DEVELOPMENT OF A TAT ENABLED ANTIBODY SELECTION PLATFORM WITHIN THE CYTOPLASM OF *Escherichia coli*

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Matthew Ryan Chang

August 2016

ABSTRACT

Antibodies are a critical part of modern medicine, comprising 20% of the novel drugs approved by the FDA in 2015. While there are several platforms available to develop antibodies against a given antigen, they generally take months to carry out or require expensive equipment and extensive technical expertise. Recently, Robinson *et al*. demonstrated that SHuffle T7-Express cells are able to form disulfide bonds within the cytoplasm, enabling the cytoplasmic production of full-length antibodies. Using this technology, this work develops a novel antibody selection platform using chloramphenicol acetyltransferase (CAT) export to the periplasm via the Tat pathway. Within each cell two constructs are expressed within the cytoplasm, an antibody and the CAT-antigen fusion, which is rapidly exported into the periplasm resulting in chloramphenicol sensitivity. In cases where the expressed antibody displays sufficient binding activity with the antigen, the resulting complex will hinder the export of the CAT-antigen fusion, restoring chloramphenicol resistance. In this study we demonstrate that Tat enabled CAT export reliably reports the binding interaction between various antigens and their antibody cognates. We have further validated this approach by utilizing it to identify known antibody-antigen pairs from a mock library with 100-fold excess of non-specific pairs. Finally, one of the successful antibody-antigen pairs was used to construct a CDR-H3 library for screening with this platform and initial results appear promising. We believe that our platform will provide a simple and cost effective method for rapid discovery of clinically relevant antibodies for use with a wide range of diseases and illnesses.

BIOGRAPHICAL SKETCH

Matthew Chang received his B.S. in Chemical Engineering from Purdue University in 2013. In 2014 he entered Cornell University's M.Eng. program in Chemical Engineering, which he completed in 2015 with the highest GPA in the class. While working on his M.Eng. degree, Matthew completed two projects, the first of which was under the guidance of Professor Detlef Smilgies, titled "Using the QCM-D Platform to Detect and Quantify Protein Adsorption and Binding to Various Functionalized Surfaces." The second project, advised by Professor Matthew DeLisa and Michael-Paul Robinson, became the basis for his Master's thesis.

This work is dedicated to my family, especially my parents Paul and Linda Chang for their unwavering support throughout my education

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Matthew DeLisa for his support and providing me the opportunity to complete my MS research in his lab. I would also like to thank Michael-Paul Robinson for his continuous support and mentorship during my two years on this project.

I would also like to thank the entire DeLisa Research Group, especially those in Room 303 for their advice and friendship, without them the past two years would not have been nearly as enjoyable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ab	Antibody
Ag	Antigen
C	Celsius
CAT	Chloramphenicol acetyltransferase
CDR	Complementary determining region
CIAP	Calf Intestinal Alkaline Phosphatase
EDTA	Ethylenediaminetetraacetic acid
IgG	Immunoglobulin G
LB	Luria-Bertani broth
MCS	Multiple cloning site
OD_{600}	Optical density at 600 nm
OE-PCR	Overlap extension polymerase chain reaction
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
SHuffle	SHuffle T7-Express
SOB	Super optimal broth
SOC	SOB + 0.2% glucose
TBS	Tris buffered saline
V	Volts
v/v	Volume per volume
w/v	Weight per volume

CHAPTER 1

INTRODUCTION

1.1 An overview of monoclonal antibodies

1.1.1 Introduction to antibodies

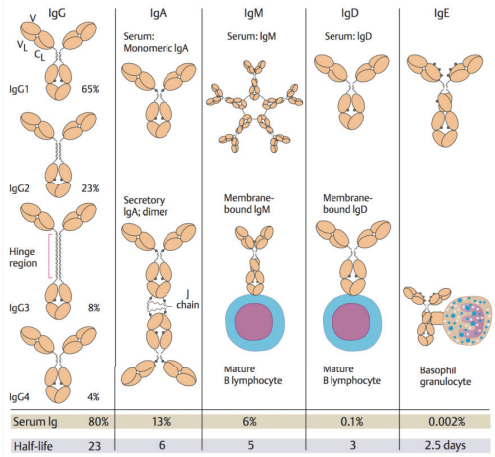

Figure 1: Immunoglobulin isotypes in mammalian cells and their respective half-lives [1]

Antibodies are immunoglobulin glycoproteins secreted by B-cells in the pancreas to identify and neutralize foreign objects in the body. They are composed of two heavy and two light chains, with each chain containing a hypervariable region, which combine to form the antigen binding site. The structure and composition of the amino acids within this hypervariable region work to determine the affinity and specificity of each antibody produced. There are five antibody isotypes produced by mammalian organisms, IgG, IgD, IgE, IgA, and IgM, each differentiated by the heavy chain that it contains. As a major component of the secondary immune response, the IgG comprises around 80% of the antibodies found in human serum and has a half-life of around 23 days, making it a prime candidate for therapeutic development [1].

1.1.2 Origins in medicine

Antibodies made their first appearance in medicine in the format of "serum therapies," which involved injecting patients with blood serum from animals immunized against a disease of interest. Despite the potential to cause fever, rash, and anaphylactic shock, by 1939 this became the standard protocol for the treatment of many infectious diseases, including diphtheria and pneumonia [2]. Following the birth of sulphonamides and other early antibiotics, serum therapy faded to the background and was eventually abandoned as a viable treatment option. Everything changed in 1975 with Kohler and Milstein's revelation of a technique to create unlimited batches of monoclonal antibodies by fusing antibody producing murine spleen cells and immortal myeloma cells. Using this technique the first therapeutic antibody, Muromonab CD3, was developed to prevent the rejection of donor kidneys in transplant patients. Muromonab CD3 selectively binds to the CD3 receptor on T cells, which is vital for antigen recognition and T cell function, thereby blocking the induction of cell-mediated lympholysis and T cell proliferative responses [3]. The hybridomas used to produce this antibody were derived from immunized murine spleen cells and therefore produced fully murine antibodies. In addition to generating a significant immune response, these murine antibodies had a short circulating half-life of 15-30 hours, leading to the necessity of frequent doses of high concentrations of the antibody [2][4]. Faced with these challenges, scientists looked to improve antibody technology by developing chimeric, humanized, and eventually fully human antibodies to avoid the downfalls associated with murine-based antibody structures. With these improvements and the development of novel high throughput screening methods, antibodies have become the most prevalent product in the biopharmaceutical market, with a broad range of indications, including transplant rejection, cancer, infection, respiratory disease and many others.

1.1.3 Current antibody therapeutics

One of the most prolific antibody therapeutics to date is the TNF-alpha class of antibodies. There are two major antibodies in this class, infliximab (Remicade) and adalimumab (Humira), both binding the TNF-alpha cytokine, thereby inhibiting its ability to bind with the canonical TNF-alpha receptors. By blocking this interaction, these antibodies interfere with the body's natural inflammatory response, making these two drugs potent treatments for arthritis, Crohn's disease, ankylosing spondylitis, and other autoimmune disorders. The major difference between these two antibodies is that Remicade is a chimeric antibody whereas Humira is the first fully human antibody to achieve regulatory approval [5].

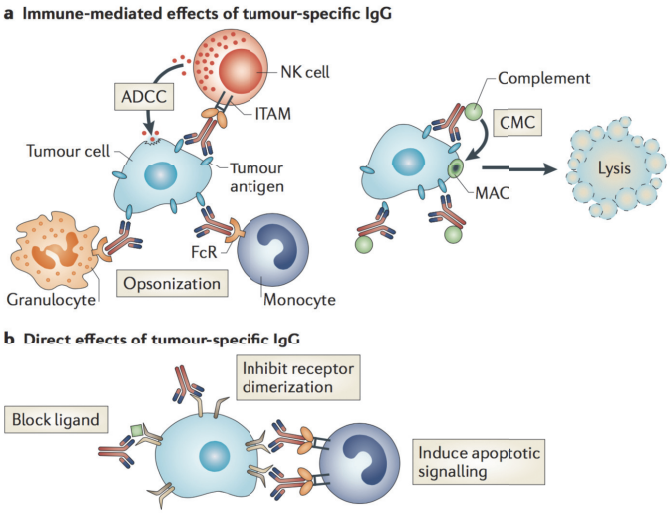

Figure 2: Biologic mechanisms of antibody induced apoptosis in cancerous cells [7]

Another field in which antibodies are currently playing an important role in is the development of modern oncology treatments. Over the past 20 years, monoclonal antibodies have become one of the most successful methods of treating hematological and solid tumor malignancies, and as a more detailed understanding of cancer biology is developed, new antibody targets will continue to be revealed [6]. As it stands, most anti-cancer antibodies are targeted towards antigens in one of seven categories: cluster

differentiation (CD) antigens (CD20, CD30, CD33), glycoproteins (EpCAM, CEA, gmA33), glycolipids (GD2, GD3), carbohydrates (Lewis-Y), vascular targets (VEGF, aVb3), growth factors (EGFR, HER2, EPhA3), and stromal and ECM antigens (FAP, Tenascin) [6]. There are three main mechanisms that therapeutic antibodies can utilize to kill cancer cells. The simplest pathway is the direct killing of the cell via the antibody binding to surface cell receptors, inhibiting cell signaling, proliferation, and eventually leading to apoptotic events [8]. The second method relies on the effector function of antibodies to activate specific cellular responses. Since antibodies are a natural part of the immune system, when they bind to cancerous cells, they are recognized by natural killer calls and macrophages, leading to antibody dependent cellular cytotoxicity (ADCC) and phagocytosis. A second form of immune reaction with antibodies is complement-dependent cytotoxicity, which requires the antibody-antigen complex on the cell surface to also interact with C1q, leading to a cascade of serum proteases [9].

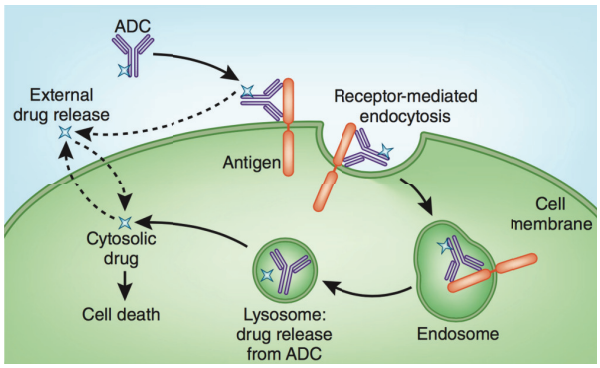

Figure 3: Antibody-drug conjugates deliver cytotoxic payloads to cancerous cells [10]

The final anti-cancer mechanism is the targeted delivery of cytotoxic payloads via antibody-drug conjugates (ADC). Monoclonal antibodies are linked to cytotoxic payloads via a cleavable linker, allowing the drug to be released in the presence of cancerous cells while maintaining stability in the blood stream [10].

1.1.4 The structure of an antibody

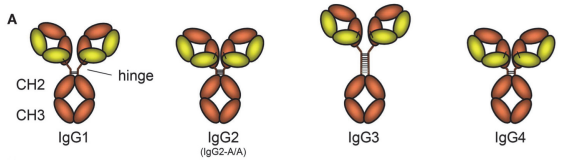

Figure 4: IgG subtypes and the disulfide bond pattern in the hinge region [11]

IgG's, (hereafter "antibodies" refer to their IgG1 format unless otherwise noted) are roughly "Y" shaped molecules that are comprised of two identical heavy (50 kDa) and light (25 kDa) chains. The heavy chain is comprised of three constant domains, CH1, CH2, and CH3, and one variable domain, VH. There are four heavy chain isotypes, each differing in the number and position of disulfide bonds in the hinge region. These four isotypes lead to the four IgG subtypes, IgG1, IgG2, IgG3, and IgG4, with each one having different effector functions and biophysical characteristics [11]. Each heavy chain also includes a highly conserved glycosylation site at Asn297 on the CH2 domain, which is important for stabilization of the CH2 domain as well as activating different effector functions [12]. The hinge region contains four interchain disulfide bonds, two between the heavy chains and one for each heavy-light chain pair.

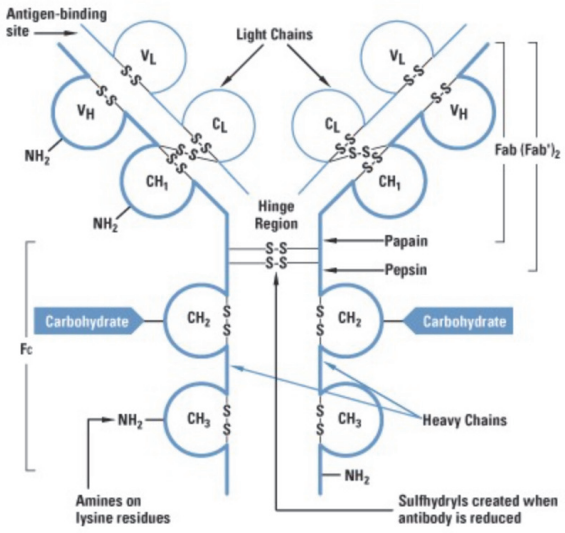

Figure 5: Basic structure of an IgG1 [13]

In addition to the four interchain bonds, there are 12 intrachain disulfide bonds that must be formed. The light chain is a compacted version of the heavy chain, comprised of one constant (CL) and one variable (VL) domain. The hypervariable regions (complimentary determining regions, CDR) are located on the ends of the variable domains on both the heavy and light chains. Each variable domain contains three CDRs, located around residues 32, 55, and 98 on the VH and residues 30, 50, and 98 for the VL [14]. The amino acid composition and loop structure within these domains determine the antigen specificity and affinity for each antibody, therefore single amino acid mutations in this region can block or restore antibody binding. Due to the symmetrical nature of antibodies, each antibody contains two independent antigen binding fragments. With native antibodies the binding fragments formed by the variable regions are typically identical, however modern antibody engineering techniques have allowed for the creation of bispecific antibodies containing two different binding regions [15].

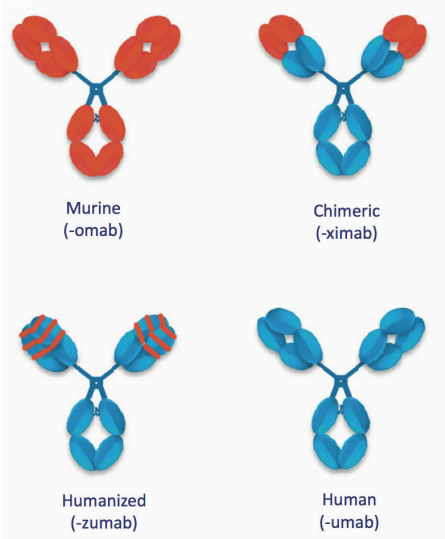

Figure 6: Murine, chimeric, humanized, and human antibody frameworks [16]

Original therapeutic antibodies were discovered by hybridomas generated from immunized mice. Since these antibodies were not human by nature, they tended to have a short circulating half life as well as cause severe immunogenic reactions following administration. In an effort to counter these adverse effects, the binding regions of the murine derived antibodies were grafted onto the constant regions of a human antibody scaffold, thereby reducing the immunogenicity of the therapeutic antibody. The first grafting strategy took the entire variable region from the murine antibody and simply used it to replace the variable region of a human antibody. The next step was grafting only the CDR regions from the murine antibody onto the human antibody, resulting in an antibody that was nearly entirely humanized. The final step in this progression was to use screening methods to identify fully human antibodies, with no murine component [17] [18].

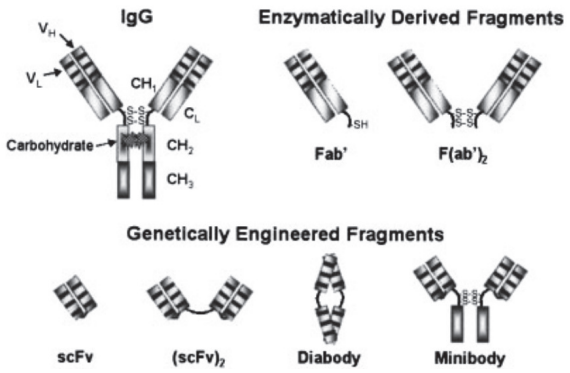

Figure 7: Antibody fragments derived from the antigen binding site [19]

A full length IgG can be broken down into various functional antibody fragments, starting with the hinge region. The first major functional fragment is the antigen binding fragment, or the Fab. The Fab consists of the entire light chain and the CH1 and VH regions of the heavy chain with a disulfide bond holding them together. The next fragment is achieved by removing the constant regions from the Fab, thereby generating an antibody fragment comprised of only the variable regions. This is

known as an scFv and it is the smallest binding unit derived from an antibody. Other fragments such as $F(ab)_2$ and $(scFv)_2$ can be formed, however they are typically combinations of these two basic fragments [18]. There are some advantages to smaller fragments, such as increased initial tumor penetration and the ability to bind certain cryptic epitopes such as immuno-evasive pathogen glycoproteins [20]. However, while smaller fragments rapidly penetrate tumors, they are also susceptible to rapid clearance, therefore it has been found that after 72 hours, IgG penetration depth was nearly identical to that of the scFv [19].

Another drawback of only using antibody fragments is that you loose the highly conserved Fc (CH2, CH3) region of the antibody, which is used to communicate with the body's immune system. Therefore when using an abbreviated antibody fragment, the ability to elicit effector functions such as antibody dependent cellular cytotoxicity and complement-dependent cytotoxicity is lost [21]. The Fc region is also important for the antibody recycling pathway. Within the low pH of an endosome, the antibody is bound by the neonatal fc receptor, which removes and protects the antibody from entering the lysosomal degradation pathway. This recycling process enables antibodies to have a long half life within the body, decreasing the need for repeated doses [16].

1.2 Current screening methods

The earliest method of developing novel antibodies involved exposing a mouse to the desired antigen, harvesting the spleen cells, and fusing the harvested cells with myelomas to create hybridomas. These hybridomas can then be screened to identify the antibodies with the desired characteristics before scale up production. While effective, this method typically takes months to carry out and therefore more rapid discovery methods are needed [22]. When screening a library, one of the most

challenging components is the coupling of the phenotypic and genotypic identities. With this in mind, there are four current library-screening techniques that have overcome this challenge; phage, yeast, ribosome, and bacterial displays.

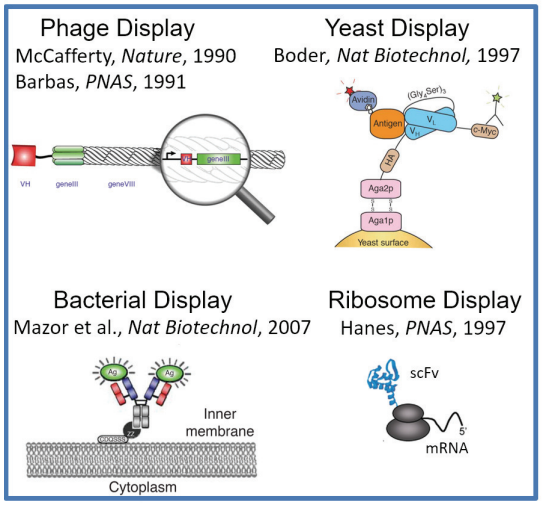

Figure 8: Overview of common antibody library screening methods

Phage display, developed in 1990 by McCafferty *et al.,* involves encoding scFvs onto the N-terminal portion of the gene III protein of a bacteriophage. Normally there are 4 copies of the gene III protein expressed on each phage, and it is responsible for the biding of the phage to the bacterial f-pilus. The library containing phages are passed over the antigen, which removes the phage that contain antibody fragments specific to that antigen. The phage can then be eluted, amplified, and the antibody fragments can be further characterized [23] [24]. While phage display is highly effective and can be used to screen massive antibody libraries, the drawback is that it is currently suitable for only scFv formats since displaying the full length IgG is not currently practical.

Yeast display is one of the few high throughput screening methods that are able to use full length IgGs. Since yeast are eukaryotic cells, they naturally form the disulfide bonds necessary for the assembly of full length IgGs. The Aga2p protein is normally used for cell-cell interactions, therefore when the antibody library is linked to this

protein, the yeast cell naturally extends it out and away from the cell body, making it available for antigen binding and decreasing the chances that it binds a marker on the yeast cell [25]. The eukaryotic nature of this platform makes it easier to screen full length IgGs, however this can also cause some issues since yeast have their own glycosylation patterns which could affect the antibody performance once it is removed from the yeast system.

Ribosome display was developed by the Plückthun group and is the only cell-free system to have widespread acceptance. With this method antibody library mRNA is translated *in vitro*, however before completion of the reaction and release of the mRNA strand, the ribosomes are cooled on ice and the ribosomal complexes are stabilized via magnesium ions. These ribosome complexes are now attached to the mRNA while displaying the translated scFv in its fully functional form [26]. One of the major advantages for this step is that the library size is not limited by any transformation efficiency, rather by the number of ribosomal units that are in the system, therefore library size is functionally unlimited.

The final method of antibody screening is bacterial display, developed by Mazor *et al.* in 2007. Fully translated heavy and light chains are exported into the periplasm of *E.coli*, where they are able to assemble into the full length IgG. Once assembled, the IgGs are captured by an Fc binding protein that is tethered to the inner membrane. Permeabilizing the outer membrane reveals spheroplasts that have full length IgGs tethered to the membrane available for antigen binding. These spheroplasts are then exposed to the antigen so that high binding antibodies can be identified for further characterization [27]. One of the issues of this technique is that it can develop high binding aglycosylated antibodies. Since these antibodies are screened and selected in

an aglycosylated format, it is difficult to predict how producing these antibodies in the glycosylated form via mammalian expression systems would affect antibody function or stability.

1.3 SHuffle T7-Express strain

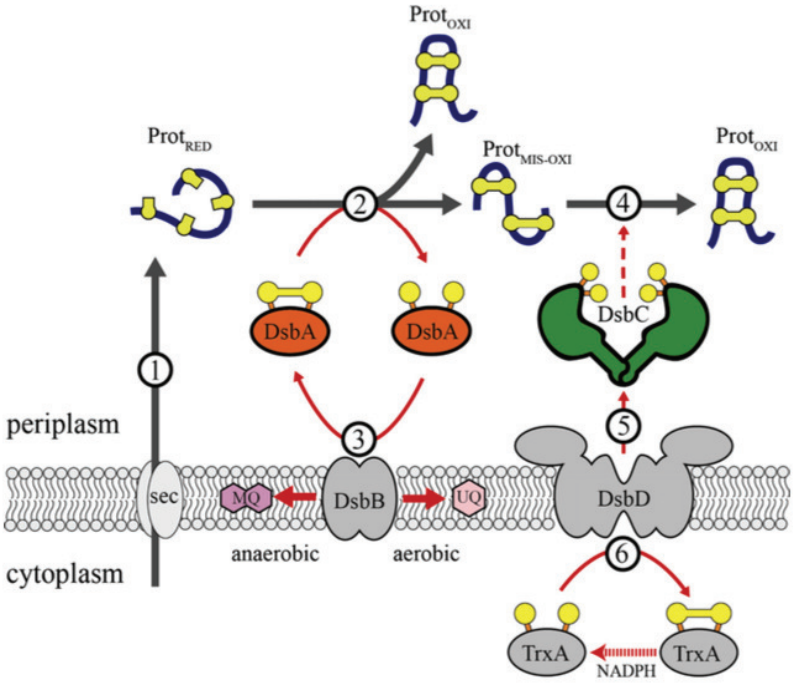

Figure 9: Periplasmic disulfide bond formation via Dsb family of proteins [31]

The cytoplasmic space in *E.coli* contains a high concentration of reducing enzymes, therefore this space is not conducive to the oxidizing nature of disulfide bond formation. Disulfide bonds are generated when the thiol groups between two cysteine residues are oxidized, forming a covalent bond. There are three events that a disulfide bond can undergo, reduction, oxidation, and isomerization, and all of these processes require the involvement of an oxioreductase to catalyze the electron transfer activity. In wild type *E.coli*, the enzymes required for disulfide bond formation are located in the periplasm, therefore initial attempts at recombinant expression of disulfide bond containing proteins required the fusion of export sequences that brought the unfolded (Sec pathway) or folded (SRP pathway) protein into the periplasm so the bonds could

be formed [28]. Disulfide bond formation in the periplasm is governed by the activity of five enzymes in the Dsb protein system (DsbA, B, C, D, G) [29]. DsbA is a strong oxidase, therefore it rapidly forms disulfide bonds with consecutive cysteine residues as they enter the periplasm. In order to ensure that DsbA activity remains high, the integral membrane protein DsbB is required to regenerate the active site via oxidization. While this DsbA/B combo is efficient for the formation of consecutive disulfide bonds, proteins requiring bonds between non-consecutive residues requires the presence of the V-shaped disulfide bond isomerase, DsbC, which rearranges existing disulfide bonds to achieve the correct protein structure [30] [31]. While periplasmic expression of disulfide bonded proteins is feasible, the main drawback for these platforms is the yield that can be achieved. When using a periplasmic based system, the yield is limited not by the amount of protein that can be translated, rather the rate at which the translated protein can be exported into the periplasm [28]. In order to overcome this challenge, scientists looked to develop a method to generate disulfide bond containing proteins within the cytoplasm of *E.coli*.

Within the cytoplasm of wild-type *E.coli*, transient disulfide bonds may spontaneously form, however they are rapidly oxidized by any one of the numerous reductase enzymes or small molecules present. Due to the abundant nature of these reducing agents, the cytoplasm has traditionally been considered a reducing environment, and is maintained in that state via the glutaredoxin and thioredoxin pathways, which in turn maintain their reductive potential via the cytoplasmic presence of NADPH [31]. The thioredoxin pathway is used to maintain proteins in their reduced state within the cytoplasm, protecting proteins from oxidative damage or aberrant enzymatic activity. In order to accomplish this, thioredoxin reductase (trxB) maintains two thiroredoxins (Trx1 and Trx2) in a reduced state, allowing them to rapidly reduce any disulfide bond

that is formed. The glutathione pathway works in complement with the thioredioxin pathway, with the glutathione reductase maintaining GshA and GshB in their reduced form. These two enzymes are used to generate reduced glutathione (GSH), a small tripeptide component of the cytoplasmic redox buffer. Within the cytoplasm, GSH directly reduces disulfide bonds by forming a GSH-protein complex, which can then be resolved by a glutaredoxin protein (Grx1, Grx2, Grx3), resulting in the reduced protein and an oxidized GSH, which can then by restored to the reduced form via glutathione reductase [32].

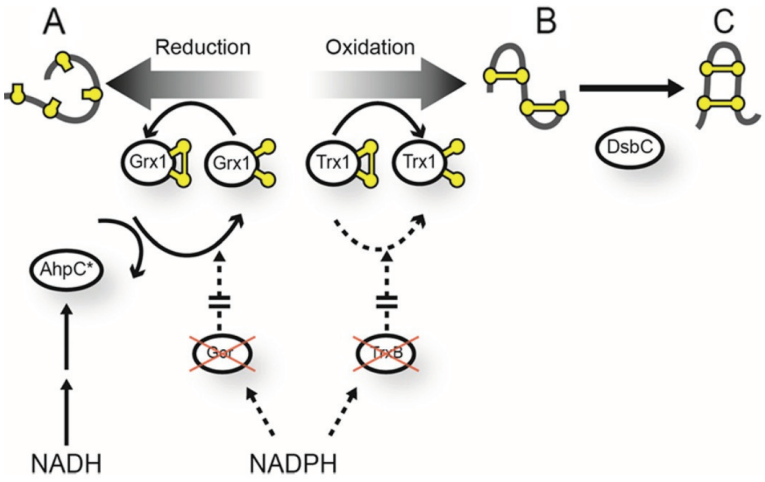

Figure 10: Pathway for the formation of disulfide bonds in the cytoplasm of SHuffle cells [31]

In order to eliminate the reducing nature of the cytoplasm, both of these pathways need to be knocked out. However strains with both the *trxB* and *gor* pathways deleted are non-viable since these reduction pathways are critical for the function and recycle of some essential proteins such as ribonucleotide reductase. Following work with the double knockout strain, Berkman *et al.* discovered that the *ΔtrxB Δgor* strains can be revived by a mutation in the peroxidase AhpC pathway [30]. Normally exhibiting peroxidase activity, these mutant AhpC* proteins were able to reduce Grx1, thereby partially restoring the reducing potential of the cell and providing a pathway for the

reduction of essential proteins required for survival. While the AhpC* mutation allows for the generation of reduced Grx1, the *ΔtrxB* knockout allows thioredoxins to remain in their oxidized form, allowing them to function as disulfide bond catalysts. Strains with these three mutations *ΔtrxBΔgor,aphC** have been designated FÅ113 and are commercially developed by Novagen under the trade name Origami. While the FÅ113 strains are able to efficiently create disulfide bonds within the cytoplasm, the native disulfide bond isomerase DsbC is only expressed in the periplasm, therefore there is limited ability to correct misfolded proteins within the cytoplasm [30]. While the ability to create only consecutive disulfide bonds would be suitable for the production of proteins containing simple disulfide bond patterns such as alkaline phosphatase or human growth hormone, they are very inefficient in the production of proteins requiring many complex and non-consecutive bonds, such as antibodies. To overcome this shortcoming, DsbC was expressed without its native signal sequence, causing it to remain in the cytoplasm and providing a mechanism to correct misfolded proteins. The *ΔtrxBΔgor,aphC** strain with the addition of cytoplasmic DsbC has been marketed as SHuffle T7 Express by New England Biolabs [30][31].

1.4 Inner membrane translocation pathways

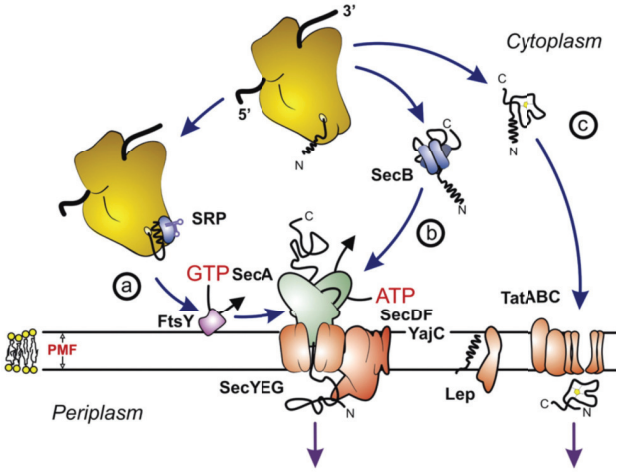

Figure 11: Mechanism of the Sec and Tat pathways [33]

There are two distinct pathways for the export of cytoplasmic proteins into the cytoplasm, each with a unique mechanism and substrate selection protocol. The majority of protein translocation is carried out by the general secretory (Sec) pathway, a proteinaceous conduit formed by an oligomeric assembly of the heterotrimeric membrane protein SecYEG and the ATPase SecA as the molecular motor. Since substrates typically thread their way through the membrane pore generated by the SecYEG, an unstructured protein is required [33]. The second major export pathway is the twin-arginine translocation (Tat) pathway, and though it has a limited number of areas homologous to the Sec pathway, the mechanism and machinery are similar to that found in the ΔpH driven protein import pathway found in plant thylakoids [34]. The Tat pathway consists of two (Tat A/Tat C) or three (TatA/TatB/TatC) membrane-integrated subunits combining to form the receptor and conducting machinery. The fundamental difference between these two pathways is the conformational preference of the substrate. The Sec pathway secretes an unfolded protein into the periplasm where it can be folded, whereas substrates for the Tat pathway fold within the cytoplasm before being shuttled to the periplasm. The Tat pathway is a critical component for pathways that require the translocation of cofactor bound proteins across the inner membrane, including the global nitrogen cycle, photosynthetic pathways involving cytochrome bc1 complexes, and many more [*33*].

In order for Tat translocation, substrates require an N-terminal signal sequence, comprised of a positively charged amino-terminal, a hydrophobic core, and a polar carboxyl-terminal region. In addition to these general building blocks, Tat signal sequences are easily recognized by the highly conserved pair of arginine residues between the amino- and hydrophobic cores, with a typical format resembling Z-R-R-x- Φ - Φ, where Z is any polar residue and Φ is any hydrophobic residue [33][35]. By

virtue of the fact that the Tat pathway only exports folded proteins, there is an inherent quality control aspect required in order for the pathway to reject partially or unassembled proteins while still selecting for the fully folded proteins and oligomers. While the mechanism behind this process is relatively unknown, some theories exist concerning the use of molecular chaperones blocking the signal sequences or the need of proteins to hide hydrophobic cores before translocation is possible [36].

1.5 Chloramphenicol and CAT

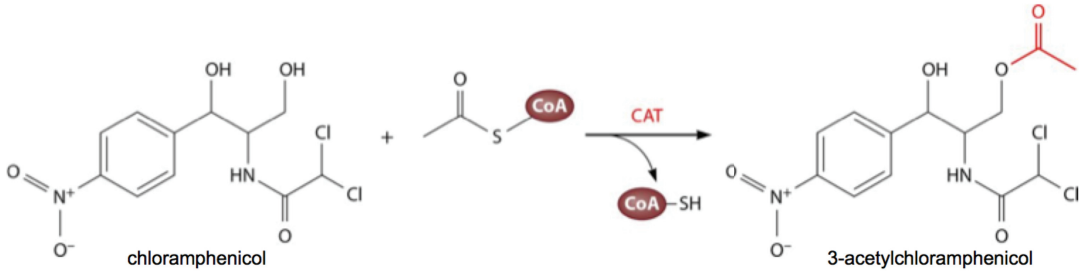

Figure 12: Mechanism action for chloramphenicol acetyltransferase [37]

Chloramphenicol (Cm) is a bacteriostatic antibiotic that exists as an amphipathic and uncharged molecule allowing it easy passage through biological membranes. Due to the stereochemical conformations around the C1 and C2 carbons, four species are possible, however only the D-threo conformation exhibits antibiotic activity [38]. Chloramphenicol inhibits bacterial growth by interacting with the 23S rRNA of the 50S subunit and inhibiting the peptidyl transferase center, thereby halting protein synthesis [39]. One mechanism that has evolved in resistant bacteria is the development of the chloramphenicol acetyltransferase protein (CAT), which acetylates the C-3 hydroxyl [37] [40], rendering the compound inert. The potency of chloramphenicol is reliant on its ability to bind and interact with the ribosomal subunits, therefore in order to maintain chloramphenicol resistance, the CAT enzyme must remain in the cytoplasm to continuously deactivate the rapid influx of chloramphenicol molecules.

1.6 This work

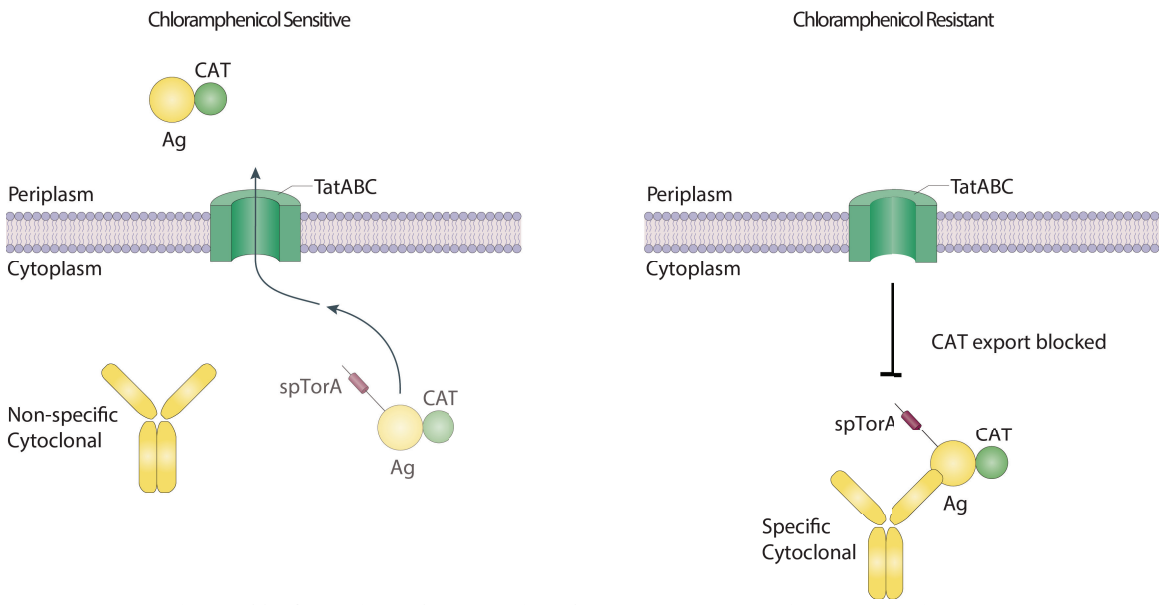

Figure 13: Overview of mechanism for Tat enabled antibody selection

The overall aim of this work is to develop a novel *E.coli* based antibody selection platform based upon the successful or blocked transport of the CAT-Ag fusion into the periplasm via the Tat pathway. If exported to the periplasm, the CAT protein will be unable to neutralize the chloramphenicol molecules and the cells will become sensitive to chloramphenicol. On the other hand, if the export of the CAT fusion is blocked, the CAT protein will be able to neutralize the chloramphenicol molecules before they can inhibit protein synthesis, allowing those cells to survive. The differential survival generated by blocking or exporting the CAT fusion will enable this platform to select novel IgG antibodies against any biological antigen that can be expressed by the cell.

In order to create this platform, three aims need to be accomplished before applying the technology to an antibody library selection. The first aim will seek to establish the basic principle upon which all future work will be based; blocked CAT export is a selectable trait. For this test, the CAT protein will be fused to an N-terminal Tat signal peptide (ssTorA) or the mutated signal peptide (ssTorA(KK)) with the twin arginine

residues replaced with a twin lysine motif, which will be used as an analog for blocked export. The second aim involves identifying antibody-antigen pairs that can be used to demonstrate the effectiveness of this selection by displaying differential growth patterns when plated on LB-Agar containing chloramphenicol. Antigens of various size and structure will be fused to the ssTorA-CAT protein and coexpressed with the matched antibody in an effort identify a set of proteins that are suitable for this system. This set will provide information regarding the characteristics of proteins that can be successfully exported by the Tat pathway and blocked by the binding action of the antibody. The final aim will be to conduct mock library studies as a final proof of concept for the platform. The mock libraries will also provide initial information regarding the conditions required for selection and robustness of the technology.

The future goal for this project is the generation of an antibody library and the application of that library to this selection platform. In this work, the two antigens that will be utilized in the first round of selections are Fos and Gcn4, both of which are basic leucine zippers. Primers carrying either the NDT or NNK degeneracy will be used to generate a CDR-H3 library based off the anti-Gcn4 IgG. With the Gcn4 selection, the goal is to recover the parent CDR-H3 as well as any other candidates that may exist. Switching the affinity of the anti-Gcn4 IgG to become an anti-Fos IgG will be the final effort to demonstrate the effectiveness of this selection platform.

CHAPTER 2

Materials and Methods

2.1 Plasmids

Table 1: List of plasmids used in this study

	Description	Source
pBAD24	Pbad promoter (Arabinose), araC regulator, pBR322 ori, Amp^R	[41]
pCOLADuet	Bicistronic T7 promoter (IPTG), Lac operon, ColA origin, Kan^R	[42]
pET28a	N-terminal 6xHis tag, T7 promoter, pBR322 origin, Kan^R	[43]

This work utilized two general plasmid formats. The antigen plasmid was based off of a pBAD24 vector with an n-terminal twin-arginine translocon signal sequence linked to the chloramphenicol acetyltransferase gene followed by the antigen of interest via a short peptide linker. The antibody plasmid utilized the bicistronic pCOLADuet vector, with the light chain in the first MCS and the heavy chain in the second. The pBad24 vector carries an ampicillin resistance marker whereas the pCOLADuet vector confers kanamycin resistance. The pET28a vector was used for the purification of antigens for use in ELISA and future SPR experiments.

2.2 Bacterial strains

Table 2: *E.coli* strains used in this study

	Description	Source
DH5alpha	Standard DNA cloning strain	
SHuffle T7-Express	*ΔtrxBΔgor aphC*DsbC*	[30][31]
BL21 (DE3)	Antigen expression for purification	

2.3 Plasmid construction

Various PCR strategies were used to develop the expression plasmids used in this work. The primary PCR technique uses a forward and reverse primer to amplify a desired gene while adding the appropriate restriction enzyme sites. For PCR of genes from ~500-1200 bp, Vent polymerase was used, outside of this range, Phusion polymerase was preferred. For some of the inserts required, one PCR was not suitable to amplify the entire gene, therefore overlap extension PCR was implemented. The desired gene was divided into two pieces and each piece was amplified using traditional PCR, with the primers for the middle of the gene containing an overlap region that is common to both fragments. Following the first PCR, 50 ng of each product was used as the template for the OE-PCR. At this step, only the forward primer for the first fragment and the reverse of the second fragment are used. For all OE-PCR regardless of the primers, the annealing temperature used is 68-70 C. Once the gene was created, both the gene and the vector were digested with the desired restriction enzymes following the NEB recommended protocol. Following digestion, CIAP was added to each vector to remove the 5' phosphate, reducing the likelihood of

self-ligation. To confirm digestion and remove the CIAP, the digested vector was ran on a 1% (w/v) agarose gel for 20 minutes at 105 V and then gel extracted. Ligation reactions were done with T4 ligase for either 1 hour at room temperature or overnight at 16 C. Following transformation, all sequences were confirmed via sequencing analysis.

2.4 Calcium chloride transformation

To create calcium chloride competent cells, the desired cell line was cultured overnight at 37 C in Luria–Bertani (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl and NaOH to pH 7.0) with the appropriate antibiotics. The next morning, the culture was used to inoculate fresh LB at a dilution of 1:100 and the culture was grown to an OD_{600} between 0.3 and 0.7, with lower OD_{600} values generating slower growing cells with a higher degree of competence and the higher values producing less competent but faster growing cells. Once at the desired OD_{600} value, the sample was cooled on ice for 15 minutes before being centrifuged for 10 min at 3000 rpm and 4 C. The media was separated from the pellet and the pellet was resuspended in the appropriate volumes of ice cold 0.1M $CaCl_2$ and 50% (v/v) glycerol. For each 10 ml of culture, the pellet was resuspended with 700 ul of $CaCl_2$ and 300 ul of glycerol. Cells were then aliquoted into 1.5 ml Eppendorf tubes and stored at -80 C until needed. For transformation, the cells were thawed on ice for 10 minutes and then mixed and incubated on ice with the plasmid or ligation product for 10 – 30 minutes. Following this incubation, the cells were placed in the 42 C water bath for 35 seconds and then placed back on ice for 2 minutes. 950 ul of SOB was added to the transformed cells before they were incubated at 37 C for one hour. The cells were then pelleted and streaked out on LB-agar plates with the appropriate antibiotics and incubated at 37 C overnight.

2.5 Electrocompetent cell transformation

Two methods of generating electrocompetent cells were used in this work, however they only differ in the type of media used and the target OD. The first method generates standard electrocompetent cells, whereas the second method generates highly competent cells suitable for library generation [44] [45]. DH5alpha cells are cultured in 5 mL of fresh LB (YENB for method 2) and allowed to grow overnight at 37 C. The following morning the cells are subcultured 1:100 for the first method and 1:1000 for the second into fresh media. The difference here is due to the target OD_{600}, with the first method looking for an OD_{600} around 0.5 while the second methods requires an OD_{600} of 0.2-0.3. Once at the target OD_{600}, cells are then cooled on ice for 1 hour before centrifugation for 15 min at 3000 rpm and 4 C. Following centrifugation, the media is removed, the cells are washed three times with 10% (v/v) cold glycerol, and then pelleted again. After the final wash, they are resuspended in the appropriate amount of 10% glycerol and aliquoted for future use.

Prior to transformation, it is critical to desalt any ligation product or plasmid solution. This is done by a floating a dialysis membrane in a plate of Millipore water, pipetting two times the transformation volume onto the membrane and allowing it to desalt for 20 minutes. Once desalted, the DNA can be mixed with a thawed tube of competent cells and then the cells can be transferred to an electroporation cuvette for electroporation via an electroporator. Immediately following electroporation, the cells are mixed with SOB and placed in a 37 C incubator for 1 hour. Following the recovery, the cells are plated on LB-agar with appropriate antibiotics and grown overnight at 37 C.

2.6 Antibody/Protein expression

Transformed colonies were used to inoculate 3-5 ml of LB supplemented with the appropriate antibiotics, and grown overnight at 30 °C. The following day, the desired volume of fresh LB supplemented with the appropriate antibiotics is inoculated 1:100 with the overnight culture and then grown until the OD_{600} was between 0.6 – 1.0. At this point, the cultures were induced with either 0.1 mM isopropyl b-D-thiogalactopyranoside (IPTG) or 0.4% L-arabinose. Following induction, cultures were incubated overnight at 30 C before centrifugation and protein extraction.

2.7 Antibody/Protein extraction

Following overnight incubation at 30 C, the culture was pelleted, the media was removed, and the cells were resuspended in lysis buffer (PBS or TBS + Halt protease inhibitor + 0.5 mM EDTA) or the column buffer if column purification was desired. The cells were then lysed via sonication (≤6 ml) or homogenization (≥6 ml) depending upon culture volume and the cellular debris was pelleted via centrifugation at 15 krpm. The supernatant was separated from the pellet and stored as the soluble fraction.

2.8 Antibody/Protein purification

Before use, capture resins are equilibrated to room temperature and flushed with 10 column volumes of column buffer to remove any storage buffer in the system. The soluble fraction was filtered with a 0.2 um syringe filter to remove any remaining particulate matter and then applied to the resin. For the protein A column, the supernatant was incubated and rocked with the column for 1 hour whereas the Ni-NTA and Amylose columns utilized a constant flow methodology in which the flow through was applied to the column multiple times. Following the binding step, the column was washed with 10 column volumes of column buffer before elution. The

elution is captured in 500 ul fractions with the concentration of each fraction determined by the NanoDrop. Once all of the protein has been eluted, the column is regenerated and stored in the appropriate conditions.

2.9 Subcellular fractionation

The fractionation protocol was adapted from Quan et al. [44]. The strains of interest were cultured overnight at 30 C in LB with the appropriate antibiotics. The following day they were subcultured at ratio of 1:100 into fresh LB with the appropriate antibiotics and grown to an OD_{600} of 0.7-1.0 at which point they were induced with 0.4% arabinose for antigen expression or 0.1mM IPTG for antibody expression. Once induced, the cells were incubated for 16 hours at 30 C before harvesting. To harvest the cells, the cultures were pelleted via centrifugation at 3000 g for 20 min at 4 C. The media was aspirated off and the pellet was gently resuspended with an inoculating loop in 1 ml tris-sucrose-EDTA solution. The resuspended cells were incubated on ice for 30 min before being gently transferred to an Eppendorf tube for another round of centrifugation at 16,000 g for 30 min at 4 C. After centrifugation, the supernatant was removed and stored as the periplasmic fraction. The pellet was then resuspended in TBS and sonicated to lyse the cytoplasmic compartment. Following sonication, the solution was centrifuged at 16,000 g for 30 min at 4 C to pellet the cellular debris and the supernatant was removed as the cytoplasmic fraction.

2.10 Spot plating

Single colonies of transformed SHuffle cells are cultured overnight in 3 – 5 ml of LB media with the appropriate antibiotics. The following day, the OD_{600} for each culture is measured and the samples are normalized to an OD_{600} of 1 by pelleting the appropriate culture volume and resuspending in 500 ul of fresh LB media. 180 ul of

fresh LB is pipetted into each well of a 96-well plate, after which 20 ul of sample is loaded into the first well in each column. 10 fold serial dilutions are then made by taking 20 ul from the first column and mixing it with the column directly below it, repeating until all of the wells have sample. Since the first row is the first dilution from the normalized sample, it has been designated 10^{-1}, the second row is 10^{-2}, and so on. It is important to change the pipette tips between each dilution to maintain the correct concentrations. LB-Agar plates are made with varying concentrations of chloramphenicol, typically 0 – 50 ug ml^{-1}, with the control plate containing the antibiotics to select for each plasmid. 5 ul of each sample is taken with a multichannel pipette and spotted on the plate in the corresponding location. If the spots are done starting with the least dilute samples then the tips do not need to be changed between dilutions. After spotting all of the samples, the plates are left to dry and then placed in 30 C for 48 hours, with imaging every 24 hours.

2.11 SDS-PAGE analysis

Extracted protein, in purified or raw lysate form, is mixed with the appropriate loading dye (plus 10% β-mercaptoethanol for reducing samples) heated at 70 C or 100 C for 15 minutes for Novex or Biorad gels respectively. Samples are typically prepared as a 60 ul solution with a protein concentration of 1 mg ml^{-1} providing a convenient means to load varying protein amounts for each sample in consecutive lanes. While the sample is incubating, the gel is submerged in running buffer (NuPage MOPS running buffer: 50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7 or Biorad TGX running buffer: 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) and the wells are flushed to remove any residual storage buffer. After loading the samples, the gel is ran at 200 V for 45 minutes or until the dye front has ran off the bottom of the gel. The gel is removed from the plastic casing and rinsed three times in Millipore water for

five minutes each. The sample is then covered with 25 ml of Bio-Safe Commassie G-250 stain and gently rocked for 1 hour. The dye is then replaced with water and the gel is rocked for 1 hour or until the bands are clearly visible and the background has been reduced to an acceptable level.

2.12 Western blot analysis

For Western blot analysis, samples are prepared and ran in an SDS-PAGE gel as described previously. The PVDF transfer membrane (8 x 7 cm) and thick blotting paper (9 x 8 cm) are cut and the PVDF membrane is activated via a 10 second soak in methanol, followed by 2 minutes in Millipore water then 15 minutes in transfer buffer. Once the SDS-PAGE gel has completed the run, the transfer stack is generated following the procedure detailed on the transfer apparatus (bottom plate, sponge, thick, gel, membrane, thick, sponge, top plate) and the transfer is ran at 20 V for 60 minutes for the Novex apparatus or 14 hours at 30 V for the Biorad one. Once the transfer is complete, the membrane is washed in PBS for 15 minutes before a 3-6% milk/PBS blocking solution is added and the membrane is rocked for 1 hour to block the non-transferred regions of the membrane. The probing antibody is mixed with a 3% milk solution and once the membrane blocking is complete, the milk is poured out and replaced with the probing solution and allowed to gently rock for 1 hour. After an hour with the antibody, the membrane is ready to be washed to remove any excess probing antibody. During this was step, the buffer is replaced every 7 minutes for an hour before it is shaken dry and the clarity substrates are added to the membrane before imaging.

2.13 Enzyme-linked immunosorbent assay (ELISA)

Costar 96-well high binding ELISA plates were coated at 4 C overnight with 50 ul of a 10 ug ml^{-1} antigen solution in a 0.05M sodium carbonate buffer (pH 9.6). For this work, antigens were typically expressed in a BL21 cell transformed with a pET28a plasmid containing an N-terminal 6x-His tag and a maltose binding protein fused to the antigen of interest, allowing for purification via a Ni-NTA or amylose resins. The next day, the wells were emptied via inversion, and blocked with 270 ul of 5% (w/v) milk in PBS for at least one hour with moderate shaking at room temperature. Wells were then emptied by inversion and quickly washed with 100 ul PBST. The samples, either raw lysate or purified antibody solutions, were then added following a serial dilution pattern, reducing the sample concentration in half with each row. Plates were incubated with gentle shaking for 1 hour, emptied via inversion, and then washed with 100 ul PBST four times for at least 5 minutes each. Following the last wash step, 50 ul of primary or secondary probing antibodies were added in a 3% (w/v) milk in PBST solution and allowed to incubate for 1 hour with gentle shaking. Plates were then emptied by inversion and washed with 100 ul PBST six times for at least 5 minutes each. During the wash step, a SIGMAFAST OPD tablet and urea hydrogen peroxide/buffer tablet were warmed to room temperature and then dissolved in 20 ml Millipore water. 200 ul OPD/urea hydrogen peroxide solution was then added to each well taking care to minimize light exposure. The plate was incubated in the dark for 30 minutes before measuring the absorbance at 450 nm. The reaction was then quenched by adding 50 ul 3M H_2SO_4 to each well and the absorbance was measured at 492 nm.

2.14 Cell mixing

Cells previously transformed with either matched or unmatched Ab/Ag pairs were cultured in fresh LB with the appropriate antibiotics overnight at 30 C. The next morning, the cultures were normalized to an OD_{600} of 1 and then the cultures were mixed according to the predetermined ratio. For our experiment, a ratio of 1:100 matched:unmatched pairing was used. After mixing the cells at the correct dilution, the mixture was spread on a LB-Agar plate and incubated at 30 C overnight. The next morning colonies are selected and colony PCR is performed to rapidly identify the antigen in each case.

