Validation of a Minimalist, Synthetic Yeast Displayed Antibody Library

By Jacob B. Lissoos

Advised by Professor James A. Van Deventer

An honors thesis submitted to the Department of Chemical and Biological Engineering

Tufts University

Medford, MA

Draft 1: April 25th, 2018

For Tammy Kaplan (Z"L), Randy Kaplan, and Sam Blekhman. Stay strong and keep fighting, I will not forget you. I promise that this is just the first step in a career dedicated to finding a cure for you.

Acknowledgements

I would first like thank my thesis advisor Professor James Van Deventer for his guidance and support during my research experience at Tufts. I remember first meeting him while interviewing him as an assignment for a class, and he ended up not only teaching me how to avoid flooding pipettes, but also heled me to find my passion within chemical engineering. Professor Van Deventer goes above and beyond his responsibilities to make sure that I reach my fullest potential as an independent scientist and thinker, and for that I am extremely grateful. I truly could not have asked for a better mentor.

I would also like to thank Professors Nikhil Nair and Yu-Shan Lin for volunteering their free time to serve on my thesis committee. I have learned so much from both of you during my time at Tufts, both in class and through research, and I am extremely grateful for all that you have taught me.

Thirdly, I would like to thank all my prior mentors outside of Tufts: Drs. Jason Wertheim, Emad Darveshi and Girma Woldemichael. The training that each of you provided me with helped me to see new perspectives on translational biomedical sciences, and I grew tremendously as a researcher during my time working with each of you. Thank you for your guidance during the time I spent with you, and for your continued support in helping me to figure out my long-term career goals.

Finally, I must thank all the members of the Van Deventer Lab, past and present, who have been my co-workers for the past two years. Thank you to Jessica Steiglitz, Haixing Kehoe, Chris Ghadban, Laura Quinto, Gregory Burman, Sheikh Mahatabuddin and more recently Arlinda Rehzodo, Mariha Islam, Kelley Potts, and Manjie Huang. You all have provided tremendous support to me during my time in the lab, both through training me and providing outside perspectives when I was trouble shooting problems. Outside of lab, you all are also amazing friends and I have really enjoyed the time I have spent with each of you. A special thank you to Jessica Steiglitz for always making sure the lab was fully stocked and running smoothly, and for always selflessly offering to help with experiments.

Abstract

Recent advances in cancer immunotherapy have demonstrated promising results in the clinic. However, the success of these therapeutics depends on the availability of a targetable surface antigen. Moreover, this treatment approach introduces selective pressures to rapidly evolving tumors which can ultimately result in a fatal drug resistant relapse. These disadvantages could be mitigated by instead targeting an underlying mechanism of tumorigenesis. Moreover, the resulting new class of therapeutics could serve as "off the shelf" treatments for multiple types of cancer.

Certain extracellular proteases that are overexpressed in the tumor microenvironment catalyze a wide range of tumorigenic processes such as metastasis, angiogenesis, activation of oncogenic signaling pathways and the activation of other proteases. They would therefore be viable targets for this new type of generalized cancer therapeutic. Targeting these enzymes with traditional small molecule or antibody approaches remains challenging because of off-target effects and bulky topology around enzyme active sites.

We hypothesize that by conjugating small molecules to strategic locations on an antibody to generate protein-small molecule hybrid inhibitors, we can create a synergy between these two inhibition strategies and therefore overcome the disadvantages of each approach. These inhibitors could have translational applications as cancer therapeutics and could also be used as molecular biology probes to further explore the role of extracellular proteases in the tumor microenvironment. Because of the novelty of this approach, little is known about which small molecules to use, or the optimal positioning of these small molecules within an antibody to promote the synergy. A randomized library approach would thus be a valuable tool to be able to simultaneously evaluate different combinations of these parameters.

Herein, we describe the validation of a previously constructed minimalist yeast displayed scFv library with variation restricted to the CDRH3 loop. All other antibody loops are un-diversified and left at the germline sequence and therefore do not significantly contribute to antigen binding. This dramatically expands the positions at which small molecules can incorporated without disrupting antigen binding, as compared to a canonical antibody. To validate the library, we isolated binding clones against a panel of model antigens and evaluate them for specificity to the target antigen, extent of off-target binding to similar antigens, and affinity. Our results show that we were able to isolate binding scFvs from the library to a panel of model target antigens using a combination of magnetic-bead based and flow cytometry approaches. Isolated binders had triple digit nano-molar affinity for target antigen and showed no off-target binding to structurally similar antigen. When taken together, these results suggest that the CDRH3 library as initially constructed can serve as a protein scaffold into which additional chemical functionality can be incorporated. This library can have broader applications in other cases where it is desired to engineer additional functionality into antibodies, such as in motif grafting.

Table of Contents

List of Figures

List of Tables

Validation of a Minimalist, Synthetic Yeast Displayed Antibody Library

Chapter 1: Introduction

1.1: The Tumor Microenvironment as a Target for Cancer Therapeutics

The World Health Organization predicts that cancer diagnosis will increase by 50% and cancer deaths will increase by 60% by 2030[10]. Recent advances in cancer immunotherapies such as chimeric antigen T cells and monoclonal antibodies have demonstrated high efficacy in the clinic[11-15]. These therapeutics function by targeting a specific tumor surface antigen with low enough basal expression in the surrounding tissue[14-18]. However, not all cancers have antigens amenable to targeting by this approach and characterizing the antigen expression profiles of rare and under-funded cancers remains a challenge[19-20]. Moreover, targeted immunotherapies introduce a selective pressure to rapidly evolving tumors, which can ultimately lead to relapse with fatal drug resistant disease[15, 18-19]. Thus, there is a clear need for a new class of cancer therapeutics that can target the underlying mechanism of cancer progression instead of specific tumor proteins. Such a therapeutic would have the potential to treat diverse types of cancer, all while reducing potential for drug resistant disease relapse.

The tumor microenvironment is an amenable target for this type of therapeutic because of its role in both early disease progression and metastasis, as well as its uniformity across different metastatic cancers[21-31]. The role of the tumor microenvironment in driving disease progression and metastasis has been linked to the overexpression of extracellular proteases including certain zinc endopeptidases[24-26, 30, 32-33] and serine hydrolases[22, 28, 31, 34-35] as well as other enzymes such as lysyl oxidases[27].

Figure 1.1 depicts different mechanisms by which these enzymes exert their tumorigenic effects. On the catalytic level, they cleave and remodel the extracellular matrix (ECM) to provide a path for tumor metastasis and angiogenesis. This ECM cleavage is necessary for the formation of pre-metastatic pockets which circulating tumor cells can eventually colonize[6, 27, 29, 32-33]. The enzymes of the tumor microenvironment also free membrane-bound signaling molecules such as EGF, FGF and WNTs into soluble form which can then regulate oncogenic signaling pathways[6, 25, 32-33]. Moreover, these enzymes activate other tumorigenic enzymes[32-33]. Interestingly, the functions of enzymes of the tumor microenvironment are not limited to catalysis[32]. Using motifs far from their active sites, they can bind to certain receptors and activate pathways that up-regulate cell migration and proliferation[36], as well as regulate local concentrations of ligands linked to cancer malignancy[37].

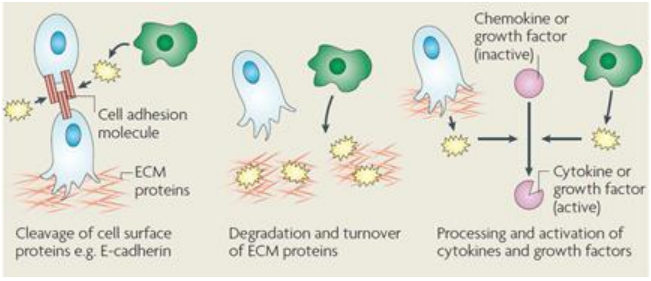

Figure 1.1. Tumorigenic proteolytic activity of the proteases of the tumor microenvironment. The proteases that are upregulated in the tumor microenvironment mediate their tumorigenic effects via 3 main mechanisms: cleavage of ECM to provide a path for metastasis and blood vessel recruitment, cleavage of membrane anchored signaling molecules to release into soluble form, and activation of other proteases. Taken from Joyce et al. (2009)[6].

Unsurprisingly, there have been numerous efforts to develop cancer therapeutics targeting the enzymes of the tumor microenvironment. Early efforts focused on small molecules that chelated catalytic ions, or mimicked enzyme substrates[22-23]. However, these inhibitors were promiscuous and had off-target interactions with different members of multi-enzyme families[23, 38-39]. This proved to be detrimental to their success since different members of the same enzyme family catalyze essential physiological functions and some are even tumor suppressive[25]. Consequently, patients in trials experienced debilitating side effects and most of these inhibitors failed out of clinical trials[23, 38-40].

In light of the shortcomings of small molecule-based inhibitors, recent efforts at targeting the enzymes of the tumor microenvironment have shifted to using immunizations or antibody libraries to

screen for selective binders that can discriminate between different members of enzyme families[41-45]. While this approach is effective at isolating specific binders to a given target, it is extremely unlikely to yield inhibitory antibodies[41, 46]. The enzyme active site is usually located within the bulkiest epitope on the surface, and therefore traditional antibodies are too sterically hindered to access the enzyme active site[47]. Still, some inhibitory antibodies against enzymes of the tumor microenvironment have been isolated. However, these were the result of complicated and labor-intensive screening campaigns[46, 48-49]. Moreover, this approach is not comprehensive, evident in the fact that inhibitory antibodies have only been isolated against a limited number of the oncogenic tumor microenvironment enzymes[25, 49]. In alternative screening strategies, camelid VHH antibody scaffolds have been used to overcome the challenging steric hinderance surrounding enzyme active sites[50]. However, camelid antibodies may be immunogenic when used in therapeutic applications[51].

Alternative protein-based inhibition strategies have attempted to exploit the fact that certain enzymes of the tumor microenvironment have natural inhibitors. For example, the Matrix Metalloproteinases (MMPs) and A Disintegrin and Metalloproteins (ADAMs) share a promiscuous set of naturally occurring inhibitors called Tissue Inhibitors of Metalloproteinases (TIMPs)[52]. Attempts have been made to use computationally guided mutagenesis and directed evolution to engineer TIMPs with boosted specificity. However, the resulting inhibitors only showed mild improvements in specificity and still had off target inhibition of other Metzincin super-family members[53-54]. Moreover, this approach is limited to enzymes that have known naturally occurring inhibitors, which is not the case for certain serine hydrolases and lysyl oxidases[22, 27].

Thus, there is an urgent need to develop a novel strategy for the engineering of a more comprehensive, selective and potent set of inhibitors for the enzymes of the tumor microenvironment. If successful, these inhibitors would be applicable as an "off the shelf" approach to treating metastatic cancer and could also be used as a molecular biology probe to elucidate more information on the roles of these enzymes in cancer progression.

1.2: Overview of Antibody Discovery Strategies

Despite numerous challenges, antibody inhibitors of some of the oncogenic enzymes of the tumor microenvironment have demonstrated their efficacy in mouse models and clinical trials[41, 43, 55]. Computational and rational design approaches to design binding proteins against a given target have yet to yield binders with a high enough affinity and high enough success rates to be useful for therapeutic applications[56-59]. Therefore, animal immunizations or directed evolution with synthetic and natural libraries remain the favored approaches to isolate antibodies[60]. Importantly, all three of these approaches have previously yielded inhibitory antibodies of specific enzymes of the tumor microenvironment[27, 41, 49].

Figure 1.2 provides an overview of the synthetic antibody library construction process. In these libraries, the location of antibody diversity, amino acid diversity profile and complimentary determining region (CDR) loop length diversity profile can all be precisely specified[61-62]. A heavy and light chain germline sequence are first chosen. These germline sequences act as the scaffold antibody for the library, on which the desired amino acid and loop length diversity can be incorporated[63]. Libraries can include single[64] or multiple[65-66] germlines. Germlines are selected based on favorable expressibility and solubility, as well as ability to give rise to binders to the specific target of interest[64, 67-69]. Primers are designed so the germline DNA can be amplified in a PCR reaction. Degenerate codons are incorporated into specific primer positions corresponding to where amino acid diversity is desired. NNS and NNK codons are favored for degeneracy incorporation because of a lower probability of stop codon incorporation, which would result in a truncated clone[63]. Carefully selected mixtures of DDNTPs ensure incorporation of a pre-

determined amino acid diversity profile[70]. Primers are also synthesized with variable length in accordance to the loop length profile[7, 63].

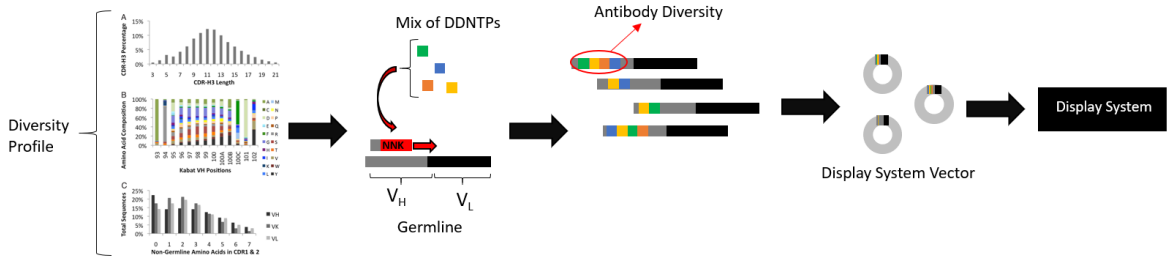

Figure 1.2. Construction of a synthetic antibody library. An antibody loop length and amino acid library diversity profile is first selected. The germline sequence is then amplified with primers designed to have degeneracy at specific positions. Custom mixtures of DDNTPs ensure the desired amino acid profiled is re-created. The resulting DNA fragments from the PCR reaction contain the germline sequence in addition to incorporated amino acid diversity. These fragments are cloned into vectors for the display system of choice so that the resulting antibody library can be screened. Specifics on display systems are discussed in section 1.3.

Synthetic library design allows for precise control over the location of diversity within the antibody. Certain locations of the antibody can be diversified, allowing them to mediate antigen binding. Other locations can be left undiversified. This property of synthetic libraries has been useful in elucidating the role of certain antibody loops in mediating antigen binding[64, 70] and in engineering additional functionality into certain areas of an antibody[71-72].

In contrast, natural antibody libraries and immunizations require the usage of full antibodies and are not as amenable to further engineering[60, 65]. Each of these strategies has additional pitfalls as compared to synthetic antibody libraries. In natural libraries, not all antibodies fold correctly and express well during screening[61, 73]. This could result in missing potential hits to a target antigen. Immunizations are limited by antigens that can be recognized as non-self and can therefore not be used to isolate antibodies against human targets that have a closely related murine analogue unless the antigen is extensively modified[60]. Thus, synthetic libraries remain the favored choice for applications where further engineering of isolated proteins is desired.

1.3: Overview of Display Technologies

A display system is needed to use directed evolution to isolate antibody binders to a given target[74]. This system provides a link between genotype and phenotype during screening, where the genotype is the DNA sequence that encodes for a given antibody and the phenotype is the antibody itself[75-76]. When a binder of interest is isolated, the corresponding DNA can be easily sequenced and used to mass produce soluble antibody. The main display technologies for antibody library screening are phage display[75], yeast display[77], ribosome display[76] and mammalian cell display[74, 78].

An overview of yeast display is presented in Figure 1.3. In wild type *Saccharomyces cerevisiae*, the α-agglutinin transmembrane protein fuses haploid cells together to form a diploid. In the *RJY100* strain engineered for yeast display, an antibody fragment is fused to the Aga2p sub-unit of α- agglutinin[77]. Because yeast are eukaryotic, post translational modifications are carried out with higher fidelity than in phage or ribosome display[73]. Yeast display up to 40,000 constructs on the cell surface[79], which creates an avidity effect during screening[9]. This means that weaker clones can be isolated than in ribosome display, which only displays 1 constructs on the surface[76]. This is important because weaker clones may be more amenable to further engineering than stronger clones.

Yeast display has several advantages over ribosome and phage display. To begin with, yeast display is compatible with flow cytometry[77]. This can be used to sort libraries with florescent activated cell sorting (FACS)[66], as a tool to assess progression of a library screen, or to estimate the K_D of an isolated antibody using titrations on the yeast surface[79]. Other assays can also be done on the yeast surface

without the need for production of soluble antibody such as target neutralization or enzyme inhibition assays[48]. Also important is the capability of yeast amber suppression machinery to incorporate non-canonical amino acids into displayed antibodies, which in turn allows for the evaluation of bioconjugates on the yeast surface[80]. For additional assays that require soluble antibody, yeast can secret small amounts of protein at the end of screening with minimal sub-cloning[81]. This can be useful for pilot testing before large scale mammalian tissue culture. Yeast displayed libraries are typically screened using antigen coated magnetic beads[9], flow cytometry[66], or bio panning[82].

Mammalian cell display has all the advantages of yeast display, with improved post translational modification fidelity and better antibody folding and expression than in yeast. However, because of transfection efficiency the library size is limited to 10^{5-7}. Such a library is too small for a naïve antibody library and is more useful for affinity maturation[78]. Moreover, grow-up after a round of sorting is limited by the doubling time of the HEK 293 cells, which is around 24 hours[78]. Thus, yeast display is extremely advantageous for the engineering of novel protein therapeutics because of all the assays that can be done on the yeast surface to verify functionality before the production of soluble protein is required.

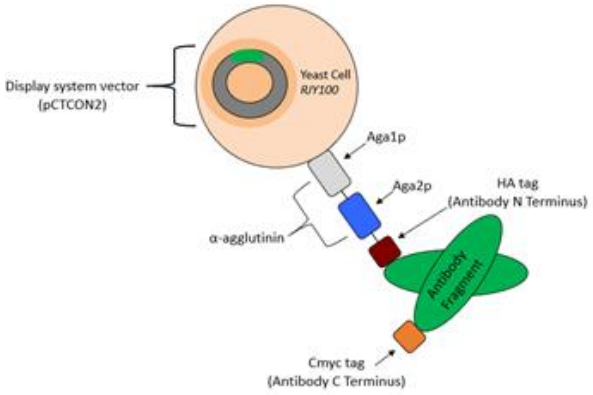

Figure 1.3. Overview of yeast display system**.** Yeast display system uses an engineered strain of *Saccharomyces Cerevisiae*, RJY100, which allows for an antibody to be fused to the Aga2p surface protein. An N-terminal HA tag and a C-terminal Myc tag are also included, which are useful in flow cytometry to label for full length vs truncated clones. Only one antibody fragment is shown on the yeast surface for simplicity, although approximately 40,000 constructs are displayed on the surface of any given cell. Figure adopted from James A. Van Deventer.

1.4: Strategies for Engineering Further Functionality Into Proteins

The functionality of canonical antibodies can be significantly expanded by engineering peptide motifs or small molecules into the antibody structure. One such strategy for further engineering antibodies is through motif grafting, which is shown in Figure 1.4. In antibody motif grafting, a canonical CDR loop of an antibody is replaced with a specific fragment from another peptide. Grafting is useful to alter the specificity of a pre-existing antibody[83-84], to humanize murine antibodies through the generation of chimeric antibodies[60] and to add additional functionality to a pre-existing antibody[71-72]. For example, antibodies that bound a given protein were made to only recognize the phosphorylated form when a phosphate binding nest was grafted into the CDRH2 loop[72]. Motif grafting can also be used to add inhibitory functionality to an otherwise non-inhibitory antibody. This strategy has been used to replace the CDRH3 of EGFR binding antibodies with the EGFR dimerization loop[71]. Resulting clones bound to and inhibited EGFR dimerization, thereby blocking a signaling pathway implicated in cancer.

In addition to protein-based functionality, chemical functionality can also be added into proteins using a variety of approaches. In bio-conjugation, a non-canonical amino acid is first added into the protein using site specific incorporation and then the desired chemical group can be conjugated to the non-canonical amino acid. This strategy has been used to add post translational modifications to proteins such as glycosylation[85] and phosphorylation[86]. The strategy has also been used to engineer hybrid therapeutics that have both protein and small-molecule components. For example, antibody drug

conjugates use this strategy to conjugate small molecule drugs to antibody framework regions for the targeted delivery of a toxic payload[87]. In some cases, a non-canonical amino acid alone can be used to engineer chemical functionality into a peptide, without the need for the conjugation of a small molecule. For example, residue specific incorporation has been used to replace proline with a boronic acid analogue in a peptide substrate of DPPIV, a serine hydrolase implicated in cancer. This modified the substrate into a potent inhibitor of the protease[88]. Different approaches use non-covalent strategies to engineer chemical functionality to proteins. For example, a promiscuous small molecule kinase inhibitor was made specific to protein kinase A by tethering it to a cyclic peptide that selectively bound to that kinase. Self-assembly of a coiled-coil motif around the small molecule was used to attach it to the cyclic peptide[89]. Thus, there are many ways to expand basal protein diversity to yield engineered proteins with expanded functionality over canonical forms.

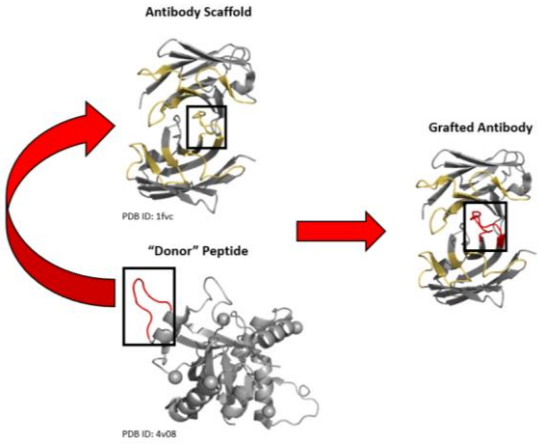

Figure 1.4. Schematic of antibody grafting. Depiction of antibody grafting. Peptide motif to be grafted into the antibody is highlighted in red. CDR loops of the antibody are highlighted in yellow, and the loop to be replaced by the grafted peptide is boxed in black. The result is the original antibody, with one CDR loop replaced by the sequence of the grafted peptide, which is highlighted in red in the final antibody structure.

1.5: Protein-Small Molecule Hybrids as a Novel Approach to Inhibiting the Enzymes of the Tumor Microenvironment

The disadvantages of small molecule and antibody inhibition approaches of the enzymes of the tumor microenvironment may be overcome by combining the two to yield protein-small molecule hybrids, which is illustrated in Figure 1.5. These proposed structures rely on a powerful synergy in which the antibody selectively delivers a small molecule close to the active site of a specific enzyme and the small molecule than inhibits the catalytic activity of the enzyme.

Protein-small molecule hybrids differ from traditional antibody-drug conjugates in that the antibody and small molecule act synergistically instead of the antibody simply delivering the small molecule to its target, and the small molecule then acting alone. This approach is powerful in that it allows for non-inhibitory antibodies to be engineered to become inhibitory. Moreover, while traditional antibodies dissociate off a target, protein-small molecule hybrids could in principle be engineered to be irreversible with certain small molecules. These inhibitors could have translational applications as cancer therapeutics and could also be used as molecular biology probes to further explore the role of extracellular proteases in the tumor microenvironment.

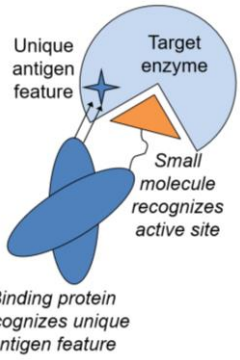

Figure 1.5. Proposed protein-small molecule hybrid function**.** The protein-small molecule hybrids combine the superior specificity of antibodies with the proven enzyme inhibition capabilities of small molecules to yield a potent synergistic inhibitor. Figure taken from James A. Van Deventer.

The small molecule will be conjugated to the antibody at a non-canonical amino acid residue that will be site specifically incorporated within the antibody. The non-canonical amino acid will be carefully positioned such that the orientation of the conjugated small molecule is optimal for interaction with an enzyme. To summarize, protein-small molecule hybrids combine traditional antibody engineering approaches overviewed in section 1.2 with methods of adding additional chemical functionality to proteins as shown in section 1.4 resulting in a novel class of enzyme inhibitors.

1.6: Construction of a Minimalist Antibody Library Capable of Accommodating Small Molecule Functionality

Because of the novelty of the protein-small molecule hybrid concept, little is known about which small molecules to use, or the optimal positioning of these small molecules within an antibody to best promote synergy. A randomized library approach would therefor be a valuable tool to be able to simultaneously evaluate different combinations of these parameters, without the need to obtain complicated structural information on antibody mode of binding.

Such a library was previously constructed by Haixing Kehoe and James Van Deventer[7]. To maximize the number of positions in the antibody that can be evaluated for small molecule incorporation, a minimalist synthetic design approach was used to diversify as little of the antibody as possible. In theory, only diversified regions of the antibody should significantly contribute to target binding, and un-diversified regions can then be explored for small molecule incorporation.

Previous studies from immunology provided cues as to what the minimal level of antibody diversity is that can still mediate binding to a target antigen. The heavy chain of an antibody alone is sufficient to mediate contact with a target antigen. For example, certain camelid species have been shown to have antibodies composed only of heavy chains[90]. Within the heavy chain, the CDRH3 loop has been shown to be the dominant mediator of antigen binding[70]. This loop is the main source of genetic diversity within the antibody[91]. Mice with antibody repertoires that had been altered to eliminate all diversity besides that of the CDRH3 yielded high affinity and specific binders to a wide range of targets when immunized[92]. In computational chemistry studies, the CDRH3 was shown to contribute most to the free energy of binding as compared to other antibody loops[93]. In antibodies capable of binding to sterically hindered targets such as the GP120 envelope of the HIV virus, abnormally long CDRH3s were found to be the main mediators of binding[94]. When taken together, these previous studies suggest that in principle, the CDRH3 alone should be sufficient to mediate antibody-antigen interactions in a minimalist antibody library. Indeed, such a library has previously been constructed in the phage display format[64], but this library was not used for grafting approaches.

The library constructed by Haixing Kehoe and James Van Deventer is modeled after the phage displayed CDRH3 library originally described by Mahon and colleagues[64]. However, our library is constructed in the yeast displayed format

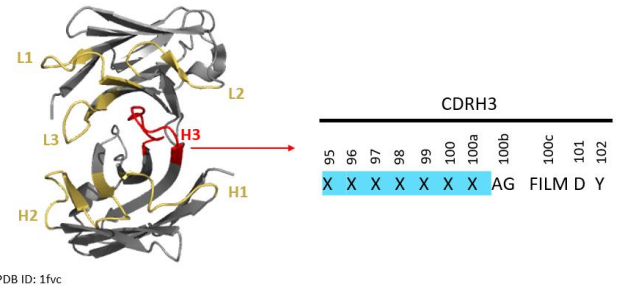

Figure 1.6. Yeast displayed library diversification approach. Shown is an scFv with the antibody framework colored in dark grey. CDR loops that would ordinarily mediate binding to target antigen but are left at the germline sequence in the library, are colored in yellow. The CDRH3, the only source of diversity in the library, is colored in red. Also shown is the CDRH3 diversification scheme. All amino acid residues in framework 2 are conserved. Thereafter, the CDRH3 region begins with a variable length stretch spanning 5-13 amino acids, which is highlighted in light blue. "X" denotes any amino acid incorporated according to the frequency found in the human B cell repertoire. Thereafter, there are 2 positions with a more restricted amino acid diversity. The final two amino acids that make up the CDRH3 are conserved as naturally found in the JH4 germline segment[8]. Numbering shown according to the Kabat scheme[3].

because of advantages in being able to screen for bioconjugates[80] and evaluate enzyme inhibition[79] all on the yeast surface. As shown in Figure 1.6, genetic diversity of the library is restricted to the CDRH3. All other antibody CDRs that usually mediate contact with antigen were left un-differentiated at the germline sequence. The antibody germline sequences were selected based off those used in the previously described CDRH3 library[64]. The DPK9-JK4 germline was used for the light chain and the DP47-JH4 germline for the heavy chain. These germlines have favorable expressibility and solubility characteristics and have been shown to be thermodynamically stable[64, 68]. These germlines appear in the human B cell repertoire with a high frequency[64].

Figure 1.6 shows the library diversification scheme. Diversification begins at the first residue in the CDRH3 according to the Kabat numbering scheme[3], and spans a variable length of 5-13 amino acid residues thereafter. The amino acid repertoire at these positions is randomized to include all 20 canonical amino acids, with the probability of incorporation mimicking that found in the human B cell repertoire. In the profile, the frequency of cystine was kept low because disulfide bonds can disrupt antibody structure[95]. Similarly, tryptophan appears infrequently in the library because this residue is implicated in polyspecificity[96]. In contrast, the frequency of tyrosine is high because this residue has been shown to be the dominant residue in mediating interactions with antigen[97]. After the variable length stretch of amino acid diversity, there are two positions with a more limited amino acid diversity. The final two positions in the CDRH3 from the JH4 segment of the germline[64] are left undiversified. The full range of possible CDRH3 loops for our antibody are 9-17, which mimics part of the distribution in the natural repertoire of 2-26[3]. Because the CDRH3 library sought to mimic human amino acid diversity and loop lengths, it is considered a natural mimic library.

The library has $1.3x10^9$ total clones. Of those clones, an estimated $1.1x10^9$ are full length, with the remainder being truncated. This library is the maximum size that can screened using yeast display at liter scale[77].

Because of the protein expression limits of the yeast display system, the library is not a full-length IgG library, and instead is a single chain variable fragment library (scFv). scFvs contain both the heavy and light chain of the antibody variable fragment, joined together by a glycine rich linker[90]. The antibody Fc region, as well as the second Fab are missing from an scFv. If full length IgGs are desired, scFvs can be cloned into an Fc at the end of screening efforts to create an scFv-Fc fusion. To summarize, the CDRH3 library will act as a protein scaffold into which additional chemical functionality can be incorporated into any of the un-diversified regions of the antibodies.

1.7: Description of Thesis Work

The baseline protein diversity of the eventual hybrid library must first be validated before additional chemical functionality can be incorporated, as shown in Figure 1.7. This is the topic of this thesis work. Other levels of library diversity will be discussed in future works.

The end goal for the library as descried in Section 1.6 will be to expand it from the basal CDRH3 amino acid and loop length diversity to include diversity in positioning and types of small molecules incorporated into the scFvs. The resulting library can be used to screen for protein-small molecule hybrids against an enzyme of interest.

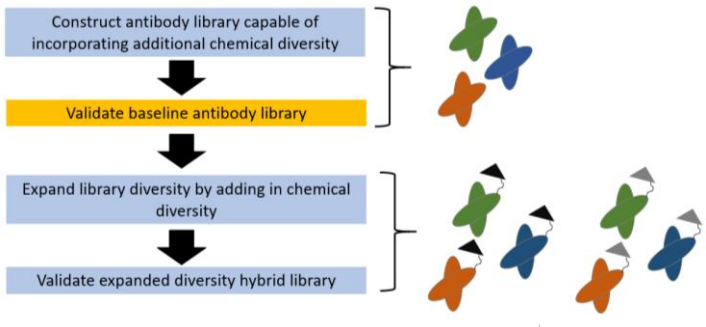

Figure 1.7. Description of thesis work. All steps required to developing the protein-small molecule hybrid library are described in sequential blue boxes. The box colored yellow is the focus of this thesis work. The level of library diversity corresponding to each step is shown to the right. Figure adapted from James A. Van Deventer.

The methods for all experiments conducted as a part of this thesis work are described in chapter 2. Before validating the library, the lab's antigen preparation and library screening protocols had to first be optimized, which will be discussed in chapter 3. Then, the library could be validated as described in chapter 4 by screening against structurally similar antigens: IgGs from different animal species. This serves as a proof of concept for the goal of the library, which is to target specific enzymes of superfamilies for inhibition. Resulting clones were probed for cross reactivity to different IgGs and their dissociation constant to target antigen. Lastly, we propose modifications to the initial library design to allow for isolation of binders to smaller enzymatic antigens, which is described in conclusions and future work.

1.8: References

1 *Cancer Facts Sheet*, <http://www.who.int/mediacentre/factsheets/fs297/en/> (2018).

2 Paradoll, D. The blockade of immune checkpoints in cancer immunotherapy *Nature Reviews Cancer* **12**, 252-264 (2012).

3 Chang, Z. & Chen, Y. CARs: Synthetic Immunoreceptors for Cancer Therapy and Beyond *Cell* **23**, 430-450 (2017).

4 Smith, I. *et al.* 2 year follow-up of transtuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet* **369**, 29-36 (2007).

5 Hinrichs, C. Cell-based Molecularly Targeted Therapy: Targeting Oncoproteins With T Cell Receptor Gene Therapy. *Journal of Clinical Investigation* **128**, 1261-1263 (2018).

6 Hing, B., Van Den Heuval, P., Prabhu, V., Zhang, S. & El-Deiry, W. Targeting tumor suppressor p53 for cancer therapy: strategies, challenges and oppertuniteis. *Current Durg Target* **15**, 80-89 (2014).

7 Joyce, J. & Pollard, J. Microenvironmental regulation of metastasis *Nature Reviews Cancer* **9**, 239-252 (2009).

8 Li, N., Fu, H., Hewitt, S., Dimitrov, D. & Ho, M. Therapeutically targeting glypican-2 via single-domain antibody-based chimeric antigen receptors and immunotoxins in neuroblastoma. *PNAS* **114**, 6623-6631 (2017).

9 Wang, R. & Wang, H. Immune targets and neoantigens for cancer immunotherapy and precision medicine. *Nature Cell Research* **27**, 11-37 (2017).

10 Alatrash, G., Jakher, H., Stafford, P. & Mittendorf, E. Cancer immunotherapies, their safety and toxicity *Expert Opinion on Drug Safety* **12**, 631-645 (2013).

11 Jespers, L., Schon, O., James, L., Verprinstev, D. & Winter, G. Crystal structure of HEL4, a soluble, refoldable human VH single domain with a germ-line scaffold *Journal of Biological Chemistry* **337**, 893-903 (20004).

12 Kabat, E., Wu, T., Perry, H., Gottesman, K. & Foeller, C. *Sequences of proteins of immunological interst* Vol. 1 (National Institutes of Health, 1991).

13 Zugazzgoitia, J. *et al.* Current Challanges in Cancer Treatment *Clinical Therapeutics* **38**, 1551-1566 (2016).

14 Komatsubara, K. & Carvajal, R. The promise and challenges of rare cancer research. *Lancent Oncology* **17**, 136-138 (2016).

15 Adams, J., Smothers, J., Srinvasan, R. & Hoos, A. Big oppertunities for small molecules in immuno-oncology *Nature Reviews Drug Discovery* **14**, 603-622 (2015).

16 Bachovchin, D. & Cravatt, B. The pharmacological landscape and therapeutic potential of serine hydrolases *Nature Reviews Drug Discovery* **11**, 52-68 (2012).

17 Coussens, L., Fingleton, B. & Matrisian, L. Matrix metalloproteinase inhibitors and cancer: trials and tribulations *Science* **295**, 2387-2392 (2002).

18 Duffy, M. *et al.* The ADAM family of proteases: new biomarkers and therapeutic targets for cancer? *Clinical Proteomics* **8** (2011).

19 Dufour, A. & Overall, C. Misisng the target: matrix metalloproteinase antitargets in inflammation and cancer *Trends in Pharmacological Sciences* **34**, 233-242 (2013).

20 Egebald, M. & Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nature reviews Cancer* **2**, 161-174 (2002).

21 Erler, J. *et al.* Lysyl oxidase is essential for hypoxia-induced metastasis *Nature* **440**, 1222-1226 (2006).

22 Larrinaga, G. *et al.* Dipeptidyl-Peptidase IV activity Is correlated with colorectal cancer prognosis. *PLOS One* **10** (2015).

23 Gupta, G. & Massague, J. Cancer metastasis: building a framework *Cell* **127**, 679-695 (2006).

24 Matrislan, L. & Lynch, C. Matrix metalloproteinases in tumor-host cell communication. *Differentiation* **70**, 561-573 (2002).

25 Menendez, J. & Lupu, R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis *Nature Reviews* **7**, 763-777 (2007).
26 Kassenbrock, K., Plaks, V. & Werbb, Z. Matrix metalloproteinases: regulators of the tumor microenvieornment *Cell* **141**, 52-67 (2010).
27 Bonnas, C., Chaou, J. & Werb, Z. Remodeling the extracellular matrix in development and disease *Nature Reviews in Molecular and Cellular Biology* **15**, 786-801 (2014).
28 Liu, R., Li, H., Liu, L., Yu, J. & Ren, X. Fibroblast activation protein: A potential theraputic target in cancer. *Cancer Biology and Therapy* **13**, 123-129 (2012).
29 Nomura, D. *et al.* Monoacyglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* **140**, 49-61 (2010).
30 Rozanov, D. *et al.* Non-proteolytic receptor/ligant interacitons associate cellualr membrane type-1 matrix metalloproteinase with the complement component c1q *The Journal of Biological Chemistry* **279**, 50321-50328 (2004).
31 D'Alessio, S. *et al.* Tissue inhibitor of metalloproteinases-2 binding to membrane-type 1 matrix metalloproteinase induces MAPK activation and cell growth by non-proteolytic mechanism *Journal of Biological Chemistry* **283**, 87-92 (2008).
32 Dove, A. MMP inhibitors: glimmers of hope amidst clinical failures *Nature* **8**, 95 (2002).
33 OVerall, C. & Lopez-Otin, C. Strategies for MMp inhibition in cancer: innovations for the post-trial era *Nature Reviews Cancer* **2**, 657-672 (2002).
34 King, J., Clingan, P. & Morris, D. Randomised double blind placebo control study of adjuvant treatment with the metalloproteinase inhibitor, Marimastat in patients with inoperable colorectal hepatic metastases: significant survival advantage in patients with musculoskeletal side-effects. *Anticancer research* **23**, 639-645 (2003).
35 Ling, B. *et al.* A novel immunotherapy targeting MMP-14 limits hypoxia, immune suppression and metastasis in tripple-negative breast cancer models. *Oncotarget* **8**, 58372-58385 (2017).
36 Devy, L. *et al.* Selective inhibition of matrix metalloprotienase-14 blocks tumor growth, invasion, and angiogenesis *Canser Research* **69**, 1517-1526 (1009).
37 Barry-Hamilton, V. *et al.* Allosteric inhibiton of lysyl oxidase-like-2 impedes the development of a pathologic microenviornment *Nature Medicine* **16**, 1009-1017 (2010).
38 Nicholson, S., Wood, C. & Devy, L. Use of MMP-9 and MMP-12 binding protiens for the treatment and prevention opf systemic sclerosis. United States of America patent (2010).
39 Appleby, T. *et al.* Biochemical characterization and structure determination of a potent, selective antibody inhibitor of human MMP9. *Journal of Biological Chemistry* **292**, 6810-6820 (Greenstein, AE).
40 Nam, D. & Ge, X. Development of a periplasmic FRET screening method for protease inhibitory antibodies. *Biotechnol Prog.* **110**, 2856-2864 (2013).
41 Li, J. *et al.* Molecular imprint of enzyme active site by camel nanobodies *The Journal of Biological Chemistry* **287**, 13713-13721 (2012).
42 Nam, D., Fang, K., Rodriguez, C., Lopez, T. & Ge, X. Generation of inhibitory monoclonal antibodies targeting matrix metalloproteinase-14 by motif grafting and CDR optimization *Protein Enigneering Design and Selection* **30**, 113-118 (2017).
43 Devy, L. & Dransfield, D. New strategies for the next generation of matrix-metalloproteinase inhibitors: selectively targeting membrane-anchored MMPs with therapeutic antibodies. *Biochemistry Research International* (2011).
44 Nam, D., Rodriguez, C., Remacle, A., Strongin, A. & Ge, X. Active-site MMP-selective antibody inhibitors discovered from convex paratope syntehtic libraries *PNAS* **113**, 14970-14975 (2016).
45 Arabi-Ghahroudi, M. Cameldi single-domain antibodies: historical perspective and future outlook. *Fronteirs in Immunology* **8** (2017).
46 Nagase, H., Visse, R. & Murphy, G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular Research* **69**, 562-573 (2006).

47 Sharbi, O. *et al.* Affinity and specificity enhancing mutations are frequent in multispecific interactions between TIMP2 and MMPs *PLOS One* **9** (2014).
48 Arkadash, V. *et al.* Development of high affinity and high specificity inhibitors of matrix metalloproteinase 14 through computaitonal design and directed evolution *The Journal of Biological Chemistry* **292**, 3481-3495 (2017).
49 Gossage, D. *et al.* Phase 1b Study of the Safety, Pharmacokinetics, and Disease-related Outcomes of the Matrix Metalloproteinase-9 Inhibitor Andexcaliximab in Patients With Rheumatoid Arthritis. *Therapeutics* **40**, 156-165 (2018).
50 Fleishman, S. *et al.* Computational design of protiens targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816-821 (2011).
51 Silwoski, G., Kothiwale, S., Meiler, J. & Lowe, E. Computational Methods in Drug Discovery. *Pharmacological Reviews* **66**, 334-395 (2014).
52 A, C. & al., e. Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74-79 (2017).
53 Huang, P., Boyken, S. & Baker, D. The coming of age of de novo protein design *Nature* **537**, 320-327 (2016).
54 Brekke, O. & Sadlie, I. Therapeutic antibodies for human diseases at the dawn of the twenty first century *Nature Reviews Drug Discovery* **2**, 52-56 (2002).
55 Adams, J. & Sidhu, S. Synthetic antibody technologies *Current Opinion in Structural Biology* **24**, 1-9 (2014).
56 Sidhu, S. & Fellouse, F. Synthetic therapeutic antibodies *Nature Chemical Biology* **2**, 682-688 (2006).
57 Liposvek, D., Mena, M., Lippow, S., Basu, S. & Baynes, B. in *Protein Engineering and Design* (eds SJ Park & JR Cochran) Ch. 4, 83-102 (CRC Press, 2010).
58 Mahon, C. *et al.* Comprehensive interrogation of a minimalist synthetic CDR-H3 library and its ability to generate antibodies with therapeutic potential. *Journal of Molecular Biology* **425**, 1712-1730 (2013).
59 Vaughan, T. *et al.* Human antibodies with sub-nanomolar affinity isolated from a large non-immunized phage display library *Nature Biotechnology* **14**, 309-314 (1995).
60 Feldhaus, M. *et al.* Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces Cerevisiae surface display library *Nature Biotechnology* **21**, 163-170 (2003).
61 Willuda, J. *et al.* High thermal stability is essential for tumor targeting antibody fragments *Cancer Research* **59** (1999).
62 Ewert, S., Huber, T., Honegger, A. & Pluckthun, A. Biophysical properties of human antibody varriable domains *Journal of Molecular Biology* **325**, 531-553 (2003).
63 Worn, A. & Pluckthun, A. Identification, classification and improvement by protein engineering *Biochemistry* **38**, 8739-8750 (1999).
64 Persson, H. *et al.* CDR-H3 diversity is not required for antigen recognition by synthetic antibodies *Journal of Molecular Biology* **435**, 803-811 (2012).
65 Kehoe, H. *Targeting the tumor microenvironment with protein-small molecule hybrids* Masters of Science in Bioengineering thesis, Tufts University (2017).
66 Miersch, S., Maruthachalam, B., Geyer, R. & Sidhu, S. Structure-directed ant tailored diversity synthetic antibdoy libraries yield novel anti-EGFR antagonists *ACS chemical biology* **12**, 1381-1389 (2017).
67 Koerber, J., Thomsen, N., Hannigan, B., Degrado, W. & Wells, J. Nature-inspired design of motif-specific antibody scaffolds *Narure Biotechnolgy* **31**, 916-921 (2013).
68 Bowley, D., Labrijin, A., Zwick, M. & Burton, D. Antigen selection from an HIV-1 immune antibody library displayed on yeast yeilds many novel antibodies compared to selection from the same library displayed on phage *Protein Engineering Design and Selection* **20**, 81-90 (2007).
69 Bradbury, A., Sidhu, S., Dubel, S. & McCafferty, J. Beyond natural antibodies: the power of in vitro display technologies *Nature Biotechnology* **19**, 245-254 (2011).

70 McCafferty, J., Griffiths, A., Winter, G. & Chiswell, D. Phage antibodies: filamentous phage displaying antibody varriable domains. *Nature* **348**, 552-554 (1990).
71 Hanes, J. & Pluckthun, A. In vitro selection and evolution of functional proteins by using ribosome display *PNAS* **94**, 4937-4942 (1997).
72 Boder, E. & Wittrup, K. Yeast surface display for screening combinatorial polypeptide libraries *Nature Biotechnology* **15**, 553-557 (1997).
73 Ho, M. & Pastan, I. in *Therapeutic Antibodies Methods and Protocols q* Vol. 525 *Methods in Molecular Biology* (ed Antony S. Dimitrov) Ch. 18, 337-352 (Springer 2009).
74 Angelini, A. *et al.* in *Yeast Surface Display* Vol. 1319 *Methods in Molecular Biology* 3-36 (2015).
75 Ackerman, M. *et al.* Highly avid magnetic bead capture: an efficient seleciton method for de novo protein engineering utilizing yeast surface dispalay *Biotechnol Prog.* **25**, 774-783 (2009).
76 Van Deventer, J., Le, D., Zhao, J., Kehoe, H. & Kelly, R. A platform for constructing, evaluating and screening bioconjugates on the yeast surface *Protein Engineering Design and Selection* **29**, 485-493 (2016).
77 Van Deventer, J., Kelly, R., Rajan, S., Wittrup, K. & Sidhu, S. A switchable yeast dispaly/secretion system *Protein Engineering Design and Selection* **28**, 317-325 (2015).
78 Zorniak, M. *et al.* Yeast display biopanning identifies human antibodies targeting glioblastoma stem-like cells *Scientific Reports* **7** (2017).
79 Aziuteu, M. *et al.* Computation-Guided Backbone Grafting of a Discontinous Moteif onto a Protein Scaffold *Science* **334**, 373-376 (2011).
80 Moroncini, G. *et al.* Motof-grafted antibodies containing the replicative interface of cellular PrP are specific for PrPSC *PNSA* **101**, 10404-10409 (2004).
81 Horiya, S., Bailey, J., Temme, S., Guillen, Y. & Krauss, I. Directed evolution of multivalent glycopeptides tightly recognized by HIV antibody 2G12. *Journal of the Amrican Chemical Society* **136**, 5407-5415 (2015).
82 Yang, A. *et al.* A chemical biology route to site-specific authentic protein modifiecations *Science* **354**, 623-625 (2016).
83 JY, A. *et al.* Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *PNAS* **109**, 16101-16106 (2012).
84 Flentke, G. *et al.* Inhibition of dipeptidyl aminopeptidase IV (DP-IV) by Xaa-boroPro dipeptides and use of these inhibitors to examine the role of DP-IV in T-cell function. *PNAS* **88**, 1556-1559 (1991).
85 Meyer, S., Shomin, C., Gaj, T. & Gosh, I. Teathering small molecules to a phage dispaly library: discovery of a selective bivalent inhibitor of protien kinase A *Journal of the Amrican Chemical Society* **129**, 13812-13813 (2007).
86 Holliger, P. & Hudson, P. Engineered antibody fragments and the rise of single domains *Nature Biotechnology* **23**, 1126-1136 (2005).
87 Tonegawa, S. Somatic generation of antibody diversity *Nature* **302**, 1079-1087 (1983).
88 Xu, J. & Davis, M. Diversity in the CDR3 region of VH is sufficient for most antibody specificities *Immunity* **13**, 37-45 (200).
89 Kunik, V. & Ofran, Y. The indistinguishability of epitopes from protein surface is explained by the distinct binding preferences of each of the six antigen-binding loops. *Protein Engineering Design and Selection* **26**, 599-609 (2013).
90 Kwong, P. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-659 (1998).
91 Huh, J., White, A., Brych, S., Franey, H. & Matsumura, M. The identification of free cystine residues within antibodies and a potential role for free cystine residues in covalent aggregation because of agitation stress *Journal of Pharmocological Sciences* **102** (2013).
92 Kelly, R., Le, D., Zhao, J. & Wittrup, K. Reduction of nonspecificity motifsi n synthetic antibody libraries *Journal of Molecular Biology* **430**, 119-130 (2018).

93 Fellouse, F. L., B, Compaan, D., Peden, A., Hymowitz, S. & Sidhu, S. Molecular recognition by a binary code *Journal of Molecular Biology* **348**, 1153-1162 (2005).

Chapter 2: Methods

2.1: Yeast Culture[79]

The *Saccharomyces cerevisiae* strain RJY100 was used for all experiments. For induction of a polyclonal population or naïve library, 10-fold the clonal diversity was first inoculated in SD-CAA media in a volume large enough to keep culture below OD_{600} 1. For single clone populations, 200uL of saturated stock culture was inoculated. Cultures were grown at 30ºC with shaking at 300RPM to saturation. Following saturation, cultures were diluted to OD_{600} 1 in SD-CAA and allowed to double to OD_{600} 2 while incubating at 30ºC and 300RPM. Following doubling, cells were spun down at 2400xg for 5 minutes to remove SD-CAA and re-suspended to OD_{600} 1 in SG-CAA. Cells were then incubated at 20ºC and 300RPM for at least 16 hours prior to flow cytometry or library sorting. All cultures are supplemented with 1x pen/strep (Sigma Aldrich).

2.2: Antigen Biotinylation

Antigen Biotinylation: Full IgG purified from non-immunized animal serum (Jackson Immuno-Research) was diluted from stock concentration to 1mg/mL in ice cold 1xPBS. Lyophilized hen egg-white lysozyme, bovine pancreases RNAase A (Sigma Aldrich) and horseradish peroxidase (Pierce Scientific) were prepared in 1mg/mL solutions in ice cold PBS. 100mM EZ-Link NHS-LC biotin solution (Pierce Scientific) was prepared in DMF (Sigma Aldrich) and mixed vigorously to dissolve by pipetting. For all protein, reactions were carried out at room temperature for 30 minutes with 100µL of sample[98]. For IgGs, 3-5-fold molar excess of biotin solution was added. For enzymes, 2-fold molar excess of biotin solution was used. Reactions were quenched with 10µL of 0.5M TRIS, 0.02%NaN_3 pH 7.4. Protein was desalted in two sequential Zeba Spin Desalting Columns with 7K MW cutoff (Thermo Fischer), which were prepared and used according to the manufacturers specifications[99]. The concentration of the flow through was measured using Nano Drop One (Thermo Fischer).

Validation Assay: The extant of biotinylation was estimated using the HABA/Avidin kit (Peirce Scientific) according to the manufacturers instructions[100]. The well plate was blanked sequentially by shaking at 150RPM for 1 minute and reading absorbance values at 500nm until readings were within ±0.05AU. This was repeated for reading absorbance with protein sample. Extent of biotinylation was then calculated using the HABA/Avidin kit online tool[101].

2.3:Library Screening[79]

Naïve Library: For naïve library sorts, 10^{10} cells from the 10^9 membered CDRH3 library that had previously been depleted for murine FC, biotin and streptavidin were inoculated in 1L SD-CAA and induced as described above. $1.5x10^{10}$ cells per antigen sort track were spun down at 2400xg for 5 minutes to remove SG-CAA media. Cells were washed 3x in ice-cold PBSA and re-suspended to a final volume of 15mL. 52.5µL of Biotin Binder Dynabeads (Thermo Fischer) were diluted to 1mL in ice cold 1xPBSA and washed 3x with ice cold PBSA on a Dynamag-2 magnet (Life Technologies). 174 pico-moles of biotinylated antigen was added to the beads and beads were incubated at 4ºC with rotation for 2 hours. Following incubation, the beads were again diluted and washed 3x in ice cold PBSA on the magnet before being re-suspended back to their original volume. 50µL of beads were added to cells and 2.5µL were saved for flow cytometry validation. Cells were incubated with beads for 2 hours at 4ºC for 2 hours. Following the incubation, the 15mL of cells were divided to 1.5mL tubes and incubated on the Dynamag for 5 minutes. Supernatant was aspirated off and discarded. Retained beads were washed with ice cold

PBSA 1x for 2 minutes on the magnet. Beads were rescued by re-suspending in 100mL SD-CAA supplemented with pen-strep and allowed to grow up to saturation at 30ºC with shaking at 300RPM.

Further Library Enrichments: The rescue was induced according as previously mentioned. $1.5x10^9$ cells were prepared the same way as the naïve library for sorting but were re-suspended to a final volume of 1mL. Beads were also prepared according to the same general methodology as the naïve library sort. In further enrichments, antigen beads were prepared by adding 42 pico-moles of biotinylated antigen to 12µL of washed beads. Biotin coated beads were prepared by adding 1µL of 50nM D Biotin (Sigma Aldrich) solution in DMSO (Sigma Aldrich) to 10µL of washed beads. In further enrichments, cells are first depleted for streptavidin and biotin followed by a second streptavidin depletion before a positive sort. For depletions, cells were incubated with 10µL of the appropriate beads for 1.5 hours at 4ºC with rotation. Supernatant was recovered and carried over to the next depletion or positive sort and beads were discarded. The positive sort and rescue was run the same as in the naïve library screen, only this time beads were washed 2x before rescue. Further library enrichments were repeated a total of 2x before progressing to flow cytometry sorting.

2.4: Flow Cytometry

FACS Sorting: $2.0x10^7$ induced cells were diluted to a volume of 1mL PBSA and washed 3x in 1mL PBSA in 1.5mL tubes by spinning down at 12,000RPM for 30 seconds, aspirating off supernatant, and re-suspending in 1xPBSA. After the final wash, cells were re-suspended in 50µL of 50nM or 250nM biotinylated antigen and 1/500 dilution of chicken anti-Cmyc antibody (Gallus Immunotech) in PBSA. The PE single color control was a Donkey IgG isolated single clone and was labeled with only 50nM biotinylated donkey IgG. The Alexa Flour 647 single color control was labeled with only a 1/500 dilution of Cmyc. Cells were incubated at room temperature with rotation for 30 minutes. After the incubation, cells were diluted with 1mL ice-cold PBSA and kept on ice. Cells were washed 3x at 4ºC and after the final wash, re-suspended in 50µL of 1/250 dilution of anti-biotin PE (eBioscience) and goat anti chicken Alex Flour 647 (Life Technologies) and allowed to incubate at 4ºC with rotation for 15 minutes. The PE single color control was only labeled with the anti-biotin PE and the Alexa Flour 647 control was labeled only with goat anti-chicken 647. After the incubation, cells were diluted with 1mL ice-cold PBSA and washed 2x at 4ºC. Cells were left in the dark and on ice until sorting. Sorting was performed by the Tufts University School of Medicine Flow Cytometry Core using a BD Aria cell sorter. An upper and lower gate was drawn for each sample (see appendix) and a minimum of 10,000 events per gate were collected. Cells were collected in 5mL SD-CAA with pen-strep and allowed to grow up to saturation.

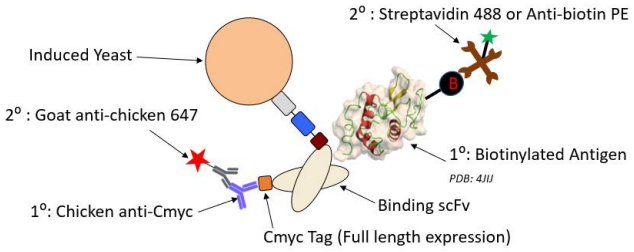

Figure 2.1. Flow cytometry labeling of induced RJY100. Yeast are first labeled with a primary label mixture consisting of biotinylated antigen and an anti-Cmyc antibody. Goat anti-chicken 647 then labels the anti-Cmyc antibody and streptavidin 488 or anti-biotin PE labels the biotin on the antigen. The secondary labels are what are read by the cytometer.

Binding Experiments: Labeling for binding experiments was performed the same as for FACS sorting. $2.0x10^6$ cells were used. In some experiments, streptavidin Alexa Flour 488 (Life Technologies) was used instead of anti-biotin PE. In these cases, a single-color Alexa Flour 488 control was prepared in the same

way as the 647 single color control, using Goat anti-chicken 488 as the secondary label. Samples were read on an Attune NxT Cytometer, and 10,000 events were collected per sample.

Bead validation experiments: 2.5µL of antigen-coated and washed beads were re-suspended in 50µL of a 1/250 dilution of either streptavidin 488 or anti-biotin PE and incubated on the rotator wheel at 4ºC for 15 minutes. Further washing and data collection was carried out as described for binding experiments.

Biotinylated Antigen Competition Experiments: $2.0x10^6$ induced cells per sample were placed in the appropriate wells of a 96 well plate. Cells were diluted to 200µL 1xPBSA and spun down at 2400xg for 5 minutes. The plate was shaken out to decant supernatant. Cells were re-suspended in 200µL PBSA and washed a total of 3x. Competition cells were re-suspended in 50uL PBSA with 1 or 5µM non-biotinylated antigen and 1/500 chicken anti-Cmyc. Secondary label control and biotinylated antigen only conditions were re-suspended in 50µL PBSA with 1/500 chicken anti-Cmyc. Cells were incubated for 30 minutes at room temperature on a plate shaker at 150RPM. After the incubation, the plate was kept on ice. Into the biotinylated antigen and competition condition, enough biotinylated antigen for a final concentration of 50 or 250mM was added. Cells were mixed and incubated on a plate shaker at 150RPM for 15 minutes at 4ºC. After the incubation, cells were diluted to 200µL ice cold PBSA and washed 3x at 4ºC. Cells were re-suspended in 50µL PBSA with 1/250 dilution of streptavidin 488 and goat anti-chicken 647 and incubated for 15 minutes on a plate shaker at 150RPM for 15 minutes at 4ºC. After 15 minutes, cells were washed 2x in PBSA and left on ice in the dark until reading on the flow cytometer.

Titrations on the Yeast Surface[79]: $1.5x10^6$ induced cells were diluted to 1mL 1xPBSA in a 1.5mL tube and washed 3x at 12000RPM for 30 seconds. After the final wash, cells were re-suspended to 1mL PBSA. 300µL of those cells were aliquoted and 0.3µL of chicken anti-Cmyc was added. The tube was vortexed gently to mix. Triplicate titration plate was prepared by making 53.3µL of 1.25µM of biotinylated antigen. 8 4-fold dilutions were performed across triplicate conditions in the plate and one PBSA blank condition. 10uL of Cmyc cells were added to each well. The plate was incubated at 150RPM at room temperature overnight. The next day, the experiment was continued with secondary labeling as described for the competition experiment. 1/500 dilutions of secondary label were used instead of the standard 1/250 dilution. 3,000 events

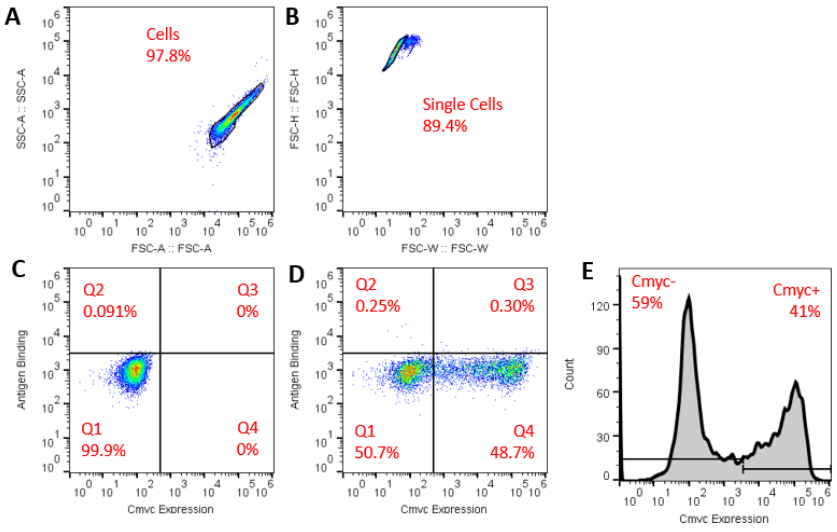

Figure 2.2. Gating analysis for flow cytometry using FlowJo software. A) Debris are gated out using SSCA vs FSCA plot. B) Doublets and triplets are gated out using FSCH vs FSCW plot. C) Vertical quadrant line is positioned such that all negative cells fall into Q1. D) Horizontal quadrant line is positioned such that no cells in the AF 647 single color control are in Q3. E) Histogram gating used to gate out un-induced cells, Cmyc-, from induced cells, Cmyc+.

were recorded per sample on the cytometer instead of the usual 10,000 events.

Data Analysis: All analysis was carried out using FlowJo software. Figure 2.2 shows flow cytometry gating. Debris were first gated out from the yeast population using the negative control sample on an SSCA vs FSCA plot. Doublets and triplets were then gated out from single cells using FSCH vs FSCW plot. In cases where scatter plots with quadrants was used, the vertical line of the quadrants was positioned by using the negative control sample in such a way that all negative cells fall in Q1. The horizontal line was positioned using the AF 647 single color control sample in such a way that minimal cells are in Q3.

In cases where histograms were used, single cells were first gated out for analysis. Induced vs un-induced cells were separated using Cmyc signal. The induced population was then placed on a histogram corresponding to antigen binding level. For titration experiments, the median florescence intensity (MFI) for the antigen binding signal was recorded for both the Cmyc positive and negative populations of a given sample. The MFI of the positive sample was subtracted from that of the negative sample. The ratio of the MFI for the sample with the highest antigen concentration to that of all subsequent samples was taken to get normalized binding. This data was then input into GraphPad Prism 7 software. A curve was fit to the data and the K_D was estimated including standard error and 95% confidence interval using the binding-saturation, one site specific binding model.

2.5: Isolation of Single Clones and Sequencing

Isolation of single clones: Following FACS sorting, the recovery yeast population was mini-prepped according to the manufacturer's instructions (Zymoprep)[102]. Instead of using 2mL of fresh culture at OD_{700} of 2, 5mL of culture was used. Instead of using the included mini-prep columns, columns from a separate E Coli mini prep kit were used (Epoch Life Sciences). Wash buffer from the E Coli mini prep kit was also used. The column was stood for 5 minutes to elute DNA instead of the recommended 2 minutes and sterile water was used instead of the provided elution buffer. The concentration and purity of plasmid DNA was measured on the nanodrop one.

Sequencing of single clones: The plasmid DNA from the yeast mini prep was transformed into chemically competent ZH1α E coli. Plasmid DNA was mixed with cells in a ratio of 1:10. Cells were incubated on ice for 30 minutes, followed by a 45 second heat shock at 42ºC and a 5-minute incubation on ice. Cells were recovered in 500µL SOC media for 1 hour at 37ºC with shaking at 200RPM, before plating on LB AMP plates and growing overnight at 37ºC. Depending on the estimated clonal diversity, anywhere from 10-25 single colonies were picked to submit for sequencing. Sanger sequencing was performed by Quintara Biosciences. Sequencing was performed using both forward and reverse primers specific to the PCT-CON 2 yeast display system plasmid.

2.5: References

1 Angelini, A. *et al.* in *Methods in Molecular Biology* Vol. 1319 3-36 (2015).

2 *Pierce Biotechnology Instructions: EZ-Link NHS-biotin reagents*, <https://www.thermofisher.com/order/catalog/product/21336> (2018).

3 *Thermo Fisher Scientific Zeba Desalting Products* <https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-purification-isolation/protein-dialysis-desalting-concentration/zeba-desalting-products.html> (2018).

4 *Pierce Biotechnology Instructions: Pierce Biotin Quantification Kit*, <https://www.thermofisher.com/order/catalog/product/28005> (2018).

5 *Thermo Fisher Scientific HABA Calculator* <https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-labeling-crosslinking/protein-labeling/biotinylation/biotin-quantitation-kits/haba-calculator.html> (2018).

6 *Zymo Research Zymoprep Yeast Plasmid Miniprep* <https://www.zymoresearch.com/zymoprep> (2018).

Chapter 3: Diagnosing Over-Biotinylation and Its Implication in Isolation of Non-Specific Binding Clones

3.1: Introduction

Biotinylating target antigens is important for both antibody library screening and flow cytometry assays. In library screening, streptavidin-coated magnetic beads are coated with biotinylated antigen to isolate binding yeast from the rest of the population[9]. In flow cytometry binding experiments, the biotin on the antigen can be labeled with a fluorescent streptavidin conjugate. This in turn allows for the detection of antigen binding on the yeast surface[79, 103].

There are two main methods to biotinylate an antigen of interest: enzymatic and chemical. In enzymatic biotinylation, a specific tag sequence is cloned into the expression vector for the antigen of interest at the location desired for biotin incorporation. An enzyme then incorporates exactly one biotin to the tag. This method is advantageous because of its robustness. Only a single biotin is incorporated per protein molecule at the same position[104]. However, this method only works with the proper tag cloned into the antigen expression vector. It thus cannot be used with purchased antigen or antigen from collaborators[103]. Moreover, uniform positioning of the biotin could mask critical antigen epitopes during library screening since the same antigen surface is positioned against the magnetic beads. In contrast, in chemical biotinylation (Figure 3.1), an NHS cross linker conjugated to the biotin reacts

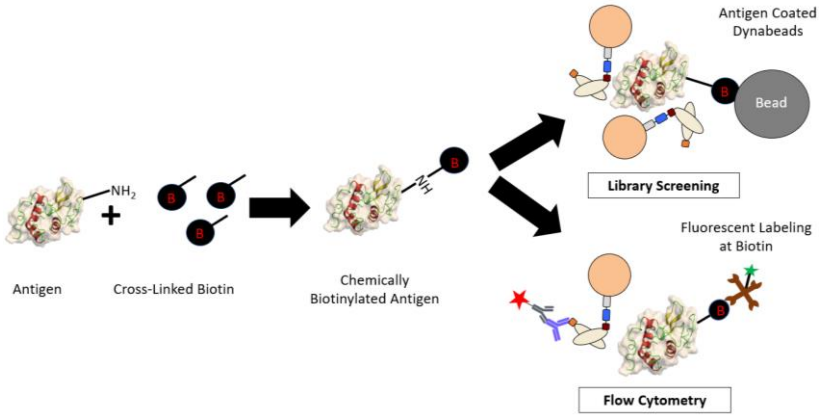

Figure 3.1. Chemical biotinylation of target antigen and its applications. In chemical biotinylation, a linker on the biotin reacts with free amines from lysine residues or the N-terminus of the antigen. The result is antigen with biotins at variable positions. This antigen can then either be used to coat streptavidin beads for library screening or for flow cytometry experiments where the biotin on the antigen is labeled with a fluorophore. The antigen is represented by the white globular protein. The biotin is represented by the black circle with a red "B" and the NHS-LC cross linker is represented by the black line. The biotin is not drawn to scale and is approximately 1/1000 the size of the antigen itself.

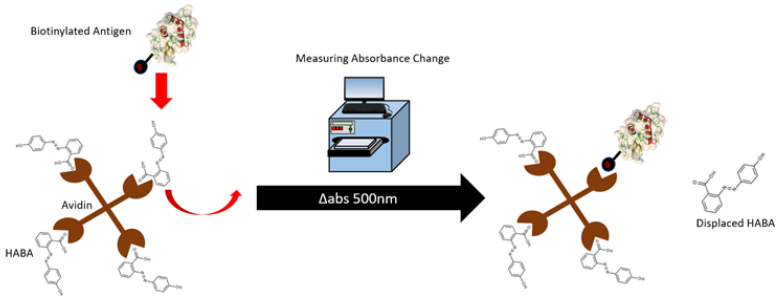

Figure 3.2. Absorbance based biotin quantification assay. An absorbance assay can be used to quantify the extent of a chemical biotinylation reaction. The main reagent of the assay is a HABA/avidin complex. When mixed with biotinylated antigen, the biotin displaces the HABA in the avidin, resulting in a change in absorbance at 500nm. This absorbance change can be used to calculate the ratio of moles biotin to moles antigen.

with free amines in the antigen. The incorporation of biotins is random and in principle occurs with equal probability at any free amine residue in the antigen[98, 105]. This is advantageous because a random distribution of biotin reduces the risk that a critical antigen epitope will be fully masked during library screening. Moreover, this method does not require any cloning and thus can be used on purchased antigen or antigen from collaborators. The disadvantage of this method is it is difficult to control the extent of reaction to get the desired number of biotins per protein molecule[103]. An uncontrolled reaction can lead to the isolation of non-specific clones that bind to biotin instead of the antigen of interest.

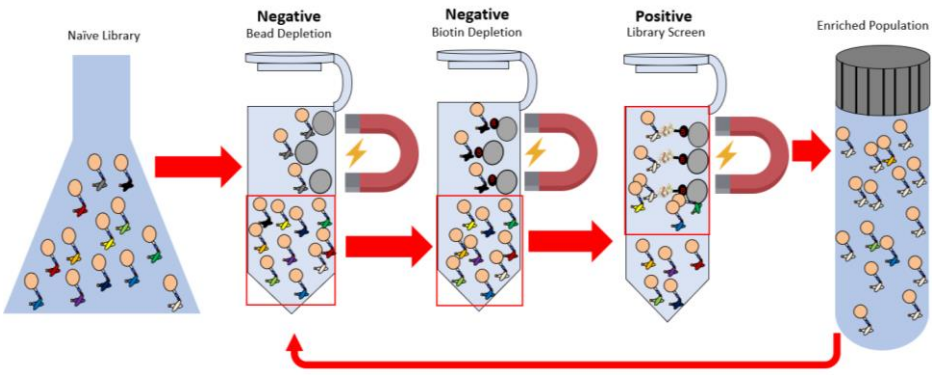

Figure 3.3. Biotin depletions to prevent isolation of non-specific binders. To prevent the isolation of clones binding to biotin instead of to the target antigen, negative selections are performed before each positive screen for both bare beads and beads coated in biotin. White yeast represent antigen binding clones, grey yeast represent bead non-specific binders and black yeast represent biotin non-specific binders. Red boxes show the population from each screening step that are carried over to the next step. Sorting is repeated until the desired enrichment is achieved in the rescue population.

In this thesis work, we prepare antigen using chemical biotinylation because we wanted to be able to screen against both antigen produced in-house and purchased antigen using consistent methodology. We use two main safeguards to mitigate the main disadvantages of this approach. Firstly, the extent of the biotinylation reaction is verified using an absorbance assay, as shown in Figure 3.2. The main reagent for this assay is Avidin complexed with 4'-hydroxyazobenzene-2-carboxylic acid (HABA), a phenyl-benzoic-acid. When the biotinylated antigen is added to the reagent, the biotin displaces the HABA because of its high affinity for the avidin, resulting in a change in absorbance. The change in absorbance can then be used to calculate the number of biotins per protein molecule[100]. The target extent of reaction is 1-2 biotins per protein molecule. This ensures that the protein is not over-biotiynlated to the point where there is a risk of isolating non-specific binders. The second safeguard is to run negative selections using biotin coated beads during library screening as shown in Figure 3.3[79]. This process removes any biotin binding clones that may be present in the population.

3.2: Over-biotinylation of Sort Antigen Masks Critical Epitopes and Leads to the Isolation of Non-specific Binders

After constructing the CDRH3 library, validation began by screening against model antigens including: bovine, donkey and rabbit IgGs. Screening an antibody library is not 100% efficient, and non-binding clones will always be retained in the enriched population. These clones most likely clumped to a bead surface or to binding yeast cells after washing and were thus carried over to the next round of sorting. Therefore, the first step after isolating single clones from a library screen is to determine which clones are genuine antigen binding and which are non-binding. Figure 3.4 shows the results of this flow cytometry binding assay for individual bovine, donkey and rabbit IgG binding clones isolated from the

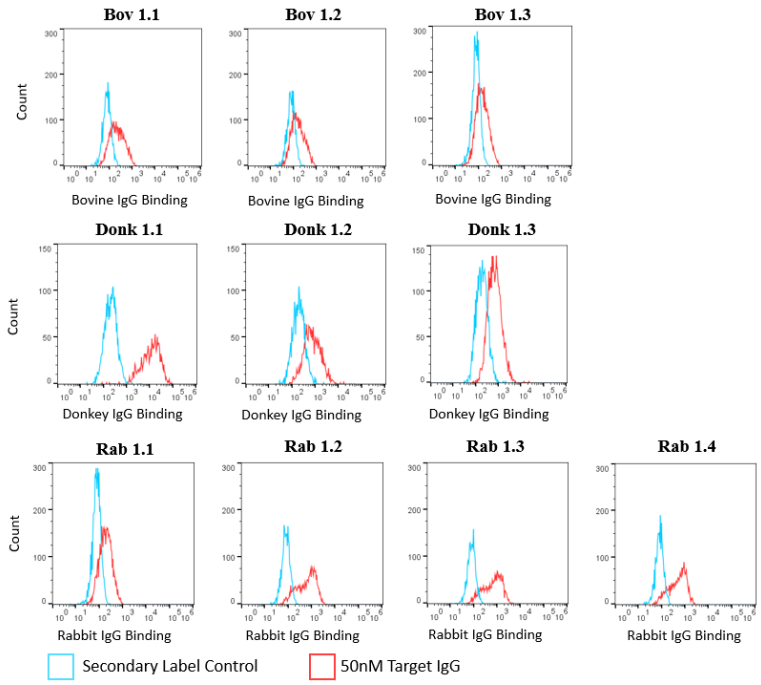

Figure 3.4. Initial binding assay for isolated IgG binding clones. All cells were labeled for Cmyc expression, which was used to gate out un-induced cells. In the antigen binding condition shown in red, cells were labeled with 50nM target antigen and anti-biotin PE secondary. In the secondary control condition shown in blue, cells were only labeled with the PE secondary. X-axis represents relative PE fluorescent signal and Y-axis represents number of events. 10,000 events were collected for each sample. Data collected by Haixing Kehoe[7].

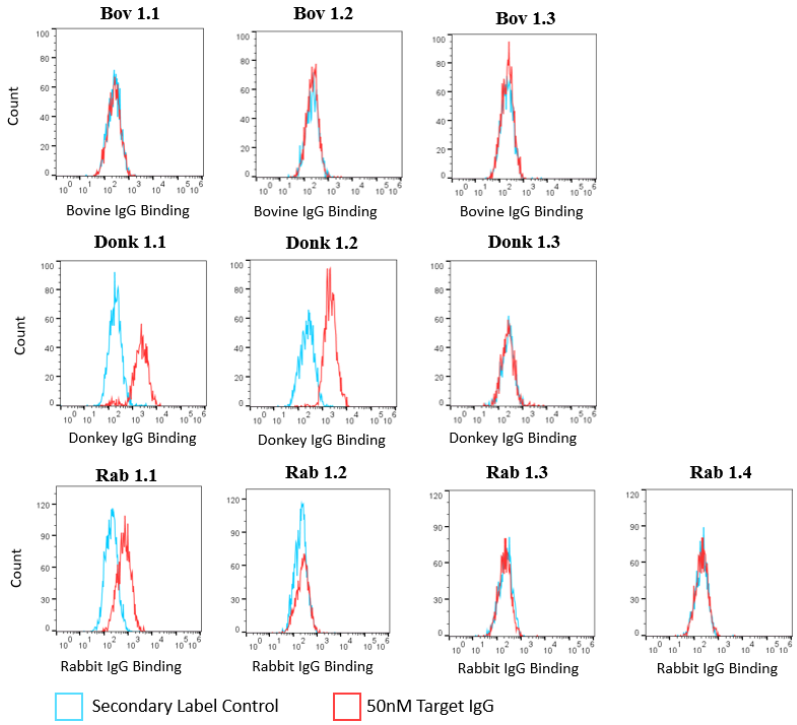

Figure 3.5. Follow-up binding assay for isolated IgG binding clones using freshly biotinylated antigen. All cells were labeled for Cmyc expression, which was used to gate out un-induced cells. In the antigen binding condition shown in red, cells were labeled with 50nM target antigen and a streptavidin-Alexa Flour 488 conjugate secondary label. In the secondary control condition shown in blue, cells were only labeled with the streptavidin. X-axis represents relative streptavidin 488 fluorescent signal and Y-axis represents number of events. 10,000 events were collected for each sample.

initial set of CDRH3 library screens[7]. Figure 3.5 shows a follow up binding assay in which cells were labeled with freshly biotinylated IgG. Clones determined to be non-binding because of carryover are not shown.

Histograms show relative antigen binding for isolated single clones labeled with target antigen and a secondary fluorophore, as compared to control samples treated with the fluorophore alone. Only induced cells are shown as events on the histograms. All single clones labeled with biotinylated antigen in

Figure 3.4 show a shift to the right along the x-axis as compared to the overlaid control sample. This shift represents an increase in the antigen binding fluorescent signal, which comes from the fluorophore that selectively binds to the biotin on the antigen surface. Thus, all clones in the initial binding assay exhibit apparent binding to their target antigen.

In contrast, in the repeat binding assay as shown in Figure 3.5, most of the clones no longer showed a shift along the x-axis. The histogram for cells labeled with antigen and fluorophore is superimposed over that for cells labeled with fluorophore only. The fact that the fluorescent signal corresponding to antigen binding remains constant between the control and antigen labeled cells indicates that clones are no longer binding to their target antigen. The exceptions are Donkey 1.1, Donkey 1.2 and Rabbit 1.1, all of which all show similar increases in antigen binding signal as in the initial binding experiment. When taken together, these results show that most isolated single clones that had previously bound to their target antigen ceased to bind to that antigen when it was freshly biotinylated. Only one third of isolated clones retained binding to the target antigen.

In analyzing the histograms of Figures 3.4 and 3.5, it is important to note that the control cells do not have a median antigen binding fluorescence signal of zero and have a basal fluorescence. This is both because cells have an autofluorescence at the wavelength of light the cytometer reads for the antigen binding signal, and because of slight non-specific sticking of the antigen secondary label to the yeast surface. Also important to note is that the histogram for the antigen-treated cells is more widely distributed than that for the control cells. This is because there is not a uniform level of construct expression on the yeast surface. Cells with more displayed constructs will show a greater level of antigen binding signal than those with less displayed constructs, causing variability.

We next sought to determine what the underlying cause of the observed diminished binding was. Antigen degradation could cause a change in the conformational shape of the protein, eliminating critical epitopes that clones used to bind to. In principle, clones that still showed binding in the follow up experiment could have bound to more thermodynamically stable regions of the antigen that were less susceptible to degradation[106]. However, this was unlikely the case since IgGs are usually very thermodynamically stable[68].

An alternative explanation could be that poorly controlled biotinylation reactions were initially used to prepare the antigen used in library screening. Over-biotinylation of sort antigen could result in the isolation of biotin binding clones rather than genuine antigen binding clones, as shown in Figure 3.6. Library depletions against biotin are run consistently during the screening process to prevent this. However, the depletions are run against D biotin and not cross-linked NHS-LC biotin. Therefore, it is possible that certain clones may bind to the 6-carbon cross linker that conjugates the biotin to the antigen surface[98]. This hypothesis is supported by the fact that synthetic antibody libraries

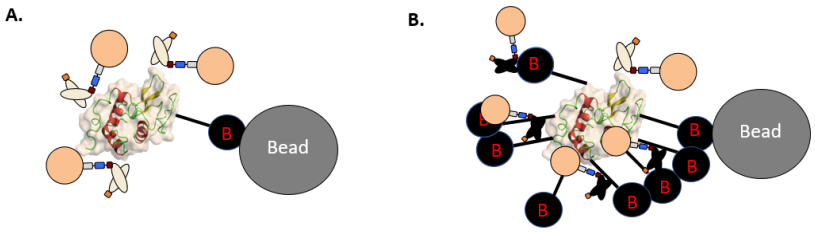

Figure 3.6. Over-biotinylation of sort antigen leads to the isolation of non-specific clones. A) The desired case of careful biotinylation of sort antigen where only 1-2 biotins are incorporated per antigen molecule. This results in the isolation of specific binders to the target of interest, which are shown in white. B) Poorly controlled biotinylation reaction in which the entire surface is covered in biotins. This results in the isolation of clones that bind to the biotin itself, the NHS-LC cross linker, or antigen that has changed shape because of the presence of excess biotin. Biotin binding clones are shown in black. There is still a genuine antigen binding clone shown in white, which is rare, but can be isolated if certain antigen epitopes have low biotin incorporation rate.

have previously been used to isolate binders capable of discriminating between the hapten and non-hapten form of a given antigen[64-66, 107]. The biotinylated form of an antigen mimics a hapten because the biotin is a small molecule that is about one thousandth the size of most model antigens. Even if isolated clones were not binding to the chemical cross linker, over-biotinylation can change the conformational shape of antigen[105]. Incorporated biotin can disrupt interactions between amino acids, resulting in a change in shape. For example, the NHS-LC biotin incorporates into the free amines of lysine residues with high frequency[98]. This could disrupt certain charge-charge interactions necessary to hold antigen in a given shape. Thus, clones isolated against over-biotinylated antigen may bind to antigen with a different conformational shape than the canonical form.

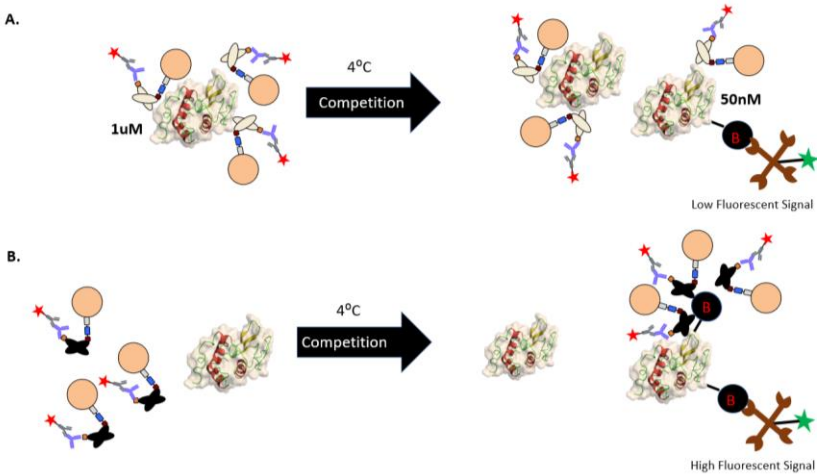

Figure 3.7. Antigen competition experiment to evaluate specificity of isolated binding clones. A) Clones are specific to non-biotinylated antigen. When incubated with a high concentration of non-biotinylated antigen, all clones bind. When biotinylated antigen is added at lower concentrations and low temperatures, clones will not bind because lower concentrations and temperatures discourage binding. Low fluorescence will be observed because the fluorophore selectively binds to the biotin on the antigen surface. B) Clones are specific to the biotinylated form of the antigen. When incubated with a high concentration of non-bitoinylated antigen, no binding will occur. When biotinylated antigen is added at lower concentrations and low temperatures, binding will still occur because the binding sites of these clones was not occupied with the non-biotinylated antigen. High fluoresce will be observed.

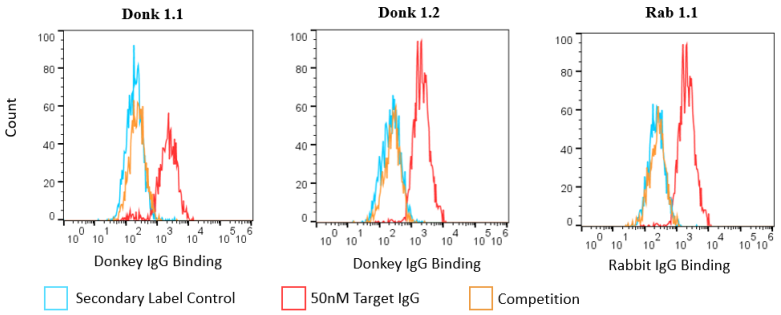

Figure 3.8. Competition experiment between biotinylated and non-biotinylated antigen for sort 1 clones that did not show diminished binding. Un-induced cells were gated out using Cmyc labeling. In the biotinylated antigen condition, cells were labeled with 50nM of biotinylated target antigen and Streptavidin 488 secondary. In the secondary control, cells were only labeled with the Streptavidin secondary. In the competition condition, cells were first labeled with 1µM non-biotinylated antigen followed by 50nM biotinylated antigen and the streptavidin 488 secondary. X-axis represents streptavidin 488 signal and Y-axis represents number of events. 10,000 events per sample were collected.

To evaluate this hypothesis, we ran an experiment that tested whether nanomolar concentrations of biotinylated antigen could compete with micromolar concentrations of non-biotinylated antigen for binding. A schematic of this experiment is shown in Figure 3.7. We tested each of the clones that showed un-diminished binding in initial binding assays. If clones are specific to the non-biotinylated form of the antigen, the biotinylated antigen cannot compete with the non-biotinylated form for binding because it is present in such low relative concentrations. In contrast, if a clone is specific to the biotinylated form of an antigen, it will only recognize that form. The biotinylated antigen will thus be

able to out-compete the non-biotinylated form independent of concentration. The results of this experiment are shown in Figure 3.8

Unsurprisingly, the shift in fluorescence of the 50nM target antigen labeled cells as compared to the control is the same as described in Figures 3.4 and 3.5 for all clones. The histogram representing the competition condition is almost directly overlaid with the secondary control condition for all clones tested. This means that the competition population had no increase in fluoresce over the control. Since the fluorescent label only binds to biotin, the lack of increase in fluorescence over the control means that clones did not bind to biotinylated antigen. If clones did not bind to the biotinylated antigen, their binding sites were occupied by non-biotinylated antigen. Thus, these isolated clones did not preferentially bind to biotinylated antigen.

The fact that these clones retained binding while others did not could potentially be attributed to the fact that they bind to an antigen epitope with little free amines. This would be the case for epitopes far from the N-terminus of the antigen and with a low amount of lysine residues[105]. Clones that showed diminished binding in Figures 3.4-3.5 also preferentially bound to biotinylated antigen in competition experiments run with the original sort antigen (Appendix, Figure A4)[7]. When taken together, these results support the hypothesis that the diminished antigen binding was caused by over-biotinylation of antigen.

In an attempt to further support our hypothesis, we compared the CDRH3 sequences of 9 clones that were shown to be biotin binders were compared to that of 6 clones that were genuine antigen binders binders and to a random sampling of 100 library members that were sequences for library validation as previously described[7]. We did so to see if biotin binding clones were enriched for in different amino acids, which could be indicative of a different mode of binding. A graph comparing the relative prevalence of each amino acid for each population is shown in Figure 3.9. The graph shows that biotin binders were slightly enriched for in alanine and tyrosine as compared to the overall library population and the genuine antigen binding population. Alanine had a relative frequency of 0.13 in biotin binders, as compared to 0.072 for genuine antigen binders and 0.064 for the library population. Tyrosine has a frequency of 0.34 in the biotin binding population, and 0.28 and 0.19 for the genuine antigen binding and library populations respectively. It is possible that this enrichment occurred because alanine and tyrosine can form favorable hydrophobic interactions with the biotin's carbon cross linker. However, these observed trends should be extrapolated with caution. This sample size is small, and more data would be needed to draw more definite conclusions. In any case, we had to assess and modify the lab's antigen preparation protocol before continuing with validation of the CDRH3 library.

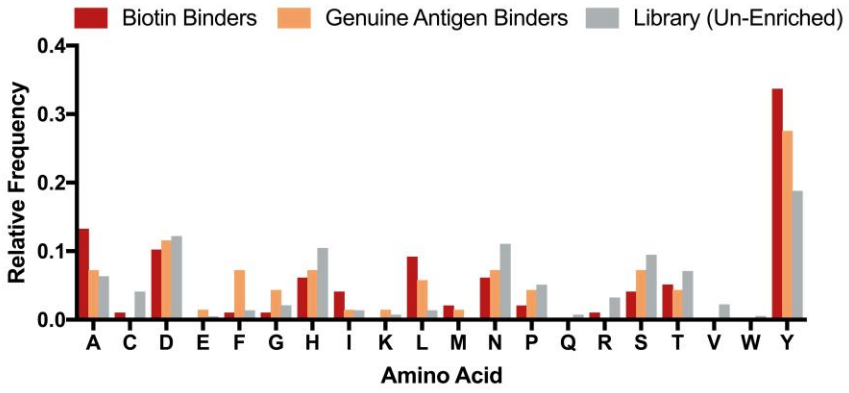

Figure 3.9. Relative amino acid frequency for biotin binding clones, genuine antigen binding clones and the un-enriched library. 9 clones were sequenced in the biotin binder population, 6 for the genuine antigen binder population and 100 for the library. Only variable portions of the CDRH3 sequences are compared in the graph. Relative frequency is defined as the ratio of number of times that amino acid appears in a given population to the total number of amino acids in a population.

3.3: Optimization of Antigen Preparation Protocol to Reduce Chances of Isolating Non-specific Binding Clones

We began the process of optimizing the chemical biotinylation protocol by analyzing lab notebook entries of prior group members to pinpoint potential experimental errors. We found that in past protocols, a 90-fold molar excess of biotin as compared to antigen was being added to the reaction. In contrast, we found that when re-preparing the antigen only a 5-fold molar excess of biotin was required at most to get 1-2 biotins per protein molecule incorporated. Moreover, the reaction kit guidelines suggest that a 20-fold molar excess of biotin should be sufficient to label an IgG with 8-12 biotins per protein molecule[98]. Based off the extreme excess of biotin added to the reaction, it is likely that all free amine residues on sort antigen had biotin incorporation during library screening.

However, even with the excessive addition of biotin, the absorbance assay was still reporting biotinylations within the target ratio of 1-2 biotins per mole of protein. We therefore needed to determine whether the assay itself was faulty. Further analysis of lab notebook entries revealed that the source of error was the fact that the absorbance assay is heavily dependent on protein concentration. After the biotinylation reaction, the antigen is placed through a series of two desalting columns to remove any un-reacted biotin and cross-linker byproduct from the solution[98-99]. Approximately 70% of the protein passes through the column, meaning that concentration decreases from the initial value. However, in the original protocol the concentration of protein was assumed to be constant. As shown in Figure 3.2, one biotin molecule displaces one avidin resulting in a change in absorbance in the validation assay[100]. If the concentration of the antigen is over-estimated, this change in absorbance will be divided among too many protein molecules, therefore underestimating the extent of biotinylation.

We hypothesized that chemical biotinylation, if carefully executed, can still yield the desired ratio of 1-2 biotins per mole of protein. We wanted to verify that the antigen was still being biotinylated using a secondary validation method in addition to the colorimetric method. Streptavidin beads are big enough to pass through the flow cytometer, so the beads were coated with biotinylated antigen. Streptavidin 488 could then be used to label free biotin of the antigen attached to the beads. The results of this validation for a wide range of different model antigens are shown in Figure 3.10. The antigens were chosen to represent different types of proteins, both IgGs and enzymes, with different molecular weights.

The histogram for the antigen coated beads for all model antigens tested shows a shift along the X-axis as compared to the streptavidin 488 control. This shift along the X-axis means an increase in streptavidin 488 signal. Since streptavidin selectively binds to biotin, this in turn corresponds to an increase in biotin detection, meaning that the antigen was biotinylated. Moreover, the histograms are relatively narrowly distributed meaning that a relatively uniform extent of biotinylation occurred across antigen molecules. The flow cytometry experiment only confirms the presence of biotin on antigen and does not provide information on the extent of biotinylation. Even so, the results of this experiment suggest that when the stoichiometric amount of biotin added to antigen is drastically decreased, chemical biotinylation of the antigen

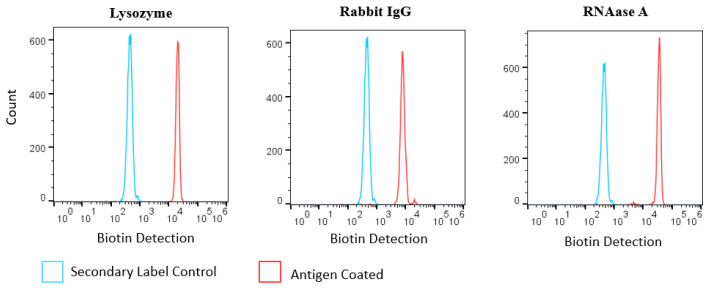

Figure 3.10. Flow cytometry validation of antigen biotinylation. Beads were prepared according to the standard protocol used for library screening and were then labeled with the fluorophore streptavidin 488. The secondary label control beads were blank beads that were then labeled with the same fluorophore as the antigen coated beads. The X-axis represents biotin detection and the Y-axis represents number of single bead events.

still occurs to enough of an extant to be useful in both library screening and flow cytometry applications.

3.4: Conclusion

In conclusion, we found that CDRH3 library clones that had previously been isolated against a panel of model IgG antigens as a proof of concept no longer bound to the same antigen when it was freshly biotinylated. We developed a hypothesis that this was because over-biotinylation of antigen used in library screening. Initially isolated clones likely either bound to the chemical cross-linker of the biotin or to conformationally altered epitopes that formed because of excess biotin. We therefor had to re-evaluate the lab's biotinylation protocol to prevent over-biotinylation from occurring in the future. We optimized the protocol to incorporate only 1-2 biotins per protein molecule. We hypothesize that this updated protocol will allow for all of the canonical antigen epitopes to be exposed during screening and will therefore result in the isolation of genuine antigen binding clones. To validate this hypothesis, we next had to re-screen the CDRH3 library using antigen prepared according to the optimized protocol.

3.5: References

1 Ackerman, M. *et al.* Highly avid magnetic bead capture: an efficient seleciton method for de novo protein engineering utilizing yeast surface dispalay *Biotechnol Prog.* **25**, 774-783 (2009).

2 Kehoe, H. *Targeting the tumor microenvironment with protein-small molecule hybrids* Masters of Science in Bioengineering thesis, Tufts University (2017).

3 Angelini, A. *et al.* in *Yeast Surface Display* Vol. 1319 *Methods in Molecular Biology* 3-36 (2015).

4 Kay, B., Thai, S. & Volgina, V. in *High throughput protein expression and purification* Vol. 498 185-198 (2011).

5 Fairhead, M. & Howarth, M. in *Site-Specific Protein Labeling* Vol. 1266 171-184 (2014).

6 *Pierce Biotechnology Instructions: EZ-Link NHS-biotin reagents*, <https://www.thermofisher.com/order/catalog/product/21336> (2018).

7 Azim-Zadeh, O. *et al.* Use of biotin derivatives to probe conformational changes in proteins *Journal of Biological Chemistry* **282**, 21609-21617 (2007).

8 *Pierce Biotechnology Instructions: Pierce Biotin Quantification Kit*, <https://www.thermofisher.com/order/catalog/product/28005> (2018).

9 Vermeer, A. & Norde, W. The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein. *Biophysical Journal* **78**, 394-404 (2000).

10 Ewert, S., Huber, T., Honegger, A. & Pluckthun, A. Biophysical properties of human antibody varriable domains *Journal of Molecular Biology* **325**, 531-553 (2003).

11 Mahon, C. *et al.* Comprehensive interrogation of a minimalist synthetic CDR-H3 library and its ability to generate antibodies with therapeutic potential. *Journal of Molecular Biology* **425**, 1712-1730 (2013).

12 de Haard, H. *et al.* A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *Journal of Biological Chemistry* **274**, 18218-18230 (1999).

13 Vaughan, T. *et al.* Human antibodies with sub-nanomolar affinity isolated from a large non-immunized phage display library *Nature Biotechnology* **14**, 309-314 (1995).

14 Feldhaus, M. *et al.* Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces Cerevisiae surface display library *Nature Biotechnology* **21**, 163-170 (2003).

15 *Thermo Fisher Scientific Zeba Desalting Products* <https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-purification-isolation/protein-dialysis-desalting-concentration/zeba-desalting-products.html> (2018).

Chapter 4: Re-screening the CDRH3 Library to Verify Specific Binders Can be Isolated

4.1: Introduction

The ultimate purpose of the CDRH3 library is to generate specific binding protein fragments that can act as scaffolds into which further functionality can be incorporated in the form of small molecules or grafted peptides. We envision the application of the library as generating protein-small molecule hybrids for the targeted inhibition of specific members of structurally similar, yet functionally diverse enzyme families. Therefore, the library must yield binding proteins capable of discriminating against structurally similar targets. However, we sought to start library screening with a panel of model antigens outlined in Table 4.1 instead of directly with a model enzyme family such as the MMPs. Starting with a panel allows us to critically evaluate the performance of the library. Specifically, we can pinpoint the types of antigens the library fails to isolate binders to and adjust the library design strategy to correct for these shortcomings.

Model Antigen	Molecular Weight (kD)
Polyclonal Bovine IgG	160
Polyclonal Donkey IgG	150
Polyclonal Rabbit IgG	150
Hen Egg White Lysozyme	14.3
Bovine Pancreases RNAase A	13.7
Horseradish Peroxidase	44

Table 4.1. Model antigens for CDRH3 library validation. Model antigens used to validate the CDRH3 library included both enzymes and IgGs with a wide range of molecular weights. Polyclonal IgGs were purchased from the serum of non-immunized animals. All enzymes were purchased as lyophilized powders.

IgGs are relatively conserved across mammalian species, all consisting of the Fc region attached to two Fabs. The Fabs contain the heavy and light chains, each chain with 3 complementary determining region (CDR) loops[90]. Moreover, IgGs are also all the same size across mammalian species at 150kD[8, 108]. The exception is bovine IgG, which has a slightly higher molecular weight at 160kD because of an abnormally long stalk and knob CDRH3[109]. Therefore, a panel of different mammalian IgGs are useful as model antigens to validate the CDRH3 library. The structural similarity across antigens in this panel mimic that of the eventual therapeutic targets of the library, such as the matrix metalloproteinases.

While the IgGs can test the capability of the library to discriminate between structurally similar targets, we also included model enzymes in the antigen panel to evaluate the ability of the library to yield binders against different sized targets with different conformational shapes. Hen egg-white lysozyme, RNAase A, and horseradish peroxidase were chosen as model enzymes because they have been previously used in library validation[96, 110-111]. This would allow us to compare the number of clones isolated and their affinity to other libraries. Moreover, as shown in Table 4.1 these enzymes range from one third the size of IgGs to a full order of magnitude smaller. They therefore allow us to determine whether the library is

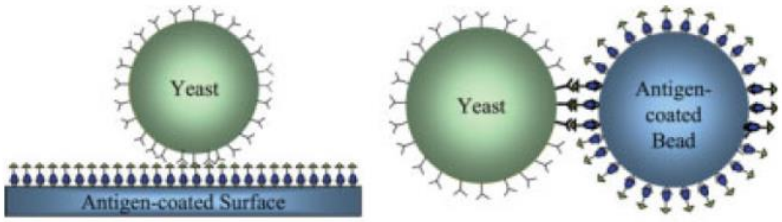

Figure 4.1. Multiple avidity of magnetic bead-based library enrichments**.** Antigen coated magnetic beads allow for a greater contact surface area with yeast displayed antibody fragments as compared to 2-D coated surfaces. This creates a multiple avidity effect which boosts the effective affinity of the yeast displayed constructs, allowing for the isolation of weaker clones. Figure from Ackerman et al. (2009)[2, 9].

limited to the size of antigens it can effectively produce binders against.

Once we selected our model antigens, we could begin screening the CDRH3 library. The main methods used to screen a yeast displayed antibody library are magnetic bead-based sorting[9], bio panning[82] and florescent activated cell sorting (FACS)[66, 79]. We initially use magnetic bead-based sorting to pare down the naïve library, after which FACS was used to achieve a more efficient enrichment.

We use magnetic bead-based sorting because the beads facilitate a greater contact surface area between the coated antigen and yeast displayed constructs as compared to 2-D surfaces, which is shown in Figure 4.1. This creates an avidity effect, boosting the effective affinity of the yeast displayed antibody fragments. The multiple avidity allows for the isolation of weaker and rarer clones than with other screening methods[9]. Weaker clones can still be viable leads for affinity maturation or for further engineering[112].

4.2: Screening CDRH3 Library Against IgG Panel Yields Novel Clones

Screening the CDRH3 library against the panel of IgGs took four rounds of magnetic bead-based sorting followed by a 5th round with FACS. Multiple rounds of enrichment are required because inefficiencies result from non-binding clones non-specifically sticking to beads or clumping to binding yeast. With each progressive round of screening, the population becomes more enriched for binding clones as less non-binding yeast are retained. Qualitative flow cytometry dot plots depicting the progressive library enrichment are shown in Figures 4.2. A quantitative summary of these results describing percent induced cells that are target binding for each round of sorting is shown in Figure 4.3.

Figure 4.2 shows flow cytometry dot plots with quadrant gating. Un-labeled cells have a basal Cmyc detection and antigen binding signal. Moreover, cells that are only Cmyc labeled have a basal antigen binding signal. To overcome this background signal, quadrants are positioned using an unlabeled control to separate Cmyc positive and negative populations and a Cmyc single color control to separate antigen binding from non-binding populations.

Cells in quadrant 4 have no Cmyc detection and no antigen binding signal. These cells are un-induced, meaning they do not have any constructs expressed on their surface. This is expected, because the induction efficiency of the yeast display system is not 100%.

Cells in quadrant 3 exhibit elevated levels of Cmyc detection, but no antigen binding signal. These cells are induced but not binding, meaning cells express constructs on the surface that do not bind to the target antigen. Cells in this quadrant have a wide distribution along the Cmyc detection axis. This is true because the expression level of constructs on the yeast surface is variable between single cells.

Cells in quadrant 2 have both Cmyc detection and antigen binding signal. These cells are induced and display constructs that bind to the antigen of interest. Cells do not appear in this quadrant until after the 3rd round of library screening for the donkey and rabbit IgG sorts, and in the 4th round of screening for the bovine IgG sort. After the initial appearance of binding cells, progressively more cells are present in the second quadrant after each successive round of sorting. This is evident in the fact that the percentage of the population in quadrant 2 increases progressively after each round of sorting for each IgG. The exception is the bottom FACS population, which does not show a higher percentage of cells in quadrant 2 as compared to the prior round of sorting. This is true because the gate for the lower FACS population was drawn close to quadrant 3, such that some induced but non-binding cells likely got isolated along with binding clones. Figures showing FACS gates are in the appendix.

The flow cytometry plots shown in Figure 2 do not just show whether a given population has binding clones but can also show the relative affinity of multiple clones in the same population. Higher

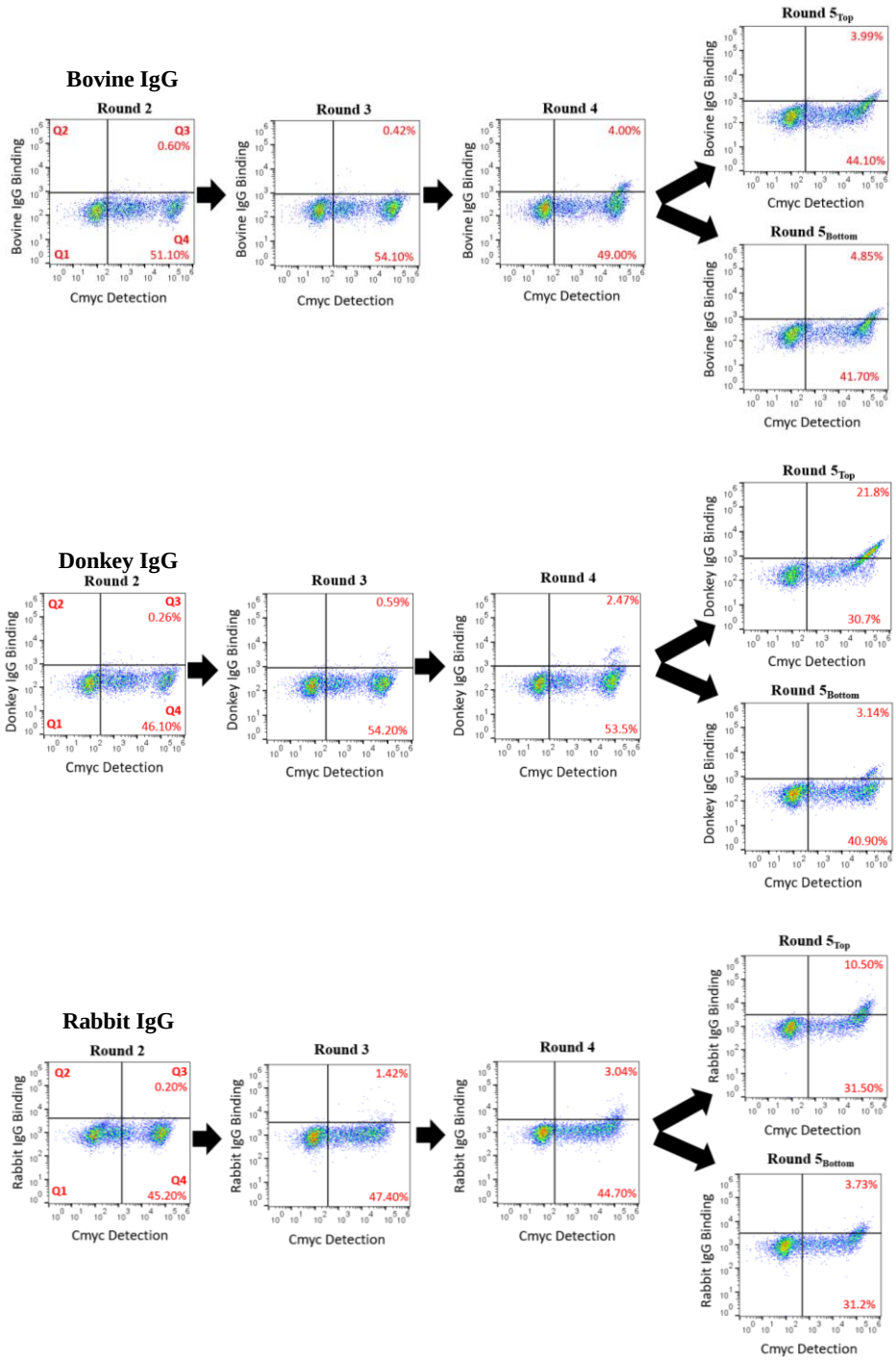

Figure 4.2. CDRH3 enrichment for IgG screening. Results are shown for rounds 2-5 of library screening. Rounds 1-4 were done using magnetic bead-based sorting and round 5 was done using FACS. At FACS, high and low gates were drawn, splitting the sort sample to "top" and "bottom." Cells were labeled with 50nM target IgG and a streptavidin-Alexa Flour 488 conjugated secondary for bovine and donkey IgG sorts or anti-biotin PE secondary for rabbit IgG sorts. X-axis represents Cmyc detection and Y-axis represents target IgG binding. The 4th quadrant represents un-induced cells, the 3rd quadrant induced but not binding cells, and the 2nd quadrant induced and binding cells. The 1st quadrant would represent un-induced cells that bind to target antigen. Since cells cannot bind without an scFv on the surface, this is not theoretically possible and no cells should be present in this quadrant. Red numbers denote the percentage of total population present in the given quadrant. Bolded red numbers indicate the quadrant number. Data not taken after round 1.

affinity clones will bind more antigen and will have a greater antigen detection signal than lower affinity clones. For all of the IgG targets, only a single binding population is evident even after 5 rounds of sorting. This suggests that enriched populations either only contain a single clone, or a group of clones with similar affinities for the target antigen.

The steady and progressive enrichment that occurred over multiple rounds of library sorting is further evident in analyzing quantitative enrichments as shown in Figure 4.3. The percent enrichment represents the percentage of all induced cells in a sort population that bind to the target antigen. Similar to the results shown by the flow cytometry scatter plots, percent enrichments above 0 first appear after the 3rd round of screening for the donkey and rabbit IgG sorts and after the 4th round for the bovine IgG sorts. The percent enrichment then increases steadily after each round of sorting.

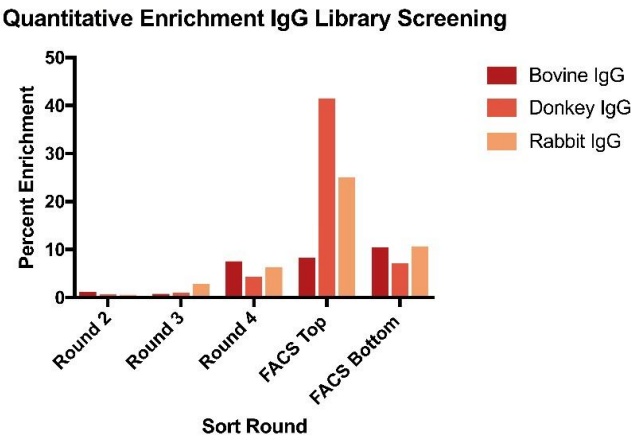

Figure 4.3. Percent enrichment represents the percent of all induced cells with any construct expression that bind to the target of interest. Percent Enrichment $= \frac{Q2}{Q2+Q3}$. Both FACS Top and FACS Bottom populations are from the 5th round of library screening.

The fact that binding clones do not appear until later rounds of screening, and that enrichment occurs steadily over multiple rounds of screening is important in that it indicates that certain clones are not over-abundant in the library such that they might out-compete and mask other clones[113-114]. For example, Figures 4.2 and 4.3 show that donkey IgG screening yielded a much higher percent of binding clones after the 5th round of screening as compared to the other IgGs in the panel. However, binding clones did not appear until the same round of sorting as rabbit IgG and appear just one round before bovine IgG binding clones. This indicates that the high extent of enrichment observed is likely due to the high affinity of binding clones instead of those clones being over-represented in the library population.

After five rounds of library screening, sufficient library enrichment was observed to isolate single clones and sequence them. This was done by miniprepping the output population from the 5th round of sorting, transforming E. coli with the isolated plasmids and picking individual colonies. Figures 4.4-4.5 show the results of this sequencing for isolated bovine and donkey IgG binding clones. Sequencing information from the Rabbit IgG population was not available at the time of this thesis.
Importantly, none of the bovine IgG binding clones isolated in the first round of sorting were also isolated during the second round. This is true for both the binding clones shown in Figure 4.4 (A) and the non-specific clones shown in the appendix. The fact that different clones are being isolated during different rounds of screening means that the library is being subjected to different selection criteria than during initial screening. This is likely due to biotins no longer masking critical antigen epitopes.

In contrast, Figure 4.5 (A) shows that for the donkey IgG screen the same clone was isolated in the second round of screening. However, this clone had previously shown to genuinely bind to target antigen. As previously discussed, it is possible the epitope of donkey IgG that this clone binds to has a low biotin incorporation rate and was thus not impacted by over-biotinylation. It is also possible this clone had such a high target affinity that it was able to recognize its epitope even with over-biotinlation. Additionally, Donkey 1.3, which was previously shown to preferentially bind to biotin was not isolated in the second round of sorting. This further supports the fact that the library is being exposed to different selection criteria than in previous rounds of sorting.

However, Donkey 1.2 which was shown to be a genuine antigen binding clone was not re-isolated. This is unsurprising given the induction efficiency of yeast and the efficiency of library screening in general[115]. It is possible that this clone was not induced or had low construct expression or was underrepresented during the naïve library sort. This would have prevented the clone from being enriched.

On the amino acid level, no clones isolated in the second round of screening in Figures 4.4 or 4.5 (A) had any tryptophan or cystine residues in their CDRH3 sequences. Tryptophan has previously been implicated in cross-reactivity and polyspecificity[96] and disulfide bonds formed by cystine can disrupt antibody structure[95]. The CDRH3 sequences had a relatively high prevalence of tyrosine. This corroborates with previous studies showing this residue is implicated in forming specific interactions with antigen[97].

In comparing Figures 4.4 (B) and (C), the upper population from FACS had less non-binding clones than the lower population. This supports the results of Figure 4.2, which shows a higher percentage of non-binding clones retained in the bottom population from FACS. Bovine 2.1 was present in both populations with approximately the same frequency. However, Bovine 2.2 and 2.3 were present in the top population with approximately twice as great a frequency as the bottom population. This could mean that these clones are slightly higher affinity than Bovine 2.1.

Figure 4.5 (B) shows a low percentage of non-binding clones, with most of the population being Donkey 2.1. The high prevalence of Donkey 2.1 after sort round 5 further supports the hypothesis that this clone has a very high affinity for its target. In fact, the K_D of this clone was previously estimated to be in the low nanomolar range (data not shown)[7]. Donkey 1 may have such a high affinity for its target because it binds to a hot spot epitope on the donkey IgG antigen[116-118]. Hot spots are regions of an antigen that form protein-protein interactions extremely favorably. They have a high proportion of tryptophan, arginine and tyrosine,

A

Clone	CDRH3 Sequence	CDRH3 Loop Length
Bov 1.1	YLDHYYYALDY	11
Bov 1.2	YYDHHYYALDY	11
Bov 1.3	YLDNHYYALD	11
Bov 2.1	HYDPYDDYNFYFGFDY	16
Bov 2.2	TYYDNSYAYTALDY	14
Bov 2.3	SSIYDGYMYGLDY	13

B

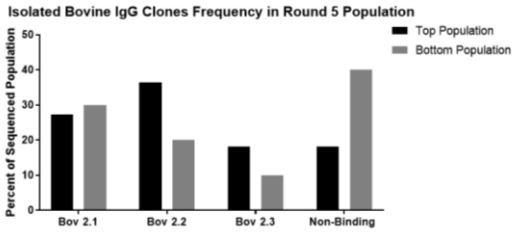

Figure 4.4. Sequencing analysis of bovine IgG binding clones. A) CDRHH3 loop sequences for all isolated binding clones. Clones highlighted in yellow are from the second round of screening. All other clones were from the first round of screening and determined to be biotin binders. CDR loop length reported according to Kabat scheme[3]. B) Relative prevalence of each isolated clone in the top and bottom FACS population. The top population is shown with black bars and the bottom population with grey bars.

A

Clone	CDRH3 Sequence	CDRH3 Loop Length
Donk 1.1	YDKTHHNPDYALDY	14
Donk 1.2	YENSYALDY	9
Donk 1.3	YYDPDRYNYAMDY	13
Donk 2.1	YDKTHHNPDYALDY	14

B

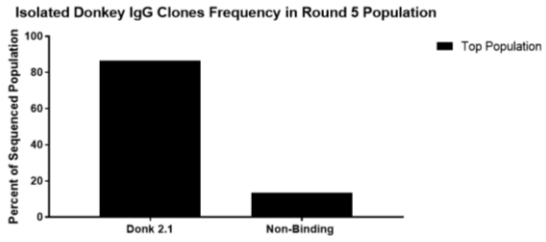

Figure 4.5. Sequencing analysis of donkey IgG binding clones. A) CDRHH3 loop sequences for all isolated binding clones. Clones highlighted in yellow are from the second round of screening. CDR loop length reported according to Kabat scheme[3]. B) Relative prevalence of each isolated clone in the top FACS population. The bottom FACS population was not enriched enough for sequencing.

residues that readily form interactions with other amino acids. Hot spots are also surrounded by a hydrophobic cushion to exclude bulk solvent from the binding interface[116-117]. Hot spots often occur clustered near other hot spots[119]. Thus, it is likely that there are other clones in the population that bound to donkey IgG with a lower affinity and were out competed by Donkey 1. To analyze the rest of the library, Donkey 1 could be produced in soluble form and used to block the epitope of antigen that this clone binds to. This would in turn allow for the isolation of clones binding to other epitopes. An alternative explanation for the high affinity of this clone is that the CDRs left at the germline sequence by chance were able to form binding interactions with the antigen. This would have boosted the affinity in comparison to clones that only use the CDRH3 to mediate binding. When taken together, the sequencing analysis of clones supports the hypothesis that the library was now being screened for antigen itself and that antigen was no longer masked by excess biotin.

4.3: Screening CDRH3 Library Against IgG Panel Yields Highly Specific Clones

The first step in evaluating the specificity of clones isolated from the CDRH3 library re-screen was to verify that clones bind to the antigen of interest instead of biotin. To do this, we tested whether low concentrations of biotinylated antigen could compete with high concentrations of non-biotinylated antigen for binding. If clones are specific to the non-biotinylated form of the antigen, the biotinylated antigen cannot compete with the non-biotinylated form for binding because it is present in such low relative concentrations. In contrast, if a clone is specific to the biotinylated form of an antigen, it will only recognize that form. The biotinylated antigen will thus be able to out-compete the non-biotinylated form independent of concentration. We tested isolated bovine IgG single clones and rabbit IgG polyclonal populations. We did not test donkey IgG, since no new clones were isolated. Biotinylated antigen competition experiments for previously isolated donkey IgG binding clones are shown in chapter 3. The results of this experiment for isolated single bovine IgG binding clones and rabbit IgG binding populations are shown in Figure 4.6.

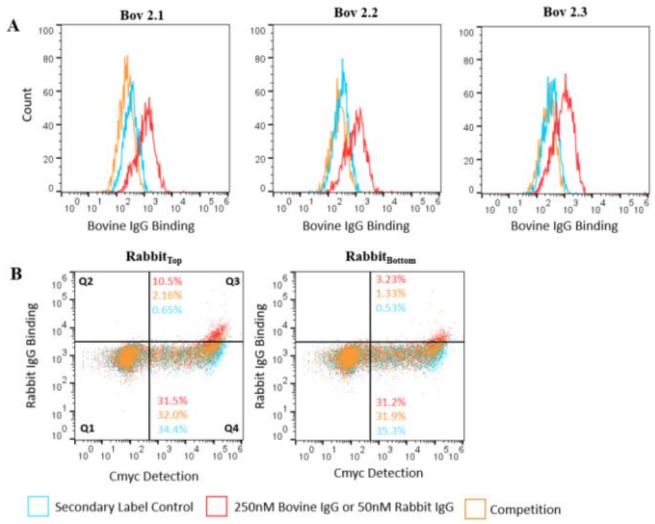

Figure 4.6. Biotinylated vs non-biotinylated antigen competition experiment for new clones isolated in the second CDRH3 library sort. A) Bovine IgG isolated single clones were labeled with 250nM biotinylated antigen in the antigen binding condition and 5uM non-biotinylated antigen with 250nM biotinylated antigen in the competition experiment. Streptavidin 488 was the secondary label in all conditions. X-axis represents antigen binding and Y-axis represents number of single celled events. Only induced cells are shown in the histogram. B) Rabbit IgG top and bottom FACS population labeled with 50nM biotinylated antigen in the antigen binding condition and 1uM non-biotinylated antigen with 50nM biotinylated antigen in the competition condition. Anti-biotin PE was used as the secondary label in all conditions. X-axis represents Cmyc signal and Y-axis represents antigen binding. Numbers in each quadrant represent the percentage of each population in that given quadrant. Bold black numbers describe the quadrant number.

In Figure 4.6 (A), 250nM biotinylated antigen was used to label bovine IgG binding clones instead of the

standard 50nM. This was done because clones are relatively weak, and thus a low antigen binding signal occurs with just 50nM. The histograms for the competition condition and the secondary label control are almost directly superimposed upon one another. In other words, the competition condition does not show a shift along the antigen binding axis as compared to the control. Since the fluorophore for antigen binding signal binds specifically to the biotin on biotinylated antigen, this means that cells in the competition condition did not bind biotinylated antigen. If cells did not bind biotinylated antigen, their binding sites were already saturated with the high concentrations of non-biotinylated antigen. Thus, all isolated bovine IgG binding clones are genuine antigen binders.

Single rabbit IgG binding clones have not yet been isolated, and therefore the competition assay was run on the entire population output from the 5th round of screening as shown in Figure 4.6 (B). A dot plot is used to illustrate these results because binding clones represent at most 10% of the overall population, which is too rare to be effectively visualized on a histogram. For both the top and bottom populations, only the dot plot for the biotinylated antigen condition has a substantial number of cells in the second quadrant, which represents induced clones binding to target antigen. The top population biotinylated antigen condition had 10.5% of the population in quadrant 2, as compared to 3.23% for the bottom population. In contrast, the competition condition had a relatively small number of cells in the second quadrant, which was comparable to the secondary control. For the competition condition, 2.16% and 1.33% of the top and bottom populations were in quadrant 2 respectively and for the control condition this was 0.65% and 0.53%. Since cells in the competition condition had a similar number of cells in the second quadrant as the control, the competition cells did not bind biotinylated antigen to a significant extent. This is because the binding sites of these clones were already fully occupied by the higher concentrations of non-biotinylated antigen. Thus, all the clones in the polyclonal rabbit IgG binding population are genuine antigen binders.

Once it was verified that isolated binders from the second round of screening were specific to antigen and not biotin, we sought to evaluate the specificity of these clones for the specific IgG they were screened against. To do this, we tested all IgG binding clones that had been verified as non-biotin binders against all other IgGs used as model antigens. These results are shown in Figure 4.7.

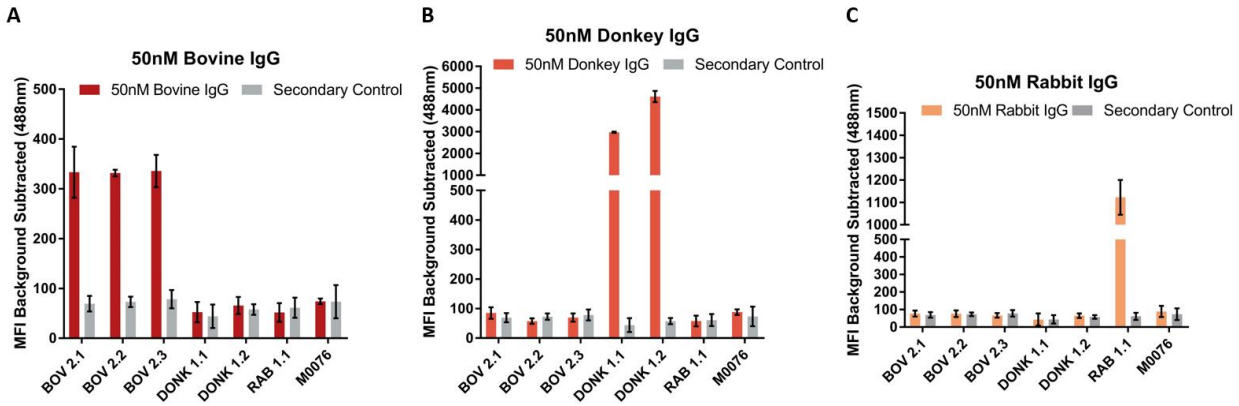

Figure 4.7. Cross-reactivity assay for all non-biotin binding IgGs isolated**.** All isolated IgG binding clones were labeled with: A) 50nm bovine IgG, B) 50nM donkey IgG, or C) 50nM rabbit IgG. Streptavidin 488 was used as the secondary fluorophore for all conditions. In the secondary only control, cells were labeled with Streptavidin 488 and no IgG. Only induced cells were analyzed and were gated out from the rest of the population using Cmyc expression. The Y-axis represents median subtracted fluorescence, meaning the median fluorescence corresponding to antigen binding signal of the Cmyc negative cells was subtracted off the values for the Cmyc positive cells. Error bars represent the standard deviation between 3 independent experimental trials. Y-axis for Donkey and Rabbit IgG labeled cells are broken at the maximum MFI value for Bovine IgG labeled cells for easier comparison.

The Y-axis for all plots in Figure 4.7 represents background subtracted median fluorescence intensity (MFI) corresponding to antigen binding signal. Background was subtracted off by subtracting the MFI value of the un-induced population from the induced population. This was done in an effort to mitigate the effects of the basal fluorescence of the cells, or of any non-specific sticking of secondary label to the cell surface. In Figure 4.7 (A), only bovine clones show a significant increase in fluorescence over the control. For all other clones, the fluorescence was the same as that of the control. Similarly, in (B) only donkey IgG binding clones showed a significant change in fluorescence as compared to the control. Lastly, in (C), only rabbit IgG binding clones showed a change in fluorescence as compared to the control. Since median fluorescence intensity corresponds to antigen binding signal, a change in florescence compared to the control represents antigen binding. Therefore, the results of Figure 4.9 show that clones only bind to the IgG they were screened against and no off-target binding occurs. The slight basal fluorescence of the control sample is likely due to slight non-specific binding with the fluorophore. M0076 is an scFv that specifically binds to matrix metalloproteinase 9[44]. Since MMP9 is an enzyme, there should in principle be no cross reactivity between this clone and any IgGs. We therefore included M0076 as a negative control. This helps to confirm that any interactions between an IgG binding clone and its target is due to genuine binding instead of non-specifically sticking to the yeast cell surface, which would register as binding on a flow cytometer. The fact that there was minimal non-specific binding between IgGs could be due to the fact that the CDRH3 library was depleted against TA99, a murine antibody before positive sorting[7]. These depletions may have eliminated any polyspecific IgG binders from the library. Indeed, depletions are often used in library screening to remove binders to structurally similar, but off-target antigens before progressing to positive screening[64, 120].

Together, these results demonstrate the ability of the CDRH3 library to isolate binders to a specific target of interest, that will not cross react with other structurally similar targets. This is important if the library will eventually be used to isolate inhibitors to specific enzymes of multi-enzyme families. Moreover, this is important if the library is ever to be used to generate therapeutics, since off-target effects can cause debilitating side-effects for patients.

4.4: Clones From CDRH3 Library Bind to Targets With Reasonable Affinity

After proving isolated clones bound only to the target they were screened against, we needed to estimate the affinity of the clones for their target to make sure they bind to their target with a high enough affinity for the intended library applications. We estimated the K_D of all isolated bovine IgG binding clones using titrations on the yeast surface and the resulting plots are shown in Figure 4.8. The K_D of all clones isolated during the first round of screening was previously determined (data not shown)[7].

Before performing the titrations on the yeast surface, we first wanted to validate that any binding observed during titrations would be due to genuine antigen binding to the yeast scFv. Figure 4.9 (A) shows this validation. We wanted to use a starting concentration of 1000nM in an attempt to capture the upper plateau region of our titration curves. However, the higher the antigen concentration, the greater the risk the antigen will non-specifically stick to the yeast surface and will still register as binding in the flow cytometer. We performed the same titration experiment with a bovine IgG binding clone (BOV 2.3) and M0076, the MMP9 binding scFv. The background subtracted MFI corresponding to antigen binding signal for M0076 is negligible at all concentrations as compared to that for the bovine IgG binding clone. There is a slight increase in MFI for M0076 at the highest antigen concentration, but this is negligible compared to the bovine IgG binding clone and is therefore not of concern. Thus, we proceeded with the titrations starting at a concentration of 1000nM.

As shown in Figure 4.8 (B)-(D), the affinity for all Bovine IgG binding clones is in the triple digit nanomolar range and were relatively close together. This supports the observation that after the 5th round

of library screening, the bovine IgG binding population blended together on flow cytometry dot-plots and sub-populations were not visible. The upper plateau for these titration curves is not observed. However, to capture this region of the curve for low affinity clones, an extremely high amount of antigen would have been required. There would have been a high risk of non-specific binding at such high concentrations, which would have made it difficult to collect accurate data.

Interestingly, the affinities of these isolated Bovine IgG binding clones were comparable to the triple digit nanomolar affinities reported in a previously described phage displayed minimalist CDRH3 scFv library. Clones isolated against the transmembrane receptor NOTCH 1 had K_D values ranging from 100-450nM[64]. However, the affinities reported for clones isolated from this library screen were relatively low as compared to other synthetic antibody libraries[62, 70, 97]. Synthetic antibody libraries have been shown to yield binders to targets with affinities in the double digit picomolar range[121-122]. While the CDRH3 may be the dominant loop in mediating antigen binding interactions, all other antibody loops also form favorable interactions with the antigen surface in canonical antibodies, boosting antibody affinity[108]. Without the other CDR loops, clones from the CDRH3 library can still recognize the antigen, but with lower affinity than if the antibody had diversity at all 6 loops[64, 66, 121].

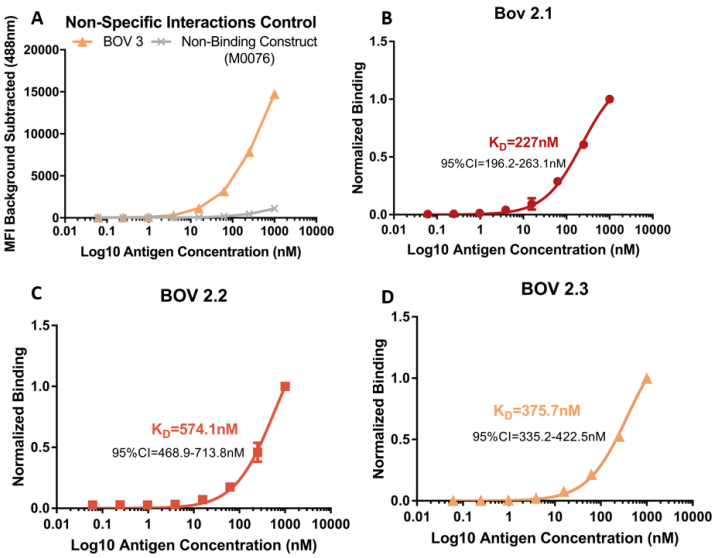

Figure 4.8. Titrations on the yeast surface to estimate affinity of isolated bovine IgG binding clones. Titrations were done by labeling cells starting with 1000nM bovine IgG and 4-fold serial dilutions. Un-induced cells were gated out using Cmyc expression and the background subtracted MFI was calculated by subtracting antigen binding signal of the un-induced population from the induced population. Normalized binding was calculated by taking the ratio of background subtracted MFI of any given dilution to that of the highest concentration sample. Results were input into GraphPad Prism and dissociation constant was estimated using a one-site specific binding model. A) Validation experiment to make sure that no non-specific binding was occurring during titration experiments. Titration was done with Bovine 2.3 alongside M0076, and scFv that binds to MMP9. Y-axis was left in terms of background subtracted MFI. B) Titration of Bovine 2.1. Error bars represent standard deviation of 3 independent experimental trials. One of the 1000nM samples was lost, so the average MFI of the other two was used to calculate normalized binding for that condition. C) Titration of Bovine 2.2. Error bars represent standard deviation of 5 independent experimental trials from 2 separate experiments. D) Titration of Bovine 2.3. Error bars represent standard deviation of 3 independent experimental trials.

While the affinities of antibodies isolated from this library are lower than those from canonical antibodies, the affinities are likely high enough to support the intended application of this library. Antibodies isolated from the library serve only as a backbone on which additional diversity in the form of a small molecule will be added. We predict that the synergy between the small molecule and the antibody in the protein-small molecule hybrids will boost the effective affinity of the antibodies. Therefore, while lower than conventional antibodies, the affinities of isolated clones reported here are likely high enough to support the usage of the CDRH3 for the engineering of protein-small molecule hybrids.

4.5: CDRH3 Library Fails to Yield Binders to Model Enzymes

While the library produced binders to IgG targets, we also sought to evaluate its ability to produce binders to enzymatic targets. This is important to ensure the library can produce binders to targets of a wide variety of sizes and shapes. Moreover, we hope to eventually generate protein-small molecule hybrids against enzymatic targets. The results after 4 rounds of magnetic bead sorting against the model enzyme targets is shown in Figure 9. 100,000 events were collected instead of the usual 10,000 and 250nM antigen was used instead of the usual 50nM in an effort to show any potential rare or weak binding clones.

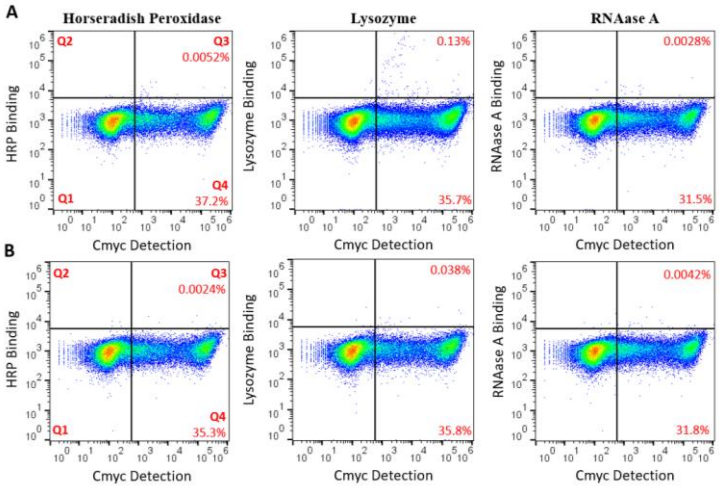

Figure 4.9. Binding assay after 4 rounds of bead-based library screening against model enzymes. A) Antigen binding samples labeled with 250nM biotinylated antigen and anti-biotin PE secondary label. B) Secondary control condition labeled only with anti-biotin PE secondary. Y-axis represents antigen binding at X-axis represents Cmyc detection. 100,000 events were recorded per sample. Red numbers represent the percentage of the population in each quadrant. Bolded red numbers show quadrant numbers.

Figure 4.9 shows that the percentage of cells in the second quadrant is approximately the same for both the antigen labeled condition in (A) and the secondary label control condition in (B). The second quadrant represents cells that bind to the target of interest. Thus, there were no binders present against any of the model enzymes after 4 rounds of bead-based screening. The only cells present in the population were the un-induced cells in the fourth quadrant or induced but non-antigen binding cells in the 3rd quadrant. Lysozyme shows slightly more cells in Q2 in (A) as compared to (B). However, these cells are very far to the left along the x-axis, suggesting low Cmyc expression. Moreover, the number of cells in Q2 for this sample as compared to what was observed after 4 rounds of sorting with IgGs is very low. Therefore, these cells are likely artifacts and not rare antigen binding clones.

We decided to explore Pymol structures of antibodies in complex with lysozyme to gain insight into the modes of binding of antibodies to the model enzymes we chose. We hoped this would help us pinpoint potential flaws in the library design that prevented us from isolating binders against these targets. We use lysozyme in our analysis because the most structures and sequencing information was available in the literature. The results of the analysis are shown in Figure 4.10.

Figure 4.10 (A) shows the structure of an antibody with CDRH3 loop length of 6,[1] the shortest loop in that specific antibody structure. Loops other than the CDRH3 mediate a significant portion of the interactions between the antibody and the antigen. Because the CDRH3 loop is so short, the other longer loops hinder it from making more contact with the enzyme surface[47]. These loops must form binding interactions with the enzyme surface in order for the CDRH3 to come close enough to bind.

In contrast, Figure 4.10 (B) shows the structure of an antibody with CDRH3 loop length 17. The CDRH3 is the longest loop in this antibody[5]. Because of its relatively long length, the CDRH3 extends

beyond the other loops and is not hindered from contacting the antigen[16, 123]. In this case, binding interactions with the CDRH3 appear to take place mostly independent of the other loops.

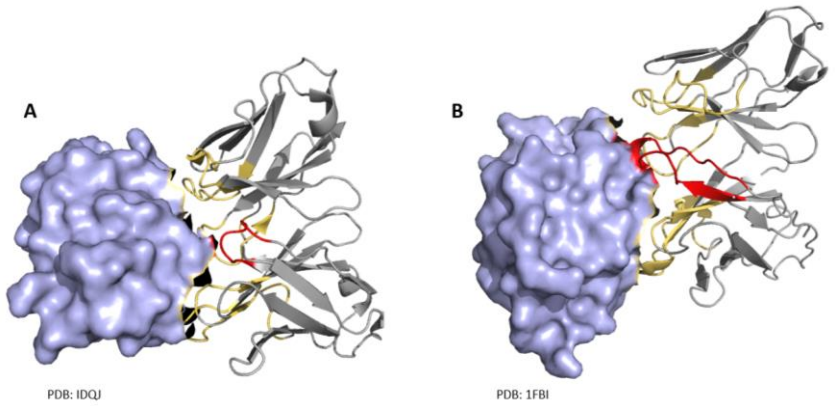

Figure 4.10. Pymol structures of different lysozyme binding antibodies. A) Fab fragment of antibody with CDRH3 loop length 6[1-2]. B) Fab fragment of antibody with CDH3 loop length 17[5]. Lysozyme antigen is shown in light blue. The framework of the antibody is shown in grey. The CDRH3 is shown in red and all other CDRs are shown in yellow.

When taken together, the results of our Pymol structure analysis show that if a CDRH3 loop is too short, the other antibody loops may hinder it from contacting the antigen surface. In the case of our antibody library, CDRH3 loop lengths were relatively short and were capped at 17. This loop length was also one of the least prevalent in the library[7]. For comparison, the longest CDRH3 loop length found in the human repertoire is 26.

In contrast to our library, a previous example of a CDRH3 library was able to yield binders to a wide range of targets with different molecular weights including enzymes. However, this library had CDRH3 loop lengths of up to 22. Moreover, 20% of isolated binders had CDRH3 loop lengths of 18-20, which were not present in our library. Thus, it is possible that we were not able to isolate binders to enzymatic targets because CDRH3 loop length was too short. The other loops were left at the germline sequence, meaning they may not have been able to form favorable interactions with the enzyme surface and may have instead hindered the CDRH3 from forming binding interactions. This is detrimental in the case of our minimalist library, because the CDRH3 is the only loop likely to form binding interactions. This hypothesis is supported by the information in Table 4.2, which compares the lengths of all the un-differentiated loops of the germlines to that of the CDRH3 in our library. In the DP47-JH4 heavy chain germline, the CDRH1 has a loop length of 5[2, 8, 124]. This is in principle too short to hinder any length CDRH3 from contacting an antigen surface. However, the CDRH2 is 17 amino acids long[2, 8, 124]. This loop length could be long enough to hinder even the longest CDRH3 loop from contacting an antigen surface. Similarly, in the DPK9-JK4 germline, all the light chain loops are longer than at least the shortest CDRH3 loop length. Particularly, the CDRL1 at 11 amino acid residues long is at least as long as three of the potential CDRH3 loop lengths[2, 8]. The other 2 light chains are also comparable in length to the shorter potential CDRH3 loop lengths. Thus, based off of the loop lengths of the germline, it is possible that the CDRH3 was prevented from forming binding interactions with the surface of our model enzymes.

Loop	V_H Length	V_L Length
CDR1	5	11
CDR2	17	7
CDR3	9-17	9

Table 4.2. Loop lengths for the DP47-JH4 heavy chain[4] and DPK9-JK4 light chain[2] germlines used in construction of the CDRH3 library. Lengths for all loops besides the CDRH3 are fixed because these were left un-differentiated at the wild type germline sequences in the library. The CDRH3 has variable length according to the diversification scheme of the library[7] as discussed in the introduction.

Potentially contradictory to this hypothesis is that in the past we were able to isolate a binder to fibroblast activation protein (FAP), an 88kDA enzymatic target (data not shown)[7]. Moreover, the CDRH3 of this clone was relatively short at 10 amino acids long. However, it is possible that in this case, by chance, the un-diversified loops were able to mediate contact with the antigen. Also interesting is the fact that we were still able to isolate binders to a wide range of IgG targets. However, the mode of binding to these IgG targets is different than to enzymes. The IgG antigen targets have loop like structures that could protrude past the un-diversified antibody loops and to the CDRH3 to allow for it to make binding contact.

4.6: Conclusion

In conclusion, by re-screening the CDRH3 with carefully biotinylated IgG, we were able to isolate clones that were genuine antigen binders and did not preferentially bind to biotinylated antigen. This validated our modified antigen biotinylation protocol and provided additional support for our hypothesis that non-specific clones had previously been isolated because of over-biotinylation.

We showed that the clones we isolated had triple digit nanomolar affinity for their target. Importantly, the clones had no affinity for off-target IgG. These results provide a strong proof of concept for the CDRH3 library, which we eventually hope can discriminate between members of multi-enzyme families.

We also found one of the potential pitfalls of the library: that it was unable to yield binders to a panel of enzymes. We used Pymol structures of antibodies bound to one of our model antigens to formulate a hypothesis that hinderance of the other antibody loops prevents the CDRH3 from contacting the enzyme surface and mediating binding. However, this is not concerning enough to warrant a re-design of the library because the eventual incorporated small molecule diversity will potentially be able to overcome this and result in the library being able to yield enzyme binders.

4.7: References

1 1. Li, Y.; Li, H.; Smith-Gill, S.; Mariuzza, R., Three-dimensional strcutures of free and antigen-bound Fab from monoclonal antilysozyme antibody HyHEL-63. *Biochemistry* **2000,** *39* (21), 6296-6309.

2. Rader, C.; Turner, J.; Heine, A.; Shabat, D.; Sinha, S.; Wilson, I.; Lerner, R.; Barbas, C., A humanized aldolase antibody for selective chemotherapy and adaptor immunotherapy *Journal of Biological Chemistry* **2003,** *332* (4), 889-899.

3. Kabat, E.; Wu, T.; Perry, H.; Gottesman, K.; Foeller, C., *Sequences of proteins of immunological interst* National Institutes of Health: 1991; Vol. 1.

4. Brezinschek, H.; Brezinschek, R.; Lipsky, P., Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. *THe Journal of Immunology* **1995,** *155* (1), 190-202.

5. Lescar, J.; Pellegrini, M.; Souchon, H.; Tello, D.; Poljak, R.; Peterson, N.; Greene, M.; Alzari, P., Crystal structure of a corss-reaction complex between Fab F9.13.7 and guinea fowl lysozyme *The Journal of Biological Chemistry* **1995,** *270* (30), 18067-18076.

6. Joyce, J.; Pollard, J., Microenvironmental regulation of metastasis *Nature Reviews Cancer* **2009,** *9* (4), 239-252.

7. Kehoe, H. Targeting the tumor microenvironment with protein-small molecule hybrids Tufts University Medford, MA 2017.

8. Jespers, L.; Schon, O.; James, L.; Verprinstev, D.; Winter, G., Crystal structure of HEL4, a soluble, refoldable human VH single domain with a germ-line scaffold *Journal of Biological Chemistry* **20004,** *337* (4), 893-903.

9. Ackerman, M.; Levary, D.; Tobon, G.; Hackel, B.; Orcutt, K.; Wittrup, K., Highly avid magnetic bead capture: an efficient seleciton method for de novo protein engineering utilizing yeast surface dispalay *Biotechnol Prog.* **2009,** *25* (3), 774-783.

10. Cancer Facts Sheet. http://www.who.int/mediacentre/factsheets/fs297/en/ (accessed April, 12, 2018).

11. Paradoll, D., The blockade of immune checkpoints in cancer immunotherapy *Nature Reviews Cancer* **2012,** *12* (4), 252-264.

12. Chang, Z.; Chen, Y., CARs: Synthetic Immunoreceptors for Cancer Therapy and Beyond *Cell* **2017,** *23* (5), 430-450.

13. Smith, I.; Proctor, M.; Gelber, D.; Guillamue, S.; Feyereislova, A.; Dowsett, M.; Goldhirsch, A.; Untch, M.; Mariani, G.; Baselga, J.; Kaufman, M.; Cameron, D.; Bell, R.; Bergh, J.; Coleman, R.; Wardley, A.; Harback, N.; Lopez, R.; Mallmann, P.; Gelmon, K.; Wilcken, N.; Wist, E.; Rovira, P.; Piccart-Gebhart, M., 2 year follow-up of transtuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet* **2007,** *369* (9555), 29-36.

14. Hinrichs, C., Cell-based Molecularly Targeted Therapy: Targeting Oncoproteins With T Cell Receptor Gene Therapy. *Journal of Clinical Investigation* **2018,** *128* (4), 1261-1263.

15. Hing, B.; Van Den Heuval, P.; Prabhu, V.; Zhang, S.; El-Deiry, W., Targeting tumor suppressor p53 for cancer therapy: strategies, challenges and oppertuniteis. *Current Durg Target* **2014,** *15* (10), 80-89.

16. Li, N.; Fu, H.; Hewitt, S.; Dimitrov, D.; Ho, M., Therapeutically targeting glypican-2 via single-domain antibody-based chimeric antigen receptors and immunotoxins in neuroblastoma. *PNAS* **2017,** *114* (32), 6623-6631.

17. Wang, R.; Wang, H., Immune targets and neoantigens for cancer immunotherapy and precision medicine. *Nature Cell Research* **2017,** *27* (1), 11-37.

18. Alatrash, G.; Jakher, H.; Stafford, P.; Mittendorf, E., Cancer immunotherapies, their safety and toxicity *Expert Opinion on Drug Safety* **2013,** *12* (5), 631-645.

19. Zugazzgoitia, J.; Guedes, C.; Ponce, S.; Ferrer, I.; Moina-PPindo, S.; Paz-Ares, L., Current Challanges in Cancer Treatment *Clinical Therapeutics* **2016,** *38* (7), 1551-1566.

20. Komatsubara, K.; Carvajal, R., The promise and challenges of rare cancer research. *Lancent Oncology* **2016,** *17* (2), 136-138.
21. Adams, J.; Smothers, J.; Srinvasan, R.; Hoos, A., Big oppertunities for small molecules in immuno-oncology *Nature Reviews Drug Discovery* **2015,** *14* (9), 603-622.
22. Bachovchin, D.; Cravatt, B., The pharmacological landscape and therapeutic potential of serine hydrolases *Nature Reviews Drug Discovery* **2012,** *11* (1), 52-68.
23. Coussens, L.; Fingleton, B.; Matrisian, L., Matrix metalloproteinase inhibitors and cancer: trials and tribulations *Science* **2002,** *295* (5564), 2387-2392.
24. Duffy, M.; Mullooly, M.; O'Donovan, N.; Sukor, S.; Crown, J.; Pierce, A.; McGowan, P., The ADAM family of proteases: new biomarkers and therapeutic targets for cancer? *Clinical Proteomics* **2011,** *8* (9).
25. Dufour, A.; Overall, C., Misisng the target: matrix metalloproteinase antitargets in inflammation and cancer *Trends in Pharmacological Sciences* **2013,** *34* (4), 233-242.
26. Egebald, M.; Werb, Z., New functions for the matrix metalloproteinases in cancer progression. *Nature reviews Cancer* **2002,** *2* (3), 161-174.
27. Erler, J.; Bennewith, K.; Micolau, M.; Dornhofer, N.; Kong, C.; Le, A.; Chi, J.; Jefferey, S.; Giaccia, A., Lysyl oxidase is essential for hypoxia-induced metastasis *Nature* **2006,** *440* (27), 1222-1226.
28. Larrinaga, G.; Perez, I.; Sanz, B.; Beitia, M.; Errate, P.; Fernandez, A.; Blanco, L.; Etxezarraga, M.; Gil, J.; Lopez, J. L., Dipeptidyl-Peptidase IV activity Is correlated with colorectal cancer prognosis. *PLOS One* **2015,** *10* (3).
29. Gupta, G.; Massague, J., Cancer metastasis: building a framework *Cell* **2006,** *127* (4), 679-695.
30. Matrislan, L.; Lynch, C., Matrix metalloproteinases in tumor-host cell communication. *Differentiation* **2002,** *70*, 561-573.
31. Menendez, J.; Lupu, R., Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis *Nature Reviews* **2007,** *7* (10), 763-777.
32. Kassenbrock, K.; Plaks, V.; Werbb, Z., Matrix metalloproteinases: regulators of the tumor microenvieornment *Cell* **2010,** *141* (1), 52-67.
33. Bonnas, C.; Chaou, J.; Werb, Z., Remodeling the extracellular matrix in development and disease *Nature Reviews in Molecular and Cellular Biology* **2014,** *15* (12), 786-801.
34. Liu, R.; Li, H.; Liu, L.; Yu, J.; Ren, X., Fibroblast activation protein: A potential theraputic target in cancer. *Cancer Biology and Therapy* **2012,** *13* (3), 123-129.
35. Nomura, D.; Long, J.; Niessen, S.; Hoover, H.; Ng, S.; Cravatt, B., Monoacyglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* **2010,** *140* (1), 49-61.
36. Rozanov, D.; SIkora, S.; Godzik, A.; Postnova, T.; Golubkob, V.; Savinov, A.; Tomlinson, S.; Strongin, A., Non-proteolytic receptor/ligant interacitons associate cellualr membrane type-1 matrix metalloproteinase with the complement component c1q *The Journal of Biological Chemistry* **2004,** *279* (48), 50321-50328.
37. D'Alessio, S.; Ferrari, G.; Cinnante, K.; Scheerer, W.; Galloway, A.; Roses, D.; Rozanov, D.; Remacle, A.; Oh, E.; Shiryaeb, S.; Strongin, A.; Pintucci, G.; Mignatti, P., Tissue inhibitor of metalloproteinases-2 binding to membrane-type 1 matrix metalloproteinase induces MAPK activation and cell growth by non-proteolytic mechanism *Journal of Biological Chemistry* **2008,** *283* (1), 87-92.
38. Dove, A., MMP inhibitors: glimmers of hope amidst clinical failures *Nature* **2002,** *8* (2), 95.
39. OVerall, C.; Lopez-Otin, C., Strategies for MMp inhibition in cancer: innovations for the post-trial era *Nature Reviews Cancer* **2002,** *2* (9), 657-672.
40. King, J.; Clingan, P.; Morris, D., Randomised double blind placebo control study of adjuvant treatment with the metalloproteinase inhibitor, Marimastat in patients with inoperable colorectal hepatic metastases: significant survival advantage in patients with musculoskeletal side-effects. *Anticancer research* **2003,** *23*, 639-645.
41. Ling, B.; Watt, K.; Barnjee, S.; Newsted, D.; Truesdell, P.; Adams, J.; Sidhu, S.; Craig, A., A novel immunotherapy targeting MMP-14 limits hypoxia, immune suppression and metastasis in tripple-negative breast cancer models. *Oncotarget* **2017,** *8* (35), 58372-58385.

42. Devy, L.; Huang, L.; Naaa, L.; Yanamandra, N.; Pieters, H.; Frans, N.; Chang, E.; Tao, Q.; Vanhove, M.; Lejune, A.; Van Gool, R.; Sexton, D.; Kuang, G.; Rank, D.; Hogan, S.; Pazmany, C.; May, Y.; Schoonbroot, S.; Nixon, A.; Landner, R.; Hoet, R.; Hendrikx, P.; Tenhoor, C.; Rabbani, S.; Valentino, M.; Wood, C.; Dransfield, D., Selective inhibition of matrix metalloprotienase-14 blocks tumor growth, invasion, and angiogenesis *Canser Research* **1009,** *69* (4), 1517-1526.
43. Barry-Hamilton, V.; Spangler, R.; Marshall, D.; SMcCauley, S.; Rodriguez, H.; Oyasu, M.; Mikels, A.; Vaysberg, M.; Gjermazien, H.; Wai, C.; Garcia, C.; Velayo, A.; Jorgensen, b.; Biermann, D.; Tsai, D.; Green, J.; Zaffrayer-Eliot, S.; Holzer, A.; Ogg, S.; Thai, D.; Neufeld, G.; Van Vlasselaer, P.; Smith, V., Allosteric inhibiton of lysyl oxidase-like-2 impedes the development of a pathologic microenviornment *Nature Medicine* **2010,** *16* (9), 1009-1017
44. Nicholson, S.; Wood, C.; Devy, L. Use of MMP-9 and MMP-12 binding protiens for the treatment and prevention opf systemic sclerosis. April 22, 2010, 2010.
45. Appleby, T.; Hung, M.; Liclican, A.; Velasquez, M.; Villasenor, A.; Wang, R.; Wong, M.; Liu, X.; Papalia, G.; Schultz, B.; Sakowicz, R.; Smith, V.; Kwon, H., Biochemical characterization and structure determination of a potent, selective antibody inhibitor of human MMP9. *Journal of Biological Chemistry* **Greenstein, AE,** *292* (16), 6810-6820.
46. Nam, D.; Ge, X., Development of a periplasmic FRET screening method for protease inhibitory antibodies. *Biotechnol Prog.* **2013** *110* (11), 2856-2864.
47. Li, J.; Xla, L.; Su, Y.; Liu, H.; Xia, X.; Lu, Q.; Yang, C.; Reheman, K., Molecular imprint of enzyme active site by camel nanobodies *The Journal of Biological Chemistry* **2012,** *287* (17), 13713-13721.
48. Nam, D.; Fang, K.; Rodriguez, C.; Lopez, T.; Ge, X., Generation of inhibitory monoclