backbone. The T4 ligase buffer provided with the enzyme contains ATP, which facilitates the formation of a phosphodiester bond between the juxtaposed 5' phosphate and 3' hydroxyl termini in the

duplex DNA. During digestion, the restriction enzymes create sticky ends on the backbone and inserts, and the T4 ligase enzyme brings these two together by ligating the two cut sites on both.

Insert	Concentration (ng/uL)	Backbone Vector	Concentration (ng/uL)
Flag50	42.8	#1	13.5
GG50	31.3	#2	11.5

Table 3: Digestion concentrations for Flag50 and GGS50

The ligation mixes were calculated using molar ratio 3:1 for insert and backbone, as shown below:

Flag50		GG50	
T4 Ligase Enzyme	1	T4 Ligase Enzyme	1
T4 Buffer	2	T4 Buffer	2
Insert	0.17	Insert	1.11
Backbone	16.83	Backbone	15.88
Total	20	Total	20

Table 4: Ligation setup for Flag50 and GGS50

Ligation was performed for 4 hours at room temperature. After ligation, the plasmids were transformed into DH5 α cells (*E. coli*) using heat shock method. Then the cells were grown overnight, on LB Amp plates at 37°C. Most of the colonies grown on the plate were transformed successfully as they contained the gene for resistance that is present in the pRS vector. 10 colonies from each plate were picked and colony PCR was performed to screen for positive colonies. On new LB Amp plates, the colony suspension of the positive

colonies were replated. They were miniprepmed using QIAquick Miniprep Kit from Qiagen. The samples were then sent to Genewiz for DNA sequencing to confirm our sequences.

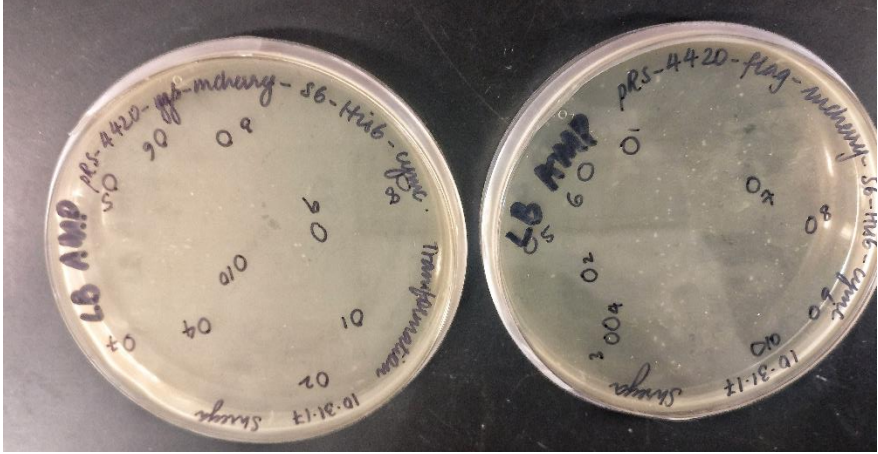


Figure 25: Transformed flagelliform50 and ggs50 in S6 vector

3.2.5 Sequencing Results

Below are the sequence alignments for Flag50 and GGS50. The enzymes are highlighted in red.

```
Tue Apr 24, 2018 13:50 PDT
Flag_50_linker.ape from 1 to 162
Alignment to
prs-4420-flag-mcherry-s6-his6-cymo-R10-mCherry_FindRev (3).seq from 1 to 1014
Matches():162
Mismatches(#):0
Gaps():852
Unattempted(.):0

  1
  1 NNNNNNNNNNNNNNNNGANGCCTGTTATCTCTCGCCCTTGCTCACCATCGTAGGTGCGCCACCCGGGCCGGCTCTCCAGGGCTGCTCCACC 45
  46 AGGACCCCGGCCTCCGGTCTCGGCCGCCCGGACCCGCTCCCCAGGGCCCGCTCCCCGGGGCCCGCACCTCTGGACTGCCCTCCCGGACCCGCA 145
  101 AGGACCCCGGCCTCCGGTCTCGGCCGCCCGGACCCGCTCCCCAGGGCCCGCTCCCCGGGGCCCGCACCTCTGGACTGCCCTCCCGGACCCGCA 200

  146 CCCCCTGGCCCACCCCT
  201 CCCCCTGGCCCACCCCTGAGGAGACGGTGACTGAGGTTCCCTTGACCCAGTAGTCCATACCATAGTAAGAACCCTACAGTAATAGATACCCATGCTTT 300

  162 ----- 162
  301 CAACTCTTAAGTTTTCATTTGAGGTAGACACTACTTTTGAATCATCTCTTGAGATGGTGAATCTGCCTTTCACAGAATCTGAATAATATGTTTCATA 400

  162 ----- 162
  401 ATTATAAGTTTGTTCATTTGCTACCCACTCCAGTCTTTCTCTGGAGACTGGCGGACCCAGTTCATCCAGTAGTCACTAAAAGTGAATCCAGAG 500

  162 ----- 162
  501 GCAACACAGGAGAGTTTCATGGGCCCTCCAGGTTGCACCAAGCCCTCCAGTCTCATCCAGTTTGACGTACCATCTTTCTTAGCATCATCTTTCTTAG 600

  162 ----- 162
  601 CATCATCTTCTTAGCAGCATCTTCTTAGCATCATCAGCAGAGGACTTAATTTCAAGCTTGGTGCCCTCCACCGAACGTCACGGAACATGTACTTTG 700

  162 ----- 162
  701 AGAGCAGAAATAAATCCCAGATCCTCAGCCTCCACTCTGCTGATCTTGAGTGTGAAATCTGTCCCTGATCCACTGCCACTGAACCTGTGGGACCCCA 800

  162 ----- 162
  801 GAAAATCGGTTGGAACTTTGTAGATCAGGACCTTTGGAGACTGGCTGGCTTCTGCANGTACCAACGTAATAGGTGTTTCCATTACTGTGTACAAGCC 900

  162 ----- 162
  901 TCTGACTAGATCTGCAAGAGATGGANGCTTGATCACCTAGACTAACAGGAAGTGATAGTGGTGGTTGAGTCATAACGACGTCGCTAGCAGCAAGGCCAA 1000

  162 ----- 162
  1001 TGGNAAAGCAGCANN 1014
```

Figure 26: Genewiz confirmation of flag50

```

Mon Apr 23, 2018 13:00 PDT
GGS_50_linker.ape from 1 to 185
Alignment to
prs-4420-ggs-mcherry-s6-his6-cymc-R3-mCherry_FindRev (1).seq from 1 to 1092

Matches():182
Mismatches(#):1
Gaps():911
Unattempted(.):0

1 -----T--TTGA-----G--T--CAT--A--ACCGTACGGCCGGCTGAACCACCAGAACCCAGAGCCACCGCTC 56
1 NNNNNNNNNNNNNNNNNNGANGNNTGTT-ATCCTCCTCGCCCTTGC-TCACCATGGTACGGCCGCTGAACCACCAGAACCCAGAGCCACCGCTC 98

57 CCCCACCTCCCACCGCTCCGCGCAGACCCACCTGATCCTCCAGACCCACCGCTTCGCGCGGATCCCCCACTGCCTCCTGATCCACCATAACCGCGCTGC 156
99 CCCCACCTCCCACCGCTCCGCGCAGACCCACCTGATCCTCCAGACCCACCGCTTCGCGCGGATCCCCCACTGCCTCCTGATCCACCATAACCGCGCTGC 198

157 CTCACCTGCCCGCCACGCGT--A--A--GTGA--A--TCC----- 185
199 CTCACCTGCCCGCCACGGCTGAGGAGACGGTGACTGAGGTTCCCTTGACCCAGTAGTCCATACCATAGTAAGAACCCTACAGTAATAGATACCCATGTC 298

185 ----- 185
299 TTCACCTCTTAAGTTGTTTCATTTGCAGGTAGACACTACTTTTGGAAATCATCTCTTGAGATGGTGAATCTGCCTTTCACAGAATCTGAATAATATGTTCA 398

185 ----- 185
399 TAATTATAAGTTTGTTCATAATTGCTGCTACCCACTCCAGTCTTCTCTGGAGACTGGCGGACCCAGTTCATCCAGTAGTCACTAAAAGTGAATCCAG 498

185 ----- 185
499 AGGCAACACAGCAGAGTTTCATGGGCCCTCCAGGTGCACCAAGCCTCCAGTCTCATCCAGTTGACGTCAACATCTTCTTACAGTATCATCTTTCTT 598

185 ----- 185
599 AGCATCATCTTCTTAGCAGCATCCTTCTAGCATCATCAGCAGAGACTTAATTTCAAGCTTGGTGCCTCCACCGAACGTCCACGGACATGTGTACTT 698

185 ----- 185
699 TGAGAGCAGAAATAAATCCCAGATCCTCAGCCTCCACTCTGCTGATCTTGAGTGTGAAATCTGTCCCTGATCCACTGCCACTGAACCTGCTGGGACCC 798

185 ----- 185
799 CAGAAAATCGGTTGGAACTTTGTAGATCAGGACCTTTGGAGACTGGCCTGGCTTCTGCAGGTACCAACGTAATAGTGTTTCCATTACTGTGTACAG 898

185 ----- 185
899 GCTCTGACTAGATCTGCAAGANATGGAGGCTTGATCACCTAGACTAACANGAAGTATAGTGGTGTGTTGAGTCATAACGACGTCCGTAGCAGCCAAAGGCC 998

185 ----- 185
999 AATGGCAAAGCAGCANNATAGCCAACAAGANATCAAAACCCCTTCATGAANTCGATCCNGGGTTTTNNNNCCNNNACGTTAAAGTANANANGNN 1092

```

Figure 27: Genewiz Confirmation for GGS50

As it can be seen, both samples paired up correctly. The mutation in GGS50 is outside our region of interest and hence does not affect our research. Part of the sample was grown up in LB Amp broth overnight, and stored with glycerol at -80°C, which would be used in the future for protein production and enzymatic assays.

Future Direction

We have the codon optimized mOx GFP ready to be inserted into our pRS vector. The insert has been obtained by IDT and will be immediately digested and ligated.

Once we have that ready, we can produce protein as per Appendix B protocol for the mOx GFP with and without the yeast codon optimization. For protein purification and characterization, the samples may be run through Ni-NTA columns, and SDS-PAGE and western blot can be performed to determine the expression of the fluorescent proteins in a small scale culture. Another fluorescent protein, mRFP has been obtained from IDT and can be used as a positive control.

Once we are successful in producing the protein, we may perform difference enzymatic assays like Sortase A and Sfp synthase reaction, which will help in protein and therapeutic nanoparticle adhesion system construction. This conjugation can be tested via flow cytometry.

Once the most efficient platform is established, we can scale up our protein production for nanoparticle conjugation and even perform flow chamber kinetic measurement.

References

1. Fortina, P., Kricka, L. J., Surrey, S. & Grodzinski, P. Nanobiotechnology: The promise and reality of new approaches to molecular recognition. *Trends in Biotechnology* 23 (2005) 168-173
2. Si, Duanyun., et al. "Biomedical evolution of nanomedicines". *Asian Journal of Pharmacodynamics and Pharmacokinetics*. 7.2 (2007):83-97
3. Muller, Rainer H., and Cornelia M. Keck. "Challenges and solutions for the delivery of biotech drugs—a review of drug nanocrystal technology and lipid nanoparticles." *Journal of biotechnology* 113.1 (2004): 151-170
4. Gatter, KEVIN C., et al. "Transferrin receptors in human tissues: their distribution and possible clinical relevance." *Journal of clinical pathology* 36.5 (1983): 539-545
5. Davis, Mark E., and Dong M. Shin. "Nanoparticle therapeutics: an emerging treatment modality for cancer." *Nature reviews Drug discovery* 7.9 (2008): 771-782
6. Danhier, Fabienne, Olivier Feron, and Veronique Preat. "To exploit the tumor micro-environment: passive and active tumor targeting of nanocarriers for anti-cancer drug delivery." *Journal of Controlled Release* 148.2 (2010): 135-146
7. Haun, Jered B., et al. "Using engineered single-chain antibodies to correlate molecular binding properties and nanoparticle adhesion dynamics." *Langmuir* 27.22 (2011): 13701-13712
8. Marks, Kevin M., and Garry P. Nolan. "Chemical labeling strategies for cell biology." *Nature methods* 3.8 (2006): 591-596

9. Freitas, Luiza Baptista de Oliveira., et al. “Multifunctional mesoporous silica nanoparticles for cancer-targeted, controlled drug delivery and imaging” *Microporous and Mesoporous Materials* 242 (2017): 271-283
10. Lippincott-Schwartz, Jennifer, and Patterson, George H. “Development and Use of Fluorescent Protein Markers in Living Cells”. *Science* 10.1126 (2003) 87-91
11. zeiss-campus.magnet.fsu.edu/articles/probes/anthozoafps
12. Jain, RK., et al. “Oligomerization of green fluorescent protein in the secretory pathway of endocrine cells”. *Biochem J.* 360.3 (2001): 645-9
13. T, Suzuki. et al. “Development of cysteine-free fluorescent proteins for the oxidative environment”. *PLoS One.* 2012;7(5):e37551
14. Plotkin JB, Kudla G. “Synonymous but not the same: the causes and consequences of codon bias” . *Nat Rev Genet* (2011) 12:32–42
15. Constantini, Lindsay M., et al. “A palette of fluorescent proteins optimized for diverse cellular environments”. *Nat Commun.* (2015); 6: 7670
16. Römer, Lin and Scheibel, Thomas. “The elaborate structure of spider silk: Structure and function of a natural high performance fiber”. *Prion* 2.4 (2008): 154–161
17. Opell, B. D. & Bond, J. E. Capture thread extensibility of orb-weaving spiders: testing punctuated and associative explanations of character evolution. *Biol. J. Linn. Soc.* 70 (2000) 107–120
18. Becker, N. et al. “Molecular nanosprings in spider capture-silk threads”. *Nat. Mater.* (2003) 2, 278–283

APPENDIX A: GENE CLONING AND PROPAGATION

MATERIALS, REAGENTS, and KITS

REAGENTS	COMPANY
Restriction Enzyme Mlu1	New England Biolabs
Restriction Enzyme Not1	New England Biolabs
Restriction Enzyme Bsiw1	New England Biolabs
100bp Ladder	New England Biolabs
1kb Ladder	New England Biolabs
T4 DNA Ligase	New England Biolabs
Vent DNA Polymerase	New England Biolabs
SOC Outgrowth Medium	New England Biolabs
Gel Loading Dye	New England Biolabs
Agarose	Invitrogen
TAE Buffer	Invitrogen
Ethidium Bromide	Invitrogen
LB Broth Base	Invitrogen
LB Agar Powder	Invitrogen
Ampicillin	Invitrogen
DH5 α	Invitrogen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen

Table 5: List of reagents

STRAINS and MEDIA

Escherichia coli strain DH5 α (New England Biolabs) was used for recombinant plasmids gene cloning. Luria-Bertani (LB) medium (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl, pH 7.5, supplemented with 50 mg/mL ampicillin) was used for bacterial growth and plasmid amplification. LB agar plates (10.0 g/L peptone, 5.0 g/L yeast extract, 5.0 g/L NaCl, 12.0 g/L agar, supplemented with 50 mg/mL ampicillin) was used for bacterial growth and selection.

METHODS

PCR (typical protocol using Vent Polymerase)

COMPONENT	VOLUME (uL)
Thermopol Buffer	5
Vent Polymerase	1
dNTPs	1
Forward Primer	1
Reverse Primer	1
Plasmid DNA	1
Ultrapure Water	Up To 40
Total	50

Table 6: PCR setup

PCR Thermocycling:

COMPONENT	TEMPERATURE (°C)	TIME
Initial Denaturation	94	2 min
30x Denaturation	94	30 sec
30x Annealing	55-62	30 sec
30x Extension	72	45 sec
Final Extension	72	5 min
Hold/Storage	4-12	∞

Table 7: PCR stages

Electrophoresis and Gel Extraction

1. Prepare 0.5 g agarose (or 1 g low-melting temperature agarose) in 50 mL TAE.
2. Melt agarose in microwave for 60 seconds or until completely melted. Cool it under cold water until it is touchable by hand.
3. Pour it in to the gel box, add 2.5 uL ethidium bromid, allow to cool and mix 2.5 µL ethidium bromide, and add the comb.
4. Dilute sample with loading buffer and add to gel (up to 30 µl for 8 well, 10 µl for 15 well).
5. Separate samples using san empty lane unless the product sizes are well separated.
6. Run at 110 V for 55 min (1%) or 1.5 hrs (2%).
7. Cut out bands with a clean razor blade and place samples into microcentrifuge tube to weigh.
8. Extract the bands using the Qiagen Gel Extraction kit as per instructions.

Plasmid Digestion

COMPONENT	AMOUNT
<i>10x NEBuffer</i>	5
<i>Restriction Enzyme(s)</i>	10 units or 1 uL each
<i>DNA</i>	1 ug/1000 ng
<i>Ultrapure Water</i>	up to 50 uL
<i>Total</i>	50 uL

Table 8: Digestion set up

Depending on the selection of enzymes, sequential digestion may be performed. If so, incubate the sample with lower temperature enzyme first (for 12 hours) or per enzyme specification. Then add the second enzyme and incubate at the required temperature for another 12 hours. Perform heat inactivation per enzyme specification.

Purify vectors using electrophoresis followed by QIAquick Gel Extraction Kit.

Purify inserts using Qiagen PCR clean-up kit.

Ligation

COMPONENT	VOLUME (uL)
<i>T4 DNA Ligase</i>	1
<i>T4 Ligase Buffer</i>	4
<i>Insert DNA *</i>	5
<i>Vector DNA *</i>	1
<i>Ultrapure Water</i>	9
<i>Total</i>	20

Table 9: Ligation set up

* = Ideally should have a 3:1 molar ratio of insert to vector.

Pipette the sample to mix and incubate it at room temperature for 4 hours. It can be transformed directly into *E. coli* or frozen and used later.

Transformation into *E. coli* (based on DH5 α from Invitrogen)

1. Place 50 μ L DH5 α cells on ice for 30 minutes.
2. Add 5 μ L containing 1 pg - 100 ng of plasmid DNA to the cell mixture.
3. Carefully flick the tube 4-5 times to mix the cells. Do not vortex.
4. Heat shock for 30 seconds at 42°C, place back on ice for 5 min and then add 950 μ L SOC media.
5. Incubate on shaker (250 rpm) at 37°C for 60 min.
6. Warm LB-Amp plates to room temperature. Mix cells thoroughly by flicking the tube and inverting.
7. Spread each dilution onto selection plate and incubate overnight at 37°C.

If using low transformation efficiency clones, spin sample and plate all cells. Grow overnight at 37°C.

Colony PCR

1. Scrape a single colony using a pipette tip or small ring and dip into 50 μ L ultrapure water.
2. Perform typical PCR except use $\frac{1}{2}$ the volumes listed above and add 5 μ L of the cells suspension. Run PCR as usual with annealing at 55°C. Check for product on a 1% agarose gel.
3. Positive samples can be plated (5 μ L) and grown up (remaining 40 μ L) in 3 mL LB-Amp for later use and sequencing, respectively.

MEDIA RECIPES

LB Broth:

Base media (for 1 L): 25 g LB powder into water, filter sterilize or autoclave.

Supplement: Add ampicillin or other antimicrobial agent (typical concentration of 50 mg/mL) once the media has cooled to ~50°C (touchable by bare hand).

LB Plates:

Base media: 40 g LB agar powder/L into water, autoclave.

Supplements: Add ampicillin or other antimicrobial agent (typical concentration of 50 mg/mL) once the media has cooled to ~50°C (touchable by bare hand).

Pour into individual plates, ensuring no bubble formation.

APPENDIX B: YEAST CULTURE AND RECOMBINANT PROTEIN PRODUCTION

MATERIALS AND REAGENTS:

REAGENT	COMPANY
Dextrose	Fisher
Galactose	Fisher
Sorbitol	Fisher
Agar	Fisher
Casamino Acids	Fisher
Yeast Extract	Fisher
Uracil	Sigma
Bacto-Peptone	Difco
Bacto-Agar	Difco
Yeast N ₂ Base Without Amino Acids	Difco

Table 10: List of reagents for yeast transformation

MEDIA RECIPES

Complex Media: YPD

Base media (for 1 L): 10 g yeast extract, 20 g peptone, 20 g dextrose Filter sterilize

For plates: Autoclave bacto-agar (15 g/L) in 90% of the final volume. Filter carbon source, yeast nitrogen base and amino acids in 10% of final volume and add to autoclaved mixture once it has cooled.

Minimal Media: SD-CAA and SG-CAA

Base media (for 1 L): 20 g dextrose or galactose, 6.7 g Yeast N₂ Base w/o Amino Acids, 5 g Casamino Acids (-ADE, -URA, -TRP), 5.4 g Na₂HPO₄, 7.46 g NaH₂PO₄

Supplements: Uracil (20 mg/L), Tryptophan (20 mg/L), BSA (1 g/L)

Filter sterilize. Note that the buffer salts do not dissolve well, can autoclave in half of the total volume (subtract ~20 ml for the solids volume).

For plates: Autoclave buffer salts, agar (15-18 g/L) and sorbitol (182 g/L) in 80% of the final volume. Filter sterilize carbon source, yeast nitrogen base and amino acids in 10% of final volume and add to autoclaved mixture once it has cooled below 50°C (touchable by bare hand).

STRAINS

BJ5464 or BJ5465: supplement with uracil, as well as tryptophan unless transformed with a pRS-based plasmid.

EBY100: supplement with tryptophan unless transformed with a pCT-based plasmid.

METHODS

Growth

Streak a single colony on YPD or SD-CAA plates and incubate at 30°C for 2 days. Inoculate 3 ml liquid culture with a single colony and grow overnight (YPD) or for 1.5 days (SD-CAA) at 30°C while shaking at 200-225 RPM.

Transformation (Lithium Acetate method from Gietz RD and RA Woods, Methods in Enzymology, 2002)

1. Inoculate 3 ml YPD with BJ5464 (transformation with pRS plasmid) or EBY100 (pCT plasmid) and grow overnight.
2. Heat single-stranded carrier DNA (2 mg/ml) in a boiling water bath for 5 min and then place on ice.

3. Spin yeast at 3000 RPM for 1 min, resuspend in 1 ml ultrapure water and transfer to a microcentrifuge tube. Spin again and discard the supernatant.
4. Layer the following on top of the pellet: 240 ml PEG 3500 (50% w/v), 36 ml LiAc (1.0 M), 50 ml carrier DNA (after vortexing) and 34 ml plasmid DNA (0.1 to 1 mg mixed in ultrapure water, typically use 2 ml). Vortex to mix and resuspend cells (pipette if necessary).
5. Incubate at 42°C for 1 hr before centrifugation and aspiration of the transformation mix. Add 1 ml ultrapure water, resuspend and plate 100 ml on selective media (SD-CAA).

Soluble Expression of Recombinant Protein (small volume assessment).

1. Inoculate 3 ml of SD-CAA+Ura with BJ5464 yeast transformed with a pRS-based plasmid. Grow 24 hrs at 30°C while shaking at 225 RPM.
2. Centrifuge cells, resuspend in 1 ml SG-CAA media, spin, and resuspend in 3 ml SG-CAA+Ura+0.1% BSA.
3. Grow for 3 days at 20°C (alternatively grow for 2 days at 30°C) while shaking at 225 RPM.
4. Recover supernatant by centrifuging at 3000xg for 10 min.