2.15 DNA mixing

Chemically competent SHuffle cells are transformed with a mixture of DNA that contains one antibody and a pool of available antigens. For initial rounds of DNA mixing, the antigen pool was mixed at a ratio of 1:1 (w) for matched and unmatched antigens. Once transformed, the cells are allowed to recover in SOC for 60 minutes before cultured overnight at 30 C in fresh LB with the appropriate antibiotics and 0.2% glucose before being plated on selective plates at varying dilutions. Following an overnight incubation at 30 C, single colonies are selected for colony PCR to identify the antigen present.

2.16 NDT library construction

The CDR H3 region for the GCN4 IgG is composed of a five amino acid chain, Gly-Leu-Phe-Asp-Tyr, with the Asp-Tyr motif highly conserved across all antibodies. In order to generate the library, degenerate primers were designed that replaced the codons for the Gly-Leu-Phe segment with an NDT-NDT-NDT segment, where N is A/C/G/T and D is A/G/T [47]. Overlap extension PCR was used to create an insert

with NcoI/AscI cut sites for insertion into the Gcn4 IgG parent vector. The Gcn4-IgG vector and the degenerate insert were then digested with NcoI and AscI as described previously, and ligated with T4 ligase overnight at 16 C. The ligation product was transformed via electroporation into specially prepared highly electrocompetent DH5alpha cells as previously described and after recovering for 90 minutes, 10 and 100 ul of the library culture were removed and plated on LB-agar the appropriate antibiotics. The remaining cell culture was added to 250 ml LB with the appropriate antibiotics and 0.2% (v/v) glucose and both liquid culture and plates were incubated overnight at 37 C. The following morning, the LB was separated into two portions; 100 ml was used for glycerol stock of the library while the remaining culture was processed to extract the library DNA. The number of colonies on the plates was counted to measure library size and individual colonies were sequenced to determine the diversity.

2.17 NNK library construction

For the NNK library degenerate primers were designed utilizing a triple NNK (N = A/C/G/T, K = G/T) codon to randomize the amino acids in the CDR H3. A single set of primers was used to amplify the entire plasmid via reverse PCR, creating a linear plasmid with blunt ends [49]. Blunt end ligation was performed for 2 hours at room temperature, followed by electroporation into highly competent DH5alpha cells as described previously. After recovering at 37 C for 90 minutes, the cell cultures were treated in method described for the NDT library.

2.18 Library selection

For selection, the library DNA was transformed as described previously into chemically competent SHuffle T7 Express cells that have previously been transformed

with the antigen containing plasmid. Following a 90 min recovery in 37 C, 10 ul of the transformation is removed and plated on LB-agar with the appropriate antibiotics to determine the library size. The remaining transformation culture is transferred to 250 ml of fresh LB with the appropriate antibiotics and 0.2 % (v/v) glucose. The plate is incubated overnight at 37 C while the liquid culture is incubated overnight at 30 C. The next day, LB-agar plates are made with varying levels of chloramphenicol (0 – 30 ug ml^{-1}). A spectrophotometer is used to measure the OD_{600} of the overnight culture and a portion is normalized to OD_{600} = 1. This normalized sample is serially diluted 10 fold and each dilution is spread on the plates with varying levels of Cm. The library plates are then incubated for 30 hours at 30 C before colonies are identified for analysis.

2.19 Library candidate analysis

Colonies identified on higher chloramphenicol concentrations are cultured overnight at 37 C in LB with the appropriate antibiotics. The next morning the potential candidates are spot plated to rapidly eliminate any false positives. The cultures are also plasmid prepped and then transformed into DH5alpha cells in order to begin the curing process. Individual colonies from this transformation are scratched on ampicillin plates followed by a kanamycin plates and grown for 10 hours at 37 C. Scratches that grow on kanamycin plates only are considered cured and are used to inoculate fresh LB with kanamycin. These cured cultures are plasmid prepped to provide the plasmid for only the antibody, which can then be back transformed into SHuffle for antibody analysis and into SHuffle with the appropriate antigen for spot plating. Candidates that demonstrate growth greater than the negative controls on this round of spot plating are sequenced, expressed, and extracted from SHuffle cells for Western blotting and ELISA experiments.

CHAPTER 3

RESULTS

3.1 CAT export is a selectable characteristic

The first portion of this work focused on establishing the groundwork for this selection by demonstrating that CAT export into the periplasm was a selectable trait. Plasmids encoding CAT in various fusion states were generated and transformed into SHuffle T7 Express cells. Single colonies were selected and cultured overnight before spot plating according to the procedure described previously.

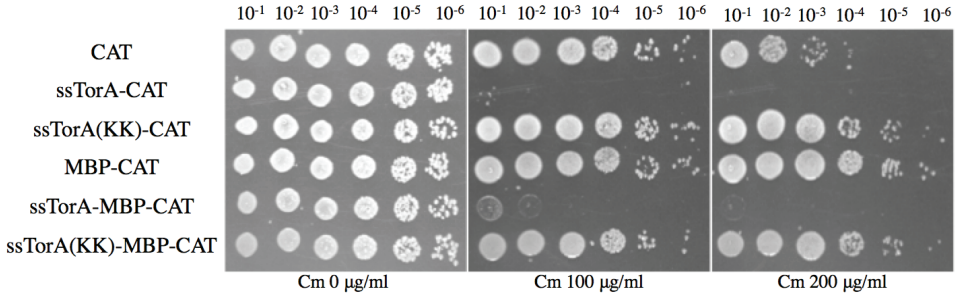

Figure 14: Spot plates demonstrate Cm sensitivity for constructs with functional export sequence

As can be seen in Figure 14, when the CAT protein is expressed without any signal sequence or protein fusions, cell growth remains even at 200 ug ml^{-1} chloramphenicol. When CAT is expressed fused to the functional Tat signal peptide, the cells ability to resist chloramphenicol decreases significantly, as demonstrated by the absence of growth at concentrations as low as 50 ug ml^{-1}. Resistance is restored by using the mutated Tat signal peptide with twin lysine residues. Expanding on this, we looked to demonstrate that even when expressed as a CAT-antigen fusion, CAT export is still selectable. To examine this, we developed MBP-CAT fusion constructs. Figure 14 demonstrates that the MBP-CAT fusion protein follows the same trends as CAT alone.

---

*The work in this section is included in the following manuscript in preparation:
Robinson, M.-P., Chang, M., Zheng, X., DeLisa, M., "A platform for selection of functional, full-length immunoglobulins within the cytoplasm of living *E. coli* cells"

3.2 Protein expression levels

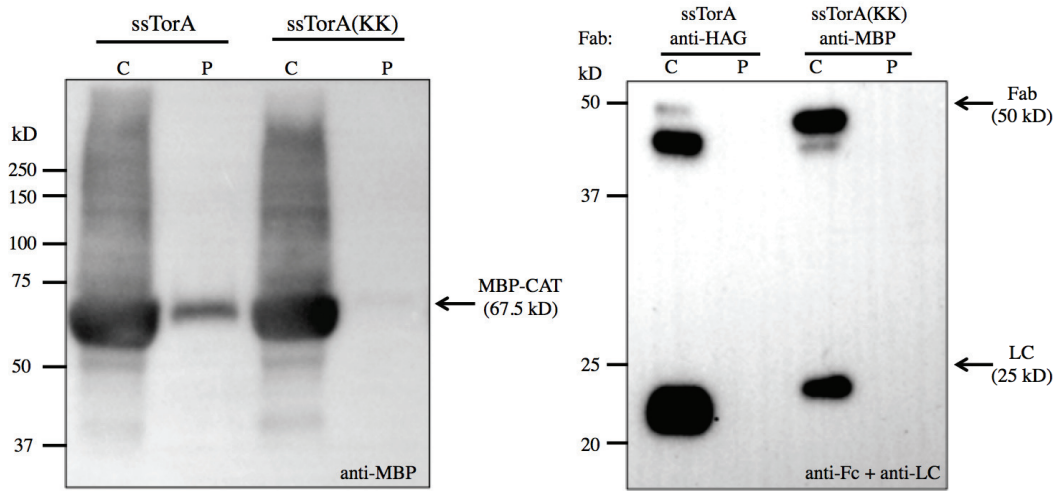

Figure 15: Subcellular fractionation to determine localization of MBP-CAT antigen and anti-MBP or anti-HAG antibody (Fab)

Periplasmic fractionalization was performed to identify the subcellular location of both the antibody and antigen proteins expressed by the two plasmids in this system. In subcellular fractionation, the outer membrane is destabilized while the inner membrane remains intact, providing a method of analyzing if proteins are located within the cytoplasm or exported to the periplasm. Cells expressing the proteins of interest were subcultured from overnight cultures and induced when the OD_{600} = 0.7-1.0. Following overnight incubation at 30 C, the samples were centrifuged and fractionated by the protocol described previously. As seen in Figure 15, when MBP-CAT is expressed with a functional ssTorA signal peptide, some of the protein is exported to the periplasm, however when the MBP-CAT construct contains the mutated ssTorA(KK) signal sequence, there is nearly no MBP found in the periplasm. Anti-GroEl probing (data not shown) of these blots demonstrated the effectiveness of this fractionation protocol and provided evidence to the purity of each fractionation sample.

Figure 15 demonstrates Fab coexpression with the MBP-CAT antigen and the resultant subcellular localization. In both of these cases, one would not expect to see Fab expression in the periplasm, since the first sample has unmatched Ab-Ag pairs and the second uses the mutated non-functional export sequence. An interesting and unexplained phenomenon in this Western blot is that when you have a matched Ab-Ag pair, such as in the second case, it can be seen that there is significantly more expression of the assembled Fab construct, whereas in the unmatched case, there is a small amount of expression of the Fab, however the majority of the light chain forms a light chain homodimer which expresses slightly below that of a Fab.

3.3 CAT-Ag orientation has dramatic effect on CAT activity

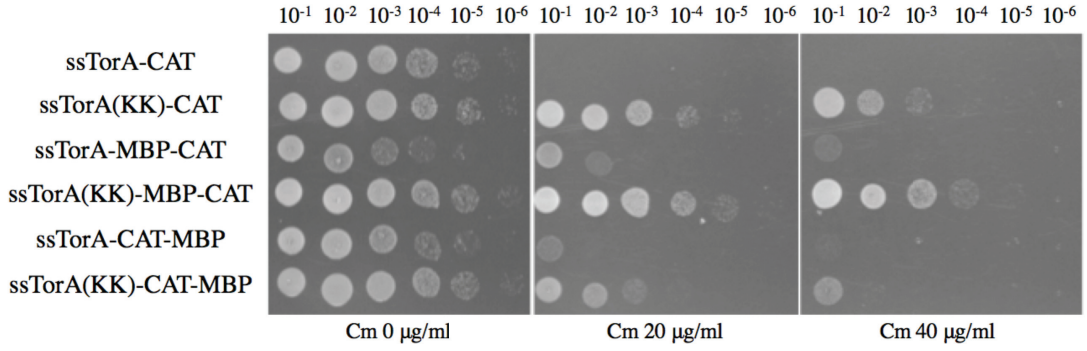

Figure 16: Order of CAT fusion plays an important role in cell survival

When designing CAT-Ag fusions, there is a choice of an N-terminal or C-terminal fusion. In order to identify the best choice, MBP-CAT and CAT-MBP fusion proteins were generated. Spot plating experiments were carried out and as demonstrated in Figure 16, the N-terminal fusion (MBP-CAT) displays significantly higher levels of resistance compared to the c-terminal fusion protein.

3.4 Full length IgG and Fab expression within the cytoplasm

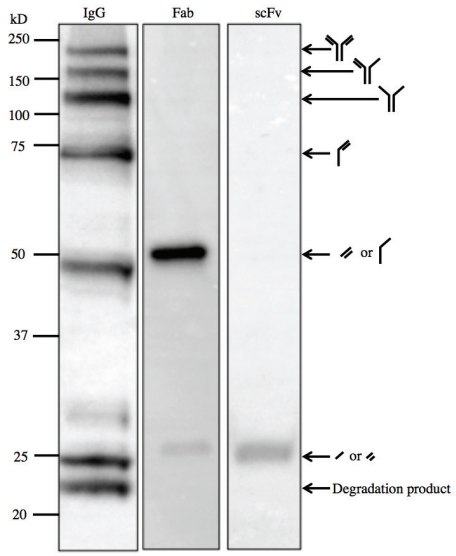

Figure 17: Efficient cytoplasmic expression of various antibody formats

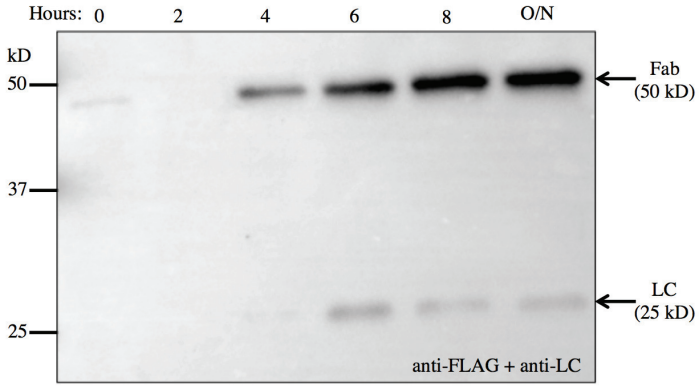

Figure 18: Time course for anti-HAG Fab expression

For successful implementation of this platform, antibodies need to be expressed and correctly folded in the cytoplasm. To confirm that this was occurring with the pCOLADuet vector, SHuffle cells transformed with the antibody construct were cultured overnight in LB followed by cellular lysis via sonication as described earlier. The samples were centrifuged to separate the soluble/insoluble fractions and then the samples were prepared as described previously for Western blotting analysis. As shown in Figure 17, sufficient quantities of all three antibody formats, IgG, Fab, and scFv, were produced at detectable levels. In addition to general expression levels, a

time trial was conducted for the anti-MBP-Fab and anti-HAG-Fab, with samples taken at 0, 2, 4, 6, 8, and 20 hours. As can be seen in Figure 18, low levels of assembled Fab can be detected by the fourth hour, and by the eighth hour, expression is nearly complete. One interesting part of this blot is that there appears to be a small amount of antibody at time 0, however this is likely due to accidental mixing during sample prep or during gel loading.

3.5 MBP fusion confers resistance with anti-MBP IgG/Fab

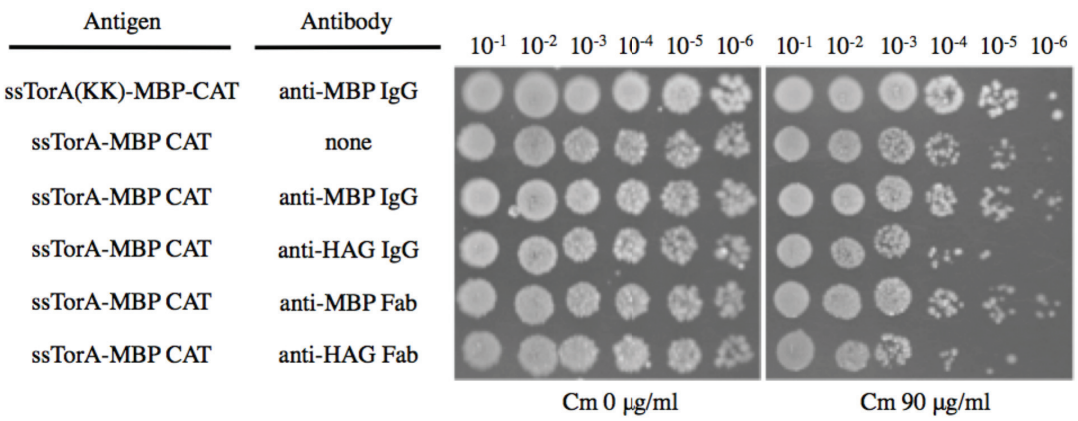

Figure 19: MBP fusions confer resistance with matched antibody in various formats

MBP is known to increase the solubility and expression of exogenous proteins within the cellular environment, therefore this was the first antibody-antigen pair tested. SHuffle cells were transformed with plasmids containing both the antibody and antigen, single colonies were then selected for overnight culturing in fresh LB + antibiotics. The next day, the samples were processed for spot plating as previously described. Figure 19 demonstrates that we can achieve a two-dilution difference between the matched and unmatched antibodies for both the IgG and Fab formats.

3.6 *gpD fusion confers resistance with anti-gPD scFv*

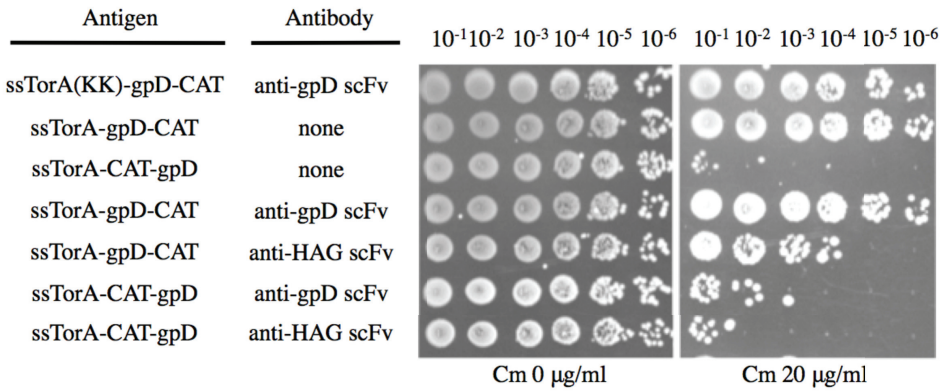

Figure 20: gpD fusions are able to confer resistance with anti-gpD scFv

With no full length IgG for gpD readily available, the anti-gpD scFv was used for this experiment. With these constructs, the gpD-CAT control does not show significant export, however the CAT-gpD construct displays the expected phenotype for the intact export sequence. Even though the control does not look as expected, the anti-gpD scFv is able to inhibit export for both antigen orientations.

3.7 *HAG fusion confers resistance with anti-HAG IgG/Fab/scFv*

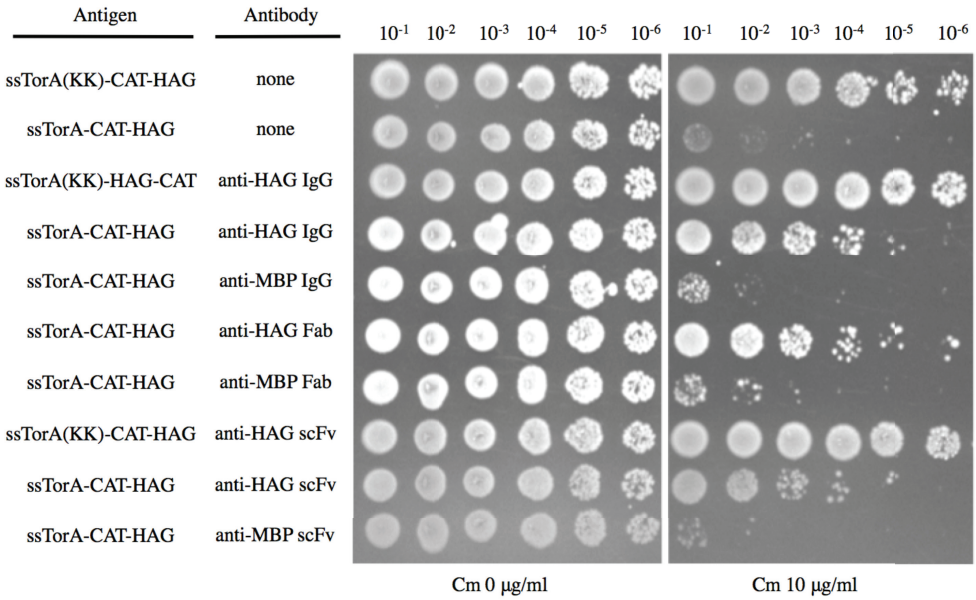

Figure 21: Multiple antibody formats can achieve the same result with the HAG antigen

Following the success with MBP and gpD-CAT fusions, new antibody-antigen pairs were explored, starting with the HAG peptide. SHuffle cells were co-transformed with CAT-HAG plus signal sequences and either the IgG, Fab, or scFv. Spot plating analysis demonstrated that the HAG constructs displayed a more robust trend compared to the MBP and gpD fusions (Figure 21).

3.8 Pre-induction of Ab/Ag overwhelms the Tat pathway

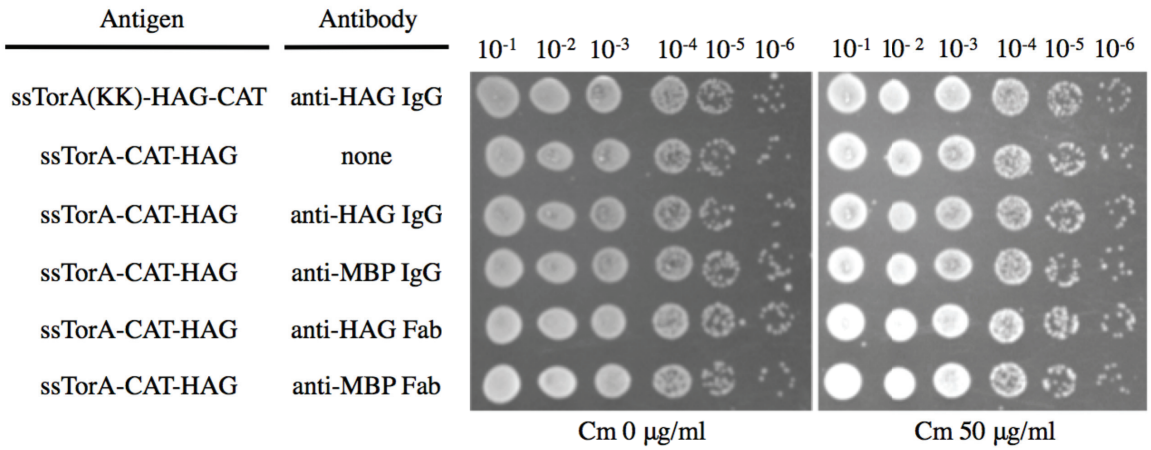

Figure 22: Pre-induction of the antigen overwhelms the Tat pathway

In an attempt to improve the difference between matched and unmatched antigen/antibody pairs, the antigen, antibody, or both proteins were induced16 hours prior to spot plating. As can be seen in Figure 22, pre-inducing the antigen prior to plating results in a saturation of the Tat pathway and overwhelms our selection. Inducing the antibody expression or the combination of both generates results similar to these.

3.9 Order of CAT-Ag impacts selection

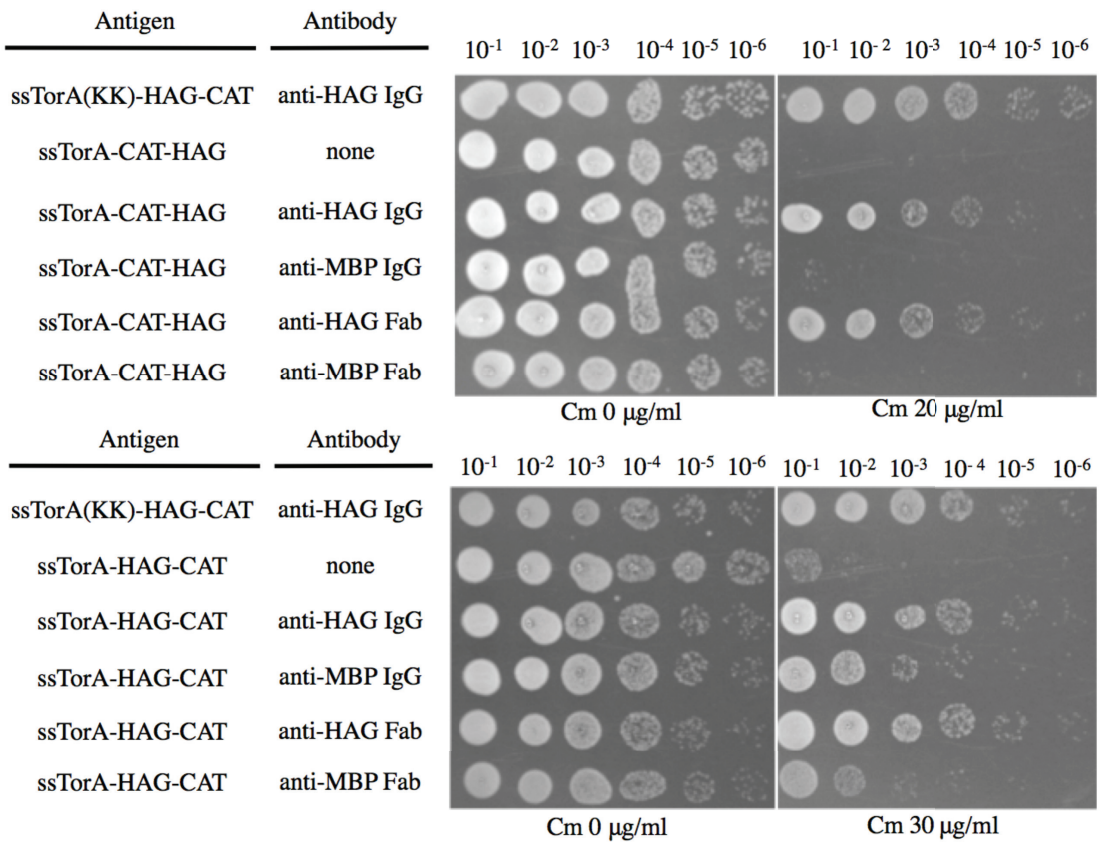

Figure 23: Optimal orientation of antigen fusion for selection opposite of the result generated with mutated signal peptides

Previously it has been demonstrated that the order of the CAT-Ag fusion plays a significant role in the level of CAT function. However the different orientations have never been tested within this antibody platform. C- and N-terminal fusions were co-transformed with either matched or unmatched antibodies and samples were spot plated as previously described. As can be seen in Figure 23, when the antigen is fused to the N-terminal CAT, the difference between the matched and unmatched antibody-antigen pair is only a couple of spots, whereas with the C-terminal fusion there is at least a four dilution difference.

3.10 *Gcn4 fusion confers resistance with anti-Gcn4 IgG/Fab*

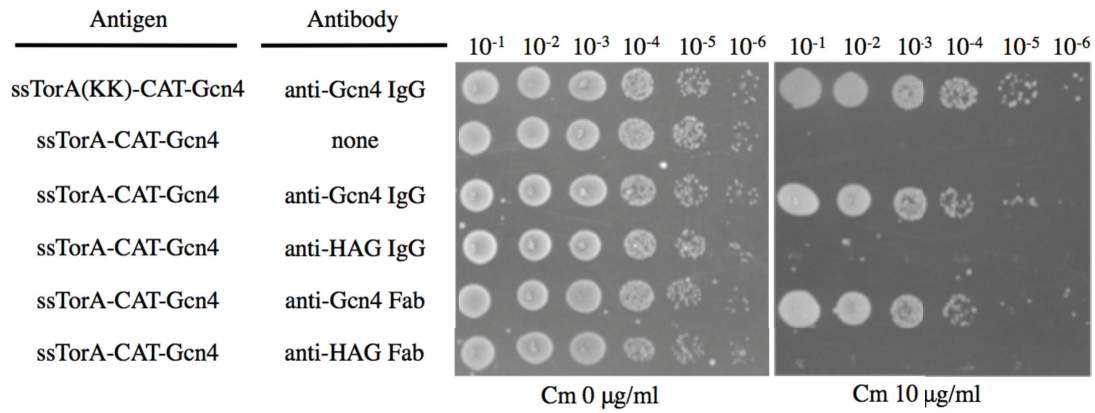

Figure 24: Gcn4 is a suitable antigen for this platform

Following the success of the smaller HAG peptide compared to MBP and gpD, the transcription factor Gcn4 was selected as the next construct. The Gcn4 peptide was cloned into the pBAD24 vector while the antibody remained on the pCOLADuet vector. Spot plating experiments in Figure 24 demonstrate that Gcn4 is a suitable antigen and the growth is very similar to that of the HAG constructs.

3.11 *anti-cMyc grafting*

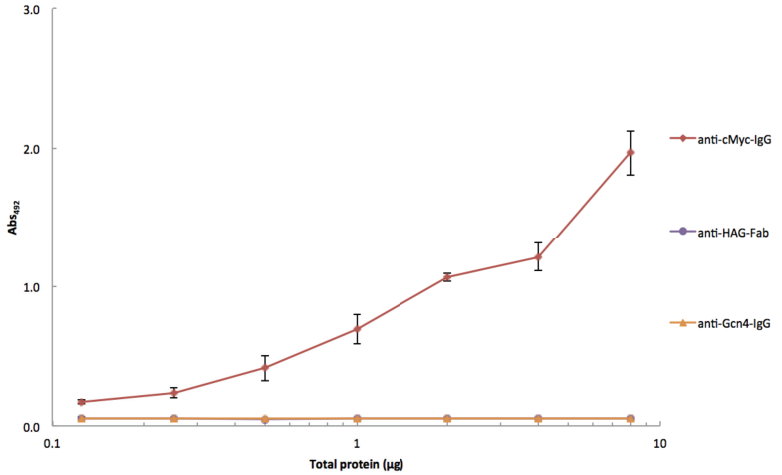

Figure 25: ELISA for grafted anti-cMyc IgG

With two successful Ab-Ag pairs in the HAG tag and Gcn4, one more case study was desired before progressing to library development. Previous work has engineered an scFv against the cMyc tag, and therefore it was decided that this would be a perfect third candidate for the system. Prior to incorporating this, the 3DX scFv needed to be grafted onto an IgG framework. Following the procedure detailed previously, the scFv for cMyc was inserted onto the framework from the HAG IgG. The first column in Figure 17 shows the western blot analysis of the grafted antibody, which displays the characteristic laddering of various fragments and the full-length construct. In addition to the western blot, an ELISA was performed with the new antibody and Figure 25 represents that work. As can be seen in the figure, the cMyc antibody displays a high binding curve, whereas the anti-HAG and anti-Gcn4 controls demonstrate no binding to the immobilized cMyc antigen.

3.12 cMyc fusion confers resistance with anti-cMyc IgG

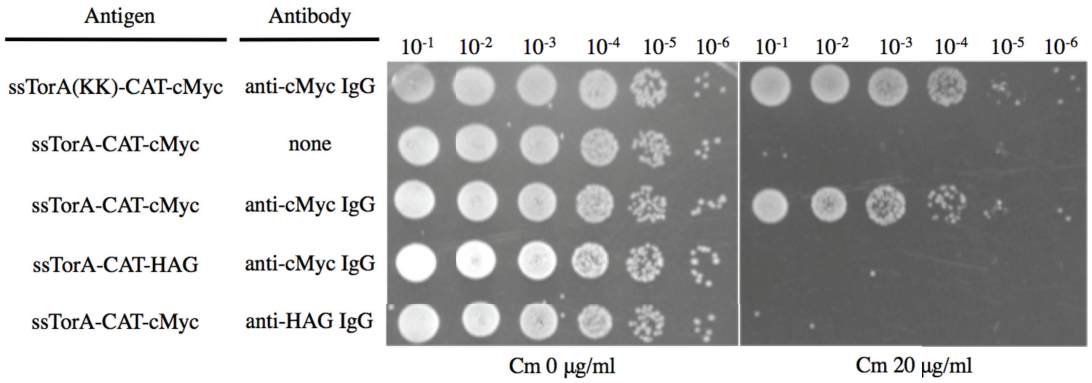

Figure 26: CAT-cMyc fusion is able to confer Cm resistance to selected cells

Following the successful grafting of the 3DX scFv into our antibody framework, the cMyc Ab/Ag pair was implemented in this system. As shown in Figure 26, the cMyc pair closely resembles that of the HA and Gcn4 pairs.

3.13 Cell mixing experiments

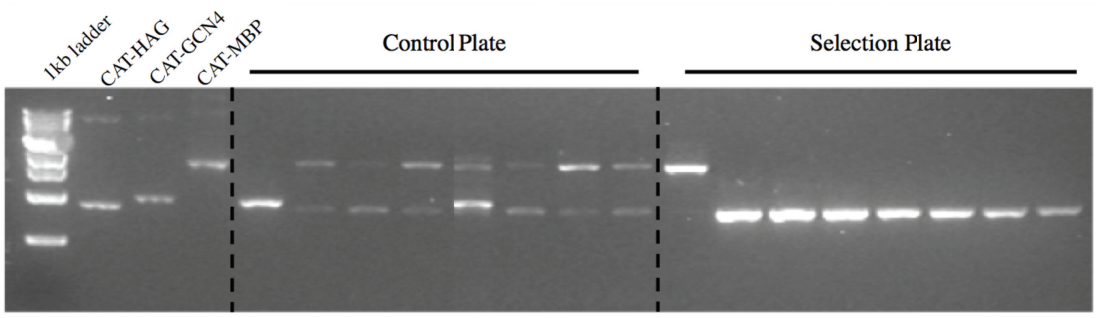

Figure 27: Cell mixing experiments are able to select for desired antigen at 1:100 cell dilution

With three successful cases identified, cell mixing experiments were used as the first attempt at applying this platform in the context of a mock selection. Cells transformed with the anti-HAG IgG and CAT-HAG, CAT-Gcn4, or CAT-MBP (with ssTorA signal sequences) were cultured separately in LB with the appropriate antibiotics overnight. The following morning the cultures were normalized to an OD_{600} = 1 and mixed at a ratio of 1:100 in favor of the unmatched pairs. In this case, there were two cultures with unmatched pairs, and they were used equally. Colony PCR analysis of the selected candidates revealed that 7/8 were the matched antigen whereas the control plate showed an erratic distribution of mixed antigens.

3.14 DNA mixing experiments

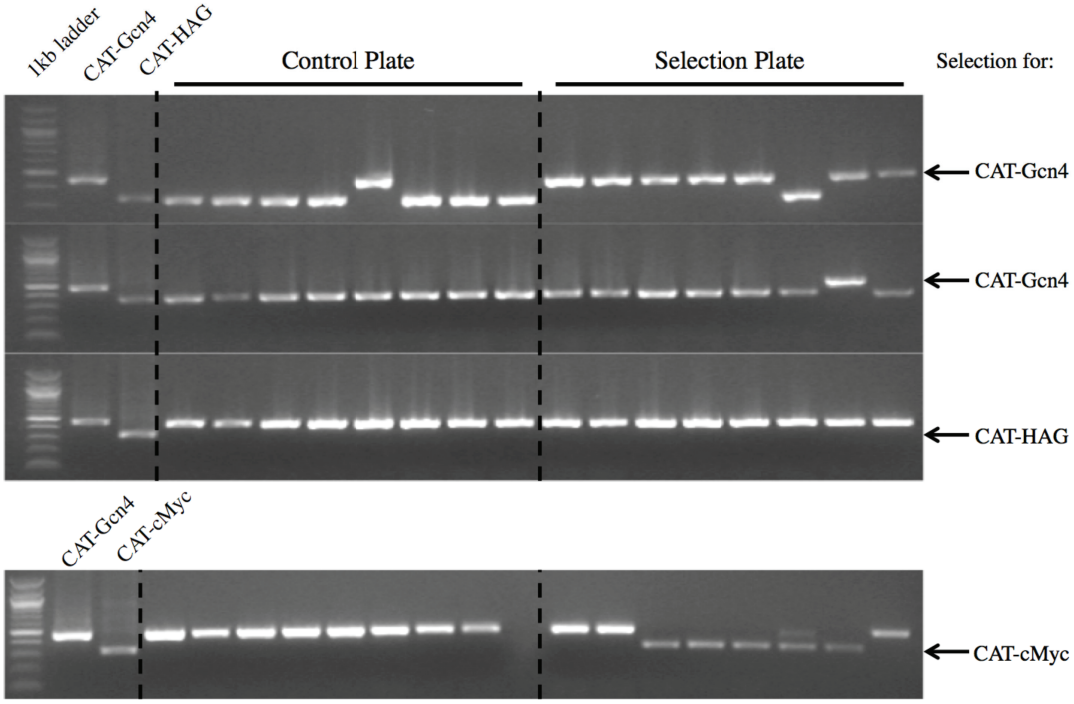

Figure 28: DNA mixing provides a mixed bag of results

Contrary to cell mixing where cells of known genotype are mixed, DNA mixing takes a DNA pool and transforms the pool into empty competent cells, providing a closer analogy to the actual library selection process. As shown in Figure 28 DNA mixing provided inconsistent results. At times the desired shift in antigen propensity appeared from the control plate to the selective plate, however there were also a significant number of instances when the control and selective plates had identical trends, either a preference for one antigen or a random assortment of antigens.

3.15 Fos is a suitable antigen for library selection

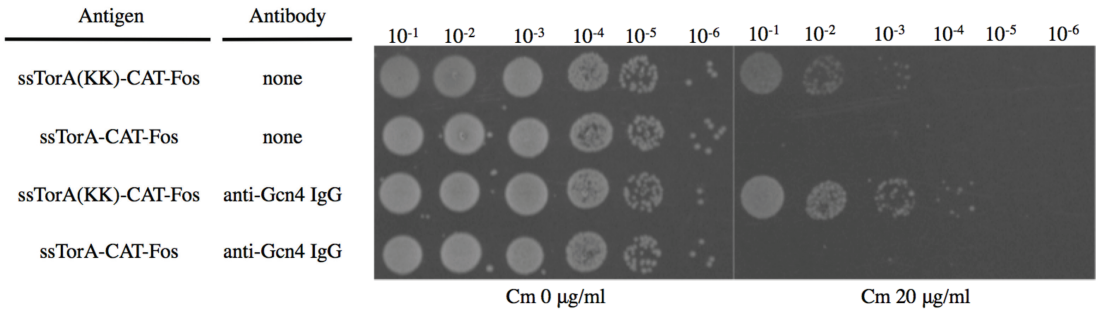

Figure 29: Fos is a suitable antigen for library selection

Prior to developing a library, a suitable antigen needs to be selected. For this work a simple selection was desired for the first attempt, therefore Fos was chosen as the initial antigen. Since Fos is another small alpha helical protein, it was decided to attempt to switch the specificity of the anti-Gcn4 IgG to anti-Fos. In order to confirm this choice, Fos constructs were spot plated with both functional and mutated signal sequences and with the Gcn4 IgG to demonstrate there was no non-specific interactions between the two.

3.16 Library diversity (NDT and NNK)

Table 3: Library diversity confirmed by sequencing analysis

NDT (RE digest)		**NDT (reverse PCR)**		**NNK (reverse PCR)**	
Sequence	CDR-H3	Sequence	CDR-H3	Sequence	CDR-H3
AAT-GTT-AGT	NVS	GCG-CAT-CGG	AHR	GCG-CAT-CGG	AHR
GGT-GAT-CAT	GDH	TGT-CTT-AGT	CLS	ACT-GAG-GCT	TEA
ATT-GGT-CGT	IGR	CGT-GTT-GAT	RVD	TGG-AGT-ATT	WSI
ATT-AAT-CTT	INL	TTT-TGT-CTT	FCL	GCG-CAG-CGG	AQR
		CGT-TTT-CAT	RFH	CAG-AGG-ACG	QRT

To confirm the diversity of the library, the naïve library was plated on LB-agar plates with 100 ug ml^{-1} kanamycin and incubated overnight at 37 C. Single colonies were selected and cultured in fresh LB with kanamycin, and grown overnight at 37 C. The plasmids were then extracted and sequenced via the Biotech Resource Center at Cornell University. The diversity was evaluated for all of the libraries generated and Table 3 shows the sequencing results from these libraries.

3.17 Fos NDT OE-PCR library 1 candidates

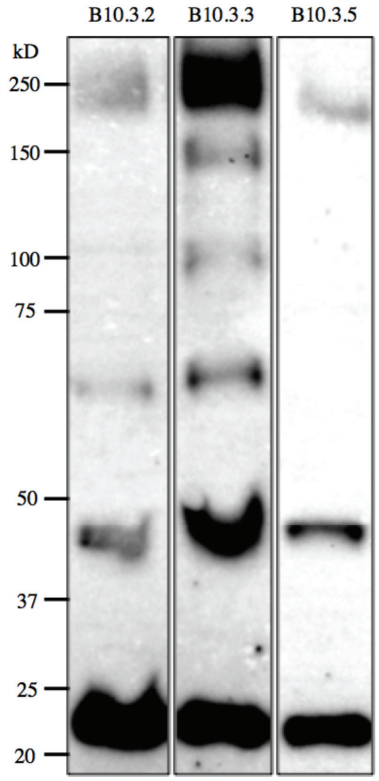

Figure 30: Western blot of Fos library candidates display antibody laddering

The first NDT library was generated via OE-PCR and standard restriction enzyme cloning. As shown by sequencing and control ligations, this library had a large number of non-functional members and also a high proportion of the parent anti-Gcn4-IgG. The library was transformed into chemically competent SHuffle cells already

containing the Fos antigen plasmid. Following an overnight recovery in non-selective media, a sample was normalized to OD_{600} = 1 and plated at dilutions ranging from 10^{-1} to 10^{-4} on plates with either 10, 20, or 30 ug ml^{-1} of chloramphenicol. 20 candidates were chosen from a range of chloramphenicol concentrations and dilution rates for further analysis. Following curing, back transformation, and sequencing, three candidates remained for characterization. On the Western blot (Figure 30), all three of these antibodies demonstrated portions of the expected laddering for IgG expression so the next step was to perform an ELISA. On the first round of ELISA, one of the hits appeared to have marginally increased binding ability to Fos compared to the wild type anti-Gcn4 IgG, however following the creation of a biological replicate, that improvement disappeared.

3.18 Gcn4 NDT OE-PCR library 1 candidates

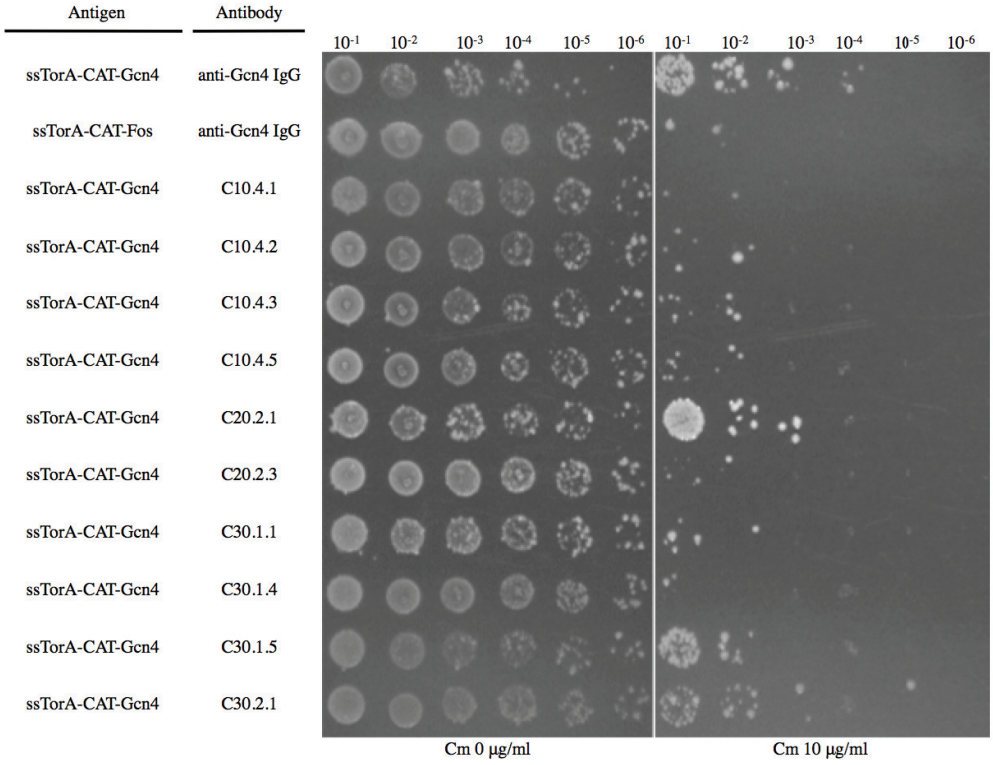

Figure 31: Back-transformation of Gcn4 NDT selection reveals three candidates

Following transformation and incubation, 20 library candidates were selected for further screening from a range of chloramphenicol concentrations and dilution rates. After the culturing the hits overnight and the first round of spot plating, four of the candidates were eliminated leaving 16 for curing and back transformations. All 16 were successfully cured after one round of patching and subsequent plasmid prep generated pCOLADuet plasmid solutions for back-transformation into chemically competent SHuffle cells. For the first set of spot plates, 10 of the 16 remaining candidates were selected. As shown in Figure 31, three of the candidates showed a pronounced difference compared to the negative control. Sequencing of these three candidates revealed that they were not actually members of the NDT library, rather they were from the original Gcn4-IgG construct that served as the backbone for the degenerate library. With this finding, it was decided to stop pursuing this library and focus on the NDT and NNK libraries generated via reverse PCR.

3.19 Gcn4 NDT and NNK reverse PCR library

In order to combat the high number of parent antibodies and nonfunctional members, new NDT and NNK CDR-H3 libraries were generated using reverse PCR. Eight members of each library were used for colony PCR to identify if the desired antibody fragment was present and six other members were sequenced to identify the library diversity. Colony PCR revealed that 7/8 of the NNK library and 8/8 of the NDT library have the desired fragments. Sequencing of six different colonies from each naïve library revealed no repeated sequences, demonstrating the diversity of each library.

CHAPTER 4

Discussion

4.1 CAT export

The first step in establishing the groundwork for this project required demonstrating that CAT export was a selectable trait. To achieve translocation of the CAT protein, an N-terminal fusion with the Tat signal peptide (ssTorA) was created. As shown in Figure 14, with the native signal sequence, the CAT protein is rapidly exported to the periplasm, making those cells vulnerable to the Cm and therefore decreasing the growth by orders of magnitude compared to the protein with the mutated signal sequence (ssTorA(KK)). While the transport of CAT was an important first step, we also needed to demonstrate that we could repeat these results with a CAT fusion, which would serve to provide the antigen in future experiments.

The first antigen chosen was the maltose binding protein. Four constructs were generated covering both N- and C- terminal fusions as well as the use of both the functional and nonfunctional signal sequences. As shown in Figure 16, the cells are able to export both orientations of the CAT-antigen constructs, however the C-terminal fusion displays a slightly more robust survival compared to the N-terminal construct. This trend can also be attributed to the presence of the MBP protein, a commonly used solubility enhancer when fused to the N-terminal of various proteins. In this case the cell may simply produce such a high level of the MBP-CAT fusion that even at maximum transport levels the cell simply cannot export all of the CAT to the periplasm. With these results, we are able to successfully demonstrate that we can

select for cells that are unable to rapidly export the CAT fusion protein, laying the groundwork for the development of our antibody selection platform.

4.2 Antibody expression

Having successfully demonstrated that CAT export is a selectable characteristic, the next aim was to determine if antibody-antigen binding could be used to block export. However before moving into analyzing antibody-antigen binding interactions, a plasmid needed to be constructed that could be coexpressed with the pBAD24 plasmid encoding the antigen. Plasmid incompatibility is generally found to exist when two plasmids share a common origin of replication, resulting in a competition for replication resources and machinery. Because of this competition, any plasmid that offers growth advantages to the cell, either by being smaller and thus easier to copy or having a lower cellular toxicity will be over produced and outnumber the other plasmid [49]. Using the table in Tolia and Joshua-Tor, a plasmid was selected that is compatible with the ColE1 origin used by the pBAD24 vector [50]. The pCOLADuet vector has a low copy number but contains the COLA origin, which is compatible with the ColE1 origin used by the pBAD vectors and was therefore chosen. In addition to having a compatible origin, the standard resistance marker for the pCOLADuet plasmid is kanamycin, which is better suited for this system than pACYCDuet, another compatible plasmid which uses chloramphenicol as the selective marker and therefore would need to be modified before using in a CAT based system.

pCOLADuet is a bicistronic plasmid, designed with two separate MCS, each with its own T7 promoter, *lac* operator, and ribosome binding site. For this study, the light chain of the antibody was cloned between the NcoI and NotI restriction sites in MCSI and the heavy chain was cloned between the NdeI and XhoI restriction sites in MCS2.

For the Fab constructs, the Fc region was removed from the heavy chain and replaced with a FLAG Tag (DYKDDDDK) flanked by an AscI restriction site. By having the light and heavy chains produced separately and with their own promoter regions, this allows for the potential to fine tune expression of either chain by engineering the RBS or promoter regions if efficient assembly is not achieved. For the scFv constructs, the promoter region for MCS2 was removed and the scFv was inserted between the NcoI and XhoI sites.

Before testing these antibodies in the selection platform, we needed to confirm that the pCOLADuet plasmid was capable of producing fully functional antibodies. All three antibody formats (IgG, Fab, scFv) were expressed, harvested, and ran on an Invitrogen NuPAGE 4-12% Protein gel in MOPS buffer. As demonstrated in Figure 17, all three formats are expressed in their fully assembled form at detectable levels. Western blotting the reduced samples demonstrates that the assembled products seen on the blots are from disulfide bonding, and not other arbitrary interactions resulting in aggregation of random fragments. An ELISA was also performed for the first IgG produced with this plasmid to compare binding ability to those antibodies produced by previously used plasmids. Anti-MBP IgGs were tested against those produced in pMAZ360 and pET21b plasmids via an ELISA and were found to have a similar degree of specificity and affinity (data not shown). Also important, the anti-HAG IgG produced in the pCOLADuet system did not bind to the MBP antigen, demonstrating that the antibodies produced via this plasmid are not only functional, but they maintain their specificity. While the pCOLADuet plasmid is expected to be compatible with the pBAD24 plasmid, it needed to be confirmed that SHuffle cells are able to express both proteins. Confirmation was achieved via periplasmic fractionation as it can be seen in

Figure 15 that SHuffle cells are able to produce both the MBP-CAT fusion as well as the desired Fab.

4.3 Testing of antibody-antigen pairs

4.3.1 MBP

Having successfully demonstrated that the pCOLADuet plasmid can produce functional antibodies in conjunction with pBAD24 expression of the antigen, it was time to look for antibody-antigen pairs that could be used to halt export of the CAT protein. Since the original proof of concept experiments used CAT-MBP fusion proteins, it was decided to start with the anti-MBP antibody in IgG and Fab formats. Cultures co-transformed with either the matched anti-MBP + MBP-CAT combination, or an unmatched control combination of anti-HAG + MBP-CAT, were spot plated on varying concentrations of chloramphenicol. As can be seen in Figure 19 when the anti-MBP IgG and MBP-CAT constructs are combined, the culture grows to a dilution of 10^{-6}, compared to the unmatched anti-HAG IgG, which only grows to 10^{-4}. The same trend can be seen for the anti-MBP Fab, which restores growth up to the 10^{-6} while the anti-HAG Fab only has growth to 10^{-4}. These results are very encouraging, however the difference between the matched and unmatched Ab-Ag pairing is not significant enough for it to be applied to a library. The primary issue can be seen with the control cultures in the first two lanes of Figure 19, ssTorA(KK)-MBP-CAT + anti-MBP IgG and ssTorA-MBP-CAT. While the positive control demonstrates the proper growth phenotype expected from the mutated ssTorA(KK), the negative control of only ssTorA-MBP-CAT continues to grow strongly to the 10^{-5} dilution. With such a significant level of growth for the strain with no blocked export, it is not surprising that this platform is unable to achieve a great enough difference between matched and unmatched samples. One theory regarding the cause of this is similar to the one

discussed earlier in the proof of concept stage. When MBP is used as an N- terminal fusion, the solubility of the fusion protein is greatly increased, which in turn increases expression levels. Therefore, by using the N-terminal fusion, the cells are simply over producing the MBP-CAT fusion so that the Tat pathway is overwhelmed and unable to export the CAT protein quick enough to keep up with production. Due to this overproduction, the cells exhibit the growth phenotype even at concentrations of 100 ug ml^{-1} chloramphenicol, which is nearly 3 times the standard concentration of chloramphenicol used. Since there is so much MBP-CAT present in the cell, the blocking that does occur is only able to increase the growth phenotype slightly, but not enough to enable a library selection.

4.3.2 gPD

The next protein used was gpD, the capsid stabilizing protein of bacteriophage lambda. After bacteriophages pack their DNA into the capsid chamber, the gpE proteins undergo a dramatic reconfiguration, increasing the shell diameter from 50-60 nm, while also decreasing the capsid thickness. In order to counter the internal pressure generated from the tightly packed DNA, gpD is used as a cementing protein, which stabilizes the mature phage capsid shell [51]. For this antigen, the only readily available antibody format was the scFv, therefore testing the ability of the IgG and Fab to block transport was not practical. As shown in Figure 20, the anti-gpD scFv is able to block transport of both antigen configurations. While the ssTorA-CAT-gpD shows strong export of the single plasmid control, the opposite configuration shows limited export since there is growth to the sixth dilution. Even though this control did not display the desired phenotype, the unmatched Ab-Ag pairings for both orientations resulted in CAT export and chloramphenicol sensitivity. When the anti-gpD scFv was used, the sensitivity was decreased, allowing the cultures to grow 1-2 dilutions further

than for the unmatched conditions. While both orientations displayed the correct trends with similar relative differences, comparing the two orientations reveals that there is a significant difference between the growth fitness of the cells. CAT-gpD constructs showed the weakest growth, with growth only extending to the third dilution compared to the gpD-CAT constructs, which had growth to the sixth dilution with the matched pair. Even with the difference in growth, the relative difference between the matched and unmatched pair remained constant. While gpD does show a difference between matched and unmatched pairing, the issue from the MBP trials still remain, since the growth difference between the matched and unmatched pairs is not significant enough to move forward with the library.

4.3.3 HAG

In following the trend of decreasing sizes, the next target chosen was the HAG tag, a peptide sequence that is adapted from residues 98-106 of the surface glycoprotein hemagglutinin from human influenza. While MBP is 40 kDa and gpD is 11.5 kDa, the HAG tag is only 0.2 kDa (9 AA), making it one of the smaller peptide moieties available for use in this platform. Once the HAG peptide was cloned into the antigen plasmid, testing began with the anti-HAG IgG and Fab. Initial experiments utilized the HAG-CAT configuration since initial testing with the MBP constructs showed that the N- terminal fusion resulted in more robust growth compared to the C- terminal fusion. However experiments using this orientation had results similar to that of MBP and gpD, with the anti-HAG IgG and Fab only providing for the recovery of two dilution factors while maintaining a large amount of background growth (Figure 23). The CAT-HAG construct on the other hand displayed remarkable recovery of 5-6 dilutions with the matched antibody-antigen pairs while having only a small amount of growth on the first dilution for unmatched pairs (Figure 23). With this knowledge, future

constructs were designed using the CAT-Ag format in an attempt to recapitulate the success had with CAT-HAG.

In addition to using the standard IgG and Fab formats, the success of the anti-gpD scFv encouraged further examination into the anti-HAG scFv. All three conformations were spot plated together in order to allow for basic comparison between the successes of each one. As seen in Figure 21, the IgG, Fab, and scFv all display a similar amount of growth for the matched or unmatched pairs, demonstrating that it is possible to use all of the antibody formats with this platform. Even though the scFv and Fab formats are simpler and easier to assemble within the cell, the IgG is the more clinically relevant molecule. Also, the Fab and scFv can be easily generated from the IgG whereas converting the fragments into a full length IgG is a challenging task that requires a significant amount of optimization. Therefore, we decided to focus on the IgG format of future antibodies, especially in regards to our antibody library creation.

4.3.4 Gcn4

With the success of the CAT-HAG constructs, the next protein chosen was Gcn4. Gcn4 is a basic leucine zipper and is primarily responsible for transcriptional response to amino acid starvation. While wild-type Gcn4 rapidly forms a homodimer, this work utilized a mutated version in which the e and g positions in the alpha helix are converted into proline residues, disrupting the salt bridges typically formed during dimerization, destabilizing the complex and allowing the Gcn4 protein to exist as a monomer [52]. As shown in Figure 24, the CAT-Gcn4 fusion is able to increase chloramphenicol resistance to the fourth dilution, whereas the unmatched pair shows no growth on plates with 20 ug ml^{-1} of chloramphenicol.

4.3.5 cMyc

With two suitable antibody-antigen pairs identified, one more example was sought after before progressing into the next stage of this project. In keeping with the trend of smaller proteins working better, the cMyc protein, a common transcription factor that is known to be associated with numerous cancers, was chosen. The cMyc antigen was available in the lab for cloning however the full length anti-cMyc IgG or Fab did not exist. To fix this, 3DX, an evolved form of the anti-cMyc scFv 9E10 was grafted onto the antibody scaffold used by the other pathways [53]. In order to graft the scFv onto the framework, it is critical to identify where the CDR-H3 and L3 ended on the parent antibody so that the entire binding region could be replaced by the scFv. This was done by identifying the Trp-Gly-XXX-Gly and Phe-Gly-XXX-Gly motifs, which typically signify the end of the CDR-H3 and L3 respectively. Once these were identified, OE-PCR was used to create an insert that contained both the variable region and the first constant region in each chain (CL, CH1) with the appropriate restriction sites for cloning into our pCOLADuet vector. After creating the new antibody construct, it was important to characterize the anti-cMyc IgG via Western blotting and ELISA. In Figure 17, column 1, you can see characteristic laddering of IgG expression in the non-reduced portion and in the reduced portion all of the bands collapse into the light and heavy chains (data not shown). After characterizing the assembly of our new antibody, the next step is to test the binding ability via an ELISA. Figure 25 demonstrates that this grafted anti-cMyc IgG is able to bind the cMyc antigen at a significantly higher level than the anti-Gcn4 IgG and the anti-HAG IgG.

With the IgG constructed, tests were ran to analyze the ability of anti-cMyc IgGs to block cellular translocation. As shown in Figure 26, the ssTorA(KK) mutant and the wild type ssTorA peptide both perform as expected, with the KK mutant allowing for

growth to the sixth dilution and the ssTorA signal sequence showing little to no growth on the first dilution. When the anti-cMyc IgG was added to the c-Myc antigen, growth was restored to the fifth dilution while the c-Myc antigen paired with anti-HAG IgG is unable to restore growth. With this data, there are now three antibody-antigen pairs demonstrated to work with this platform.

4.4 Pre-library experiments

4.4.1 Cell mixing

Cell mixing experiments were the first form of a mock selection that was performed. For this experiment, cells with known antibody – antigen combinations were cultured overnight so that the OD_{600} could be normalized and the samples could be mixed at varying dilution rates. The dilution rate that we chose was 100:1 in favor of the unmatched pair. The samples were mixed together and plated immediately on selective media. As can be seen in Figure 27, the nonselective control plate has an assortment of cell genotypes present, however once you move over to the selective plate you eliminate the colonies with mixed plasmids and are able to select 7/8 as the matched antigen-antibody pairs.

4.4.2 DNA mixing

The DNA mixing experiments were performed as a way to mimic standard library set-ups, where instead of having cells with known components, the DNA pool is transformed into a batch of cells for culturing and plating. Results with this experiment were very mixed. With the lower ratio of matched to unmatched antigens a decent amount of selection was reported. However once the ratio was increased to 1:50 or 1:100, the results tended to favor one antigen, regardless of the antibody. This is most likely explained by the fact that there is a transformation bias against the larger

plasmids. Therefore, when selecting against the Gcn4 antibody, the majority of the controls are randomized, but the platform is able to select only the Gcn4 antigens on the selective plates (Figure 28). However when you use the HAG antibody, the smaller construct is able to enter the cells easier, therefore the controls contain a significantly higher proportion of HAG antigen already (Figure 28). While this is definitely an issue for this experiment, this difference in plasmid size is not one that would be encountered during the library experiments for this platform. For the actual selection, the antibody library plasmids will be transformed into cells that already have the desired antigen, therefore all of the plasmids will be the same size. Here an antigen library is utilized instead of the antibody library since that allows for the simplest characterization. With the antigens differing in size, we are able to use colony PCR to rapidly identify (via plasmid size difference) which antigen has been selected for by that antibody, whereas to determine which antibody had been selected out of a pool would require either multiple colony PCR steps with different primers or a full sequencing analysis. For these reasons, the decision was made to progress with our actual library even though the DNA mixing experiments had not been perfected.

4.4.3 Fos as a suitable antigen

In order to develop a library against a given antigen, a suitable antigen needs to be identified. As a first round test, it was chosen to take the anti-Gcn4 IgG and switch its specificity to Fos. Fos was chosen since it is another basic leucine zipper, and therefore it has structure similar to Gcn4. To ensure that there are no native interactions between Fos and the anti-Gcn4 IgG, spot plating experiments were carried out. As can be seen in Figure 29, the ssTorA and ssTorA(KK) controls perform as expected with the positive control displaying growth to the fourth dilution and the negative showing a few colonies on the first dilution. The paired CAT-Fos + anti-

Gcn4 IgG shows no growth on the plates which means that Fos is a suitable antigen for our Gcn4-IgG specificity switching library.

4.5 Library generation

The initial library was designed with two goals, the first was that it would be used to take the anti-Gcn4 IgG and convert it into anti-Fos IgG. The second goal would be to select this library against the Gcn4 antigen to see if the original binding domain could be recovered from the antibody pool. In addition to recovering the WT antibody, it would be interesting if scFv hits originating in the Waraho and der Maur studies were also selected from this IgG platform [54][55]. While Fab and scFv fragments are simpler to express and assemble within *E.coli* cells, the benefits of focusing on selecting full length IgGs outweigh the difficulties. IgGs are more clinically relevant than the smaller fragments due to their stability, extended circulation time, and ability to access effector functions. In addition to these benefits, the IgG format can be rapidly converted into Fab and scFv formats via enzymatic digestion or cloning strategies, however grafting the shorter binding fragments onto the IgG framework is significantly more difficult, required many optimization steps.

The libraries used in this study were generated by randomization of the CDR-H3 regions of the anti-Gcn4 IgG. This region was chosen since this CDR-H3 is only five amino acids long (GLFDY) and the last two are highly conserved across the antibody space. CDR-H3 is also the most prominent contact point between the Gcn4 IgG and the Gcn4 antigen, therefore it has the greatest influence on antibody binding. Since Fos has similar structural characteristics to that of Gcn4, it was imagined that by making slight variations to the parent anti-Gcn4 IgG it would be able to switch the specificity of the IgG towards Fos.

The initial library was designed using a triple NDT (N=A/C/G/T; D=A/G/T) degeneracy to replace the three variable residues in the CDR-H3 of the anti-Gcn4 IgG. While the NDT degeneracy does not encode for all 20 amino acids, the 12 amino acids it does allow is intended to cover the entire spectrum of functional groups present on amino acids. In addition to covering all functional groups/types of amino acids, this degeneracy also does not include any stop codons and has a 1 codon, 1 amino acid setup, which allows creation of a library without any amino acid bias. From a practical standpoint, Reetz et al. demonstrated that for a three amino acid NDT library, only 5,200 library members are required to attain 95% coverage, compared to the 98,000 required for a full NNK library [47]. While this method is well suited for using the library to select the parent GLF antibody out of the library, the limited diversity could provide an issue for the Fos library. Even though Gcn4 and Fos share many structural characteristics, antibody binding is extremely finicky so just covering all the types of amino acid functional groups may not be enough to switch the specificity.

The first NDT library was designed using OE-PCR and standard restriction enzyme cloning techniques. The first attempts at creating this library used the NdeI and AscI sites flanking the Fab portion of the heavy chain, however agarose gel analysis revealed that the vector digestions were incomplete and undigested plasmid constantly remained in the vector preparations. When extracting the vector from the agarose gel, it was attempted to cut only the digested portion, however following ligation and test transformations it was revealed that over 50% of the first library was composed of parent antibody or nonfunctional ligation products. This was confirmed following colony PCR and subsequent sequencing of the naïve library. From this library, only 20% of the colonies that were tested contained an insert. However, the antibodies that

could be sequenced were not the parent antibody, therefore there was some degree of diversity within this library.

In the hopes of generating a library that contained a higher proportion of functional members, the OE-PCR was expanded to encompass the entire Fab portion and therefore use the NcoI and AscI restriction sites. While agarose gel analysis showed that this digestion was still incomplete, the control ligation showed only 10% of the library was composed of parent or non-functional self ligated members. With this in mind, we decided to proceed with this as our first library.

The second and third libraries were designed with CDR-H3 NDT and NNK degeneracy respectively. The difference between these libraries and the first library was the method in which the PCR was carried out. For this library, reverse PCR was used to amplify the entire plasmid, eliminating the need for vector digestion and therefore eliminating the presence of both non-functional self-ligated and non-digested parent members. Two primers were designed such that the reverse primer carried the desired degeneracy while the forward served to create a blunt-ended linear version of the entire plasmid, which could be then ligated together with a blunt-end ligase. Colony PCR and sequencing of the naïve libraries generated with this method revealed that 12/12 NDT library members contained the insert and of the six that were sequenced none were repeats nor the parent antibody. The NNK library had similar results, with 10/12 showing the correct insert and the five that successfully sequenced were diverse with no repeats.

4.6 Library selection

The first library selection utilized our NDT OE-PCR generated library and was selected against Fos. From the 20 candidates initially picked from the library plates, three of them made it through curing, back transformations, and spot plating without being eliminated. Western blot analysis of these three also showed that all three expressed an antibody and sequencing revealed that each had a different sequence. In order to determine if the new antibody was able to bind Fos, an ELISA was performed. On the first round of the ELISA, one of the candidates appeared to represent a slight improvement over that of the wild-type anti-Gcn4 IgG, however after retransforming and extracting the antibody, this was discovered to be a false result. While the first library did not successfully switch the specificity of the anti-Gcn4 IgG into an anti-Fos IgG, this was an expected result given the low diversity of the library.

The second attempt with this NDT OE-PCR library was an attempt to select the Gcn4 antibodies from the pool. Following recovery and plating as previously describe, 20 candidates were picked for further analysis. Of those 20, it was found that four of the candidates were the undigested parent antibody, as identified by sequencing analysis. While this result is not ideal, this is still an important finding and demonstrates that this selection is functional, however the library needs to be improved.

Having produced new NDT and NNK libraries using reverse PCR, new candidates have been generated and are currently being screened. Included in the new libraries are the NDT and NNK libraries against Gcn4 and the NNK library against Fos. There are two of each library for the Gcn4 since after the first library was plated, the control plates had colonies with various morphologies, suggesting that there was some form of

bias on the cultures. In order to ensure that this would not affect the library, fresh cells were transformed with the libraries to provide a duplicate in case the first set does not produce any suitable hits. For the Fos library, candidate colonies have been selected, however they have not began to enter the screening process.

CHAPTER 5

FUTURE WORK

5.1 Screening with current CDR-H3 libraries

The first step to finalize development of this platform is to use the current CDR-H3 libraries to select against a novel antigen. While attempts are being made to select against Fos, the Gcn4 antigen that the parent antibody binds has many similar structural and chemical properties, therefore the conversion from Fos to Gcn4 may pose less of a challenge than the conversion from Gcn4 to a peptide chain such as HAG. Therefore to demonstrate the ability of this selection platform, an attempt should be made to use the current libraries to identify proteins against a structurally diverse target.

5.2 Development of other CDR libraries

Antibodies have six CDR regions (H1-H3, L1-L3) however the libraries generated in this work only focus on the CDR-H3 region. In order to identify a truly novel antibody as well as demonstrate the broad applications of this platform, it is important to develop libraries that carry degeneracies in other CDR regions or combinations of these CDR regions. Creating these libraries will not only provide an opportunity to identify antibodies against a wider range of targets, but will also demonstrate the ability of this platform to accommodate and screen a library that is significantly larger and more diverse.

In addition to developing libraries that cover other CDR regions, it would be beneficial to also develop a library that covers CDRs of varying length. The libraries used in this study only comprise a five amino acid CDR with three variable residues,

however CDR-H3 has been known to extend to 22 amino acids, therefore in order to truly search for novel antibodies, the library needs to be expanded to include CDRs of varying length [56]. While the generation of these larger libraries will allow for the selection of novel antibodies against a target, they will also provide a means to affinity mature existing antibodies to generate stronger binding or more soluble antibodies by increasing the selective pressure.

5.3 Adaptation of clinically derived library

While synthetic libraries are useful for small-scale selections, there is a limit to the diversity and size that can be achieved. Clinically derived libraries on the other hand are exponentially larger and have a level of diversity that is unattainable from a synthetic standpoint. Therefore in order to achieve the true potential of this system it will become prudent to adapt a clinically derived library to this platform, which will provide the opportunity to sample a part of the antibody space that is simply inaccessible with synthetic libraries.

5.4 Selection of a biologically relevant antigen

The critical test for this platform will be its ability to screen a library to select a novel antigen against a clinically relevant antigen. While it is nice to select antibodies against common model and laboratory proteins, the true impact of this platform will not be realized until it is demonstrated that it is suitable to identify antibodies against novel targets that hold clinical relevance, such as cancer targets or autoimmune disease targets.

5.5 Other protein-protein interactions

While this platform was initially designed using antibody-antigen interactions, it would be interesting to see if other types of protein-protein interactions could be analyzed and reported by this platform. The ability to analyze other protein-protein interactions would significantly increase the range of interactions that could be interrogated. For example, the oncogenic signaling network requires numerous protein-protein interactions to develop, and this platform has potential to develop a cost effective method to identify interactions but also search for novel protein based inhibitors [57].

5.6 Protein epitope mapping

In this work it has been demonstrated that smaller antigens are more efficiently exported and provide a better basis for library selection. While a difference is observed in spot plating experiments with the larger proteins such as MBP and gpD, the background growth is to high for efficient library selections. One way to overcome the size limitations in this platform would be to take larger proteins and run them through epitope prediction software, breaking the larger protein into smaller, representative pieces that would be suitable for this platform. While this would provide a method to select antibodies against larger targets, this also could provide a method to analyze where a given antibody binds on the target protein. After identifying which fragment of the protein the antibody binds to, that fragment could be broken down further and further until you have a decent understanding of where on the protein this antibody binds.

5.7 Exploring the library population

In this work the selective pressure used to screen the library is always applied via solid LB-Agar plates. While the selection on solid media allows for the isolation of individual colonies compromised of identical cells, this method does not allow for real-time visualization of the progression of the selection process. An alternative approach for this technology would be to apply the selective pressure while the cells are maintained in liquid media, creating cell cultures containing a heterogeneous mixture of cell populations. Using next-generation sequencing, DNA from a large numbers of cells can be sequenced, providing a method to analyze these cultures and identify trends to visualize the path towards selection over time [58]. In addition to visualizing trends, this provides a method to count the number of occurrences for each library member, allowing for easy identification of high priority candidates.

5.8 High throughput analysis

As the size of the libraries used with this platform grows, the number of prospective candidates will also increase due to the nature of this platform to select antibodies with a wide range of affinities and solubility. The development of high throughput protocols for library selection and candidate characterization will lessen the burden and allow for the candidates to be processed in a timely manner. As a first step towards the adoption of high throughout technologies, next generation sequencing can be used to rapidly identify candidates that are represented multiple times, decreasing the number of candidates that need to be characterized.

5.9 Screening the library using next generation sequencing

In addition to following the evolution of the selection process in liquid culture, next generation sequencing provides a means to streamline the selection process by simplifying the amount of upfront screening required to identify candidates [59]. As the cultures are grown in liquid LB with selective media, sequential rounds of selection can be performed by subculturing and allowing the cells to repopulate fresh selective media. Throughout these rounds of selection, next generation sequencing can be used to observe the evolution of the library and to identify when candidates have been sufficiently enriched, signifying successful selection. Using multiple rounds of selection combined with next generation screening will allow for the focused characterization of a small number of candidates, greatly reducing the amount of energy required to eliminate false positives and other unsuccessful candidates.

CHAPTER 6

CONCLUSION

In this work it has been demonstrated that by hindering the transport ability of a CAT fusion protein through the Tat translocon, chloramphenicol sensitivity can be altered, thereby this can be used as a selective trait. To hinder the transport of the CAT complex, this platform takes advantage of the specificity and strength of antibody – antigen binding pairs. When the CAT-Ag fusion is in the unbound state, the Tat translocon is able to rapidly export the fusion into the periplasm, where the CAT protein is unable to neutralize chloramphenicol molecules, rendering the cell sensitive to chloramphenicol. When an antibody against the antigen is coexpressed with the CAT-Ag complex, the antibody binds to the CAT-Ag fusion, thereby inhibiting transport through the Tat translocon and allowing the CAT fusion to deactivate the chloramphenicol in the cytoplasm, enabling the cells to survive.

Recently, Robinson *et al.* demonstrated that full length IgGs are able to be expressed within the cytoplasm of *E.coli* cells, providing the basis for the development of a selection platform to screen antibody libraries within the cytoplasm of *E.coli* [60]. This work has demonstrated that the Ab-Ag interactions between MBP, gpD, HAG, Gcn4, cMyc and their respective antibodies are able to confer at least a two fold dilution difference when spot plated on selective media. In addition to displaying an even greater dilution difference compared to larger antigens, chloramphenicol concentrations were identified that effectively eliminated all background growth from the unmatched Ab-Ag cases for the smaller antigens.

Having successfully demonstrated via spot plating experiments that various Ab-Ag pairs confer greater resistance to chloramphenicol, mock library experiments were developed. We identified two types of mock libraries to explore, cell mixing and DNA mixing. Since smaller antigen constructs are more efficiently transported, the Gcn4 and HAG pairs were selected for these mock library studies. With the cell mixing experiments, satisfactory results were achieved, whereas with the DNA mixing the results were not as consistent. While in theory the DNA mixing experiments are more relevant to the final platform design, there are some practical differences that cause issues to arise with the DNA mixing that would not be relevant during the antibody library screening process. Therefore, even with lackluster results from the DNA mixing, the decision was made to push forward and screen the first antibody library.

The first library that was developed was chosen for simplicity and ease of use. By using a CDR-H3 NDT degeneracy, the number of library members required for full coverage would be significantly reduced. This degeneracy also provides an advantage by decreasing the number of non-functional members that would be created with some of the other degeneracy schemes from premature stop codons. The major issue with the first attempt at generating the library was the fact that the final product included a large quantity of non-functional self-ligated members as well as the non-digested parent antibody. In addition to the issues involving the creation of the library, the NDT degeneracy only provides 12 of the 20 possible amino acids, therefore the antibody selection potential is severely limited. To overcome these two issues, reverse PCR was utilized to develop two new libraries containing the NDT and NNK degeneracies, eliminating the need for any digestion steps, therefore significantly reducing the non-functional and non-digested members that plagued the first library.

Two initial library selections were attempted, the first attempted to switch the anti-Gcn4 IgG specificity towards Fos while the second attempted to recover the anti-Gcn4 IgG wild-type from the library pool. While the initial Fos selection was unsuccessful at identifying an anti-Fos IgG, this was not unexpected given the deficiency in the library creation and the limited amino acid diversity covered in the NDT degeneracy. The Gcn4 selection faired slightly better by isolating several candidates that turned out to be the original, undigested parent antibody. Even though the aim here was to recover the anti-Gcn4 antibody encoded by the NDT degeneracy, the sheer numbers of the undigested parent in the library made this unlikely. Current work is focused on applying the newly generated reverse PCR NDT and NNK libraries to the same selections as previously attempted, however the candidates acquired from these selections have not been processed and characterized yet.

APPENDIX

Appendix 1: Antibiotics and reporter concentrations

Antibiotic/Reporter	**Stock concentration (1000x)**	**Solvent**
Ampicillin (Amp)	100 mg/ml	Water
Carbenicillin (Carb)	100 mg/ml	Water
Chloramphenicol (Cm)	30 mg/ml	Ethanol
Kanamycin (Km)	50 mg/ml	Water
Spectinomycin (Sp)	50 mg/ml	Water
IPTG	0.1 M	Water

Antibiotics were used at the listed concentration unless otherwise noted

REFERENCES

1. Burmester, È., & Pezzutto, A. (2003). Color Atlas of Immunology. Thieme.
2. Marks, L. (2012). The birth pangs of monoclonal antibody therapeutics: The failure and legacy of centoxin. mAbs, 4(3), 403–412. http://doi.org/10.4161/mabs/19909
3. Todd, P. A., & Brogden, R. N. (n.d.). A Review of its Pharmacology and Therapeutic Potential Muromonab CD3, *899*(1 989), 871–899.
4. Reichert, J. M., Rosensweig, C. J., Faden, L. B., & Dewitz, M. C. (2005). Monoclonal antibody successes in the clinic, *23*(9), 1073–1078.
5. Yamada, T. (2011). Therapeutic Monoclonal Antibodies. *The Keio Journal of Medicine*.
6. Scott, A. M., Allison, J. P., Wolchok, J. D., & Hughes, H. (2012). Monoclonal antibodies in cancer therapy. *Cancer Immunity*, *12*(May), 1–9.
7. Weiner, G. J. (2015). Building better monoclonal antibody-based therapeutics. *Nature Publishing Group*, *15*(6), 361–370. http://doi.org/10.1038/nrc3930
8. Fesik, S. W. (2005). Promoting apoptosis as a strategy for cancer drug discovery. *Nature Reviews, Cancer*, *5*(October), 876–885. http://doi.org/10.1038/nrc1736
9. Natsume, A. (2009). Improving effector functions of antibodies for cancer treatment : Enhancing ADCC and CDC. *Drug Design, Development and Therapy*, *3*, 7–16.
10. Senter, P. D., & Sievers, E. L. (2012). perspective The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma. *Nature Biotechnology*, *30*(7), 631–637. http://doi.org/10.1038/nbt.2289
11. Vidarsson, G., Dekkers, G., & Rispens, T. (2014). IgG subclasses and allotypes : from structure to effector functions. *Frontiers in Immunology*, *5*(October), 1–17. http://doi.org/10.3389/fimmu.2014.00520

12. Jefferis, R. (2005). Glycosylation of Recombinant Antibody Therapeutics. *Biotechnology Progress*, *21*(1), 11–16.
13. *Thermo Scientific Avidin-Biotin Technical Handbook*. (2009). Thermo Fisher Scientific.
14. Wang, W. E. I., Singh, S., Zeng, D. L., King, K., Nema, S., & Al, W. E. T. (2007). Antibody Structure , Instability , and Formulation. *Journal of Pharmaceutical Sciences*, *96*(1), 1–26. http://doi.org/10.1002/jps
15. Kontermann, R. E., & Brinkmann, U. (2015). Bispecific antibodies. *Drug Discovery Today*, *20*(7), 838–847. http://doi.org/10.1016/j.drudis.2015.02.008
16. Buss, N. A. P. S., Henderson, S. J., McFarlane, M., Shenton, J. M., & De Haan, L. (2012). Monoclonal antibody therapeutics: History and future. *Current Opinion in Pharmacology*, *12*(5), 615–622. http://doi.org/10.1016/j.coph.2012.08.001
17. Harding, F. A., Stickler, M. M., Razo, J., & Dubridge, R. B. (2010). The immunogenicity of humanized and fully human antibodies Residual immunogenicity resides in the CDR regions. *mAbs*, *2*(3), 256–265.
18. Winter, G., & Harris, W. J. (1993). Humanized antibodies. *Immunology Today*, *14*(6), 243–6. http://doi.org/10.1016/0167-5699(93)90039-N
19. Robinson, M. K., Weiner, L. M., & Adams, G. P. (2004). Improving Monoclonal Antibodies for Cancer Therapy. *Drug Development Research*, *61*, 172–187. http://doi.org/10.1002/ddr.10345
20. Nelson, A., & Reichert, J. M. (2009). Development trends for therapeutic antibody fragments. *Nature Biotechnology*, *27*.
21. Presta, L. G. (2008). Molecular engineering and design of therapeutic antibodies. *Current Opinion in Immunology*, *20*(4), 460–470. http://doi.org/10.1016/j.coi.2008.06.012
22. Milstein, C., & Alerts, E. (1975). Continuous cultures of fused cells secreting antibody of predefined specificiy. *Nature*, *256*(5517).

23. McCafferty, J., Griffiths, a D., Winter, G., & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, *348*(6301), 552–554. http://doi.org/10.1038/348552a0
24. Barbas, C. F., Kang, A. S., Lerner, R. a, Benkovic, S. J., Iii, C. F. B., Kang, A. S., … Benkovict, S. J. (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(18), 7978–7982. http://doi.org/10.1073/pnas.88.18.7978
25. Boder, E. T. & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology*, 15, 553-557.
26. Hanes, J., & Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences*, *94*(May), 4937–4942.
27. Mazor, Y., Blarcom, T. Van, Mabry, R., Iverson, B. L., & Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nature Biotechnology*, *25*(5), 563–565. http://doi.org/10.1038/nbt1296
28. Marco, A. De. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. *Microbial Cell Factories*, *18*. http://doi.org/10.1186/1475-2859-8-26
29. Sevier, C. S., & Kaiser, C. A. (2002). Formation and transfer of disulphide bonds in living cells. *Nature Reviews, Molecular Cell Biology*, *3*(November), 836–847. http://doi.org/10.1038/nrm954
30. Lobstein, J., Emrich, C. a, Jeans, C., Faulkner, M., Riggs, P., & Berkmen, M. (2012). SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories*, *11*(1), 56. http://doi.org/10.1186/1475-2859-11-56
31. Berkmen, M. (2012). Production of disulfide-bonded proteins in Escherichia coli. *Protein Expression and Purification*, *82*(1), 240–251. http://doi.org/10.1016/j.pep.2011.10.009

32. Prinz, W. a, Åslund, F., Beckwith, J., & Holmgren, A. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. *Journal of Biological Chemistry*, *272*(25), 15661–15667. http://doi.org/10.1074/jbc.272.25.15661
33. Natale, P., Brüser, T., & Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane — Distinct translocases and mechanisms. *1778*, 1735–1756. http://doi.org/10.1016/j.bbamem.2007.07.015
34. Delisa, M. P., Tullman, D., & Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proceedings of the National Academy of Sciences*, *100*(10).
35. Robinson, C., Matos, C. F. R. O., Beck, D., Ren, C., Lawrence, J., Vasisht, N., & Mendel, S. (2011). Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria. *Biochimica et Biophysica Acta - Biomembranes*, *1808*(3), 876–884. http://doi.org/10.1016/j.bbamem.2010.11.023
36. Fisher, A. C., Delisa, M. P., Fisher, A. C., & Delisa, M. P. (2004). A Little Help from My Friends : Quality Control of Presecretory Proteins in Bacteria MINIREVIEW A Little Help from My Friends : Quality Control of Presecretory Proteins in Bacteria. *Journal of Bacteriology*, *186*(22). http://doi.org/10.1128/JB.186.22.7467
37. Röttig, A., & Steinbüchel, A. (2013). Acyltransferases in bacteria. *Microbiology and Molecular Biology Review*, *77*(2), 277–321. http://doi.org/10.1128/MMBR.00010-13
38. Brock, T. D. (1961). Chloramphenicol. *Bacteriology Reviews*, *25*(1), 32–48.
39. Vazquez, D. (1966). Binding of chloramphenicol to ribosomes The effect of a number of antibiotics. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, *114*(2), 277–288. http://doi.org/10.1016/0005-2787(66)90309-1

40. Shaw, W. V. (1967). The Enzymatic Acetylation of Chloramphenicol by Extracts of R Factor-resistant Escherichia coli*. *The Journal of Biological Chemistry*, *242*(4), 687–693.
41. Guzman, L. M., Belin, D., Carson, M. J., Beckwith, J., Guzman, L., Belin, D., & Carson, M. J. (1995). Tight regulation , modulation , and high-level expression by vectors containing the arabinose PBAD promoter . These include : Tight Regulation , Modulation , and High-Level Expression by Vectors Containing the Arabinose P BAD Promoter. *Journal of Bacteriology*, *177*(14).
42. Zverev, V. V., & Khmel, I. A. (1985). The nucleotide sequences of the replication origins of plasmids ColA and ColD. *Plasmid*, *14*(3), 192–199. http://doi.org/10.1016/0147-619X(85)90002-2
43. Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J., & Studier, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*, *56*(1), 125–135. http://doi.org/http://dx.doi.org/10.1016/0378-1119(87)90165-X
44. Sharma, R. C., & Schimke, R. T. (1996). Preparation of electro-competent E. coli using salt-free growth medium. *BioTechniques*, *20*(1), 42–44. http://doi.org/10.1007/s13239-011-0061-7
45. Wu, N., Matand, K., Kebede, B., Acquaah, G., & Williams, S. (2010). Enhancing DNA electrotransformation efficiency in Escherichia coli DH10B electrocompetent cells. *Electronic Journal of Biotechnology*, *13*(5), 1–9. http://doi.org/10.2225/vol13-issue5-fulltext-11
46. Quan, S., Hiniker, A., Collet, J.-F., & Bardwell, J. C. A. (2013). Bacterial Cell Surfaces. *Methods in Molecular Biology*, *966*, Chapter 22. http://doi.org/10.1007/978-1-62703-245-2
47. Reetz, M. T., Kahakeaw, D., & Lohmer, R. (2008). Addressing the numbers problem in directed evolution. *ChemBioChem*, *9*(11), 1797–1804. http://doi.org/10.1002/cbic.200800298

48. Kanwar, M., Wright, R. C., Date, A., Tullman, J., & Ostermeier, M. (2013). Protein Switch Engineering by Domain Insertion. *Methods in Enzymology*, *523*. http://doi.org/10.1016/j.surg.2006.10.010.Use
49. Tolia, N. H., & Joshua-Tor, L. (2006). Strategies for protein coexpression in Escherichia coli. *Nature Methods*, *3*(1), 55–64. http://doi.org/10.1038/nmeth0106-55
50. Velappan, N., Sblattero, D., Chasteen, L., & Pavlik, P. (2007). Plasmid incompatibility: more compatible than previously thought? *PEDS*, 1–5. http://doi.org/10.1093/protein/gzm005
51. Lander, G. C., Evilevitch, A., Jeembaeva, M., Potter, C. S., Carragher, B., & Johnson, J. E. (2008). Bacteriophage Lambda Stabilization by Auxiliary Protein gpD: Timing, Location, and Mechanism of Attachment Determined by Cryo-EM. *Structure*, *16*(9), 1399–1406. http://doi.org/10.1016/j.str.2008.05.016
52. Zeng, X., Zhu, H., Lashuel, H. a, & Hu, J. C. (1997). Oligomerization properties of GCN4 leucine zipper e and g position mutants. *Protein Science*, *6*(10), 2218–2226. http://doi.org/10.1002/pro.5560061016
53. Fujiwara, K., Poikonen, K., Aleman, L., Valtavaara, M., Saksela, K., & Mayer, B. J. (2002). A single-chain antibody/epitope system for functional analysis of protein-protein interactions. *Biochemistry*, *41*(42), 12729–12738. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12379115
54. Waraho, D., & Delisa, M. P. (2009). Versatile selection technology for intracellular protein – protein interactions mediated by a unique bacterial hitchhiker transport mechanism. *Proceedings of the National Academy of Sciences*, *106*(10).
55. Maur, A. A. Der, Zahnd, C., Fischer, F., Spinelli, S., Cambillau, C., Chem, J. B., … Plu, A. (2002). Direct in Vivo Screening of Intrabody Libraries Constructed on a High. *Journal of Biological Chemistry*, *277*. http://doi.org/10.1074/jbc.M205264200

56. Mahon, C. M., Lambert, M. A., Glanville, J., Wade, J. M., Fennell, B. J., Krebs, M. R., … Finlay, W. J. J. (2013). Comprehensive interrogation of a minimalist synthetic CDR-H3 library and its ability to generate antibodies with therapeutic potential. *Journal of Molecular Biology*, *425*(10), 1712–1730. http://doi.org/10.1016/j.jmb.2013.02.015
57. Ivanov, A. A., Khuri, F. R., & Fu, H. (2013). Targeting protein-protein interactions as an anticancer strategy. *Trends in Pharmacological Sciences*, *34*(7), 393–400. http://doi.org/10.1016/j.tips.2013.04.007
58. Goodwin, S., Mcpherson, J. D., & Mccombie, W. R. (2016). Coming of age : ten years of next- generation sequencing technologies. *Nature Publishing Group*, *17*(6), 333–351. http://doi.org/10.1038/nrg.2016.49
59. Ravn, U., Gueneau, F., Baerlocher, L., Osteras, M., Desmurs, M., & Malinge, P. (2010). By-passing in vitro screening — next generation sequencing technologies applied to antibody display and in silico candidate selection, 1–11. http://doi.org/10.1093/nar/gkq789
60. Robinson, M., Ke, N., Lobstein, J., Peterson, C., Szkodny, A., Mansell, T. J., … Berkmen, M. (2015). Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria. *Nature Communications*. http://doi.org/10.1038/ncomms9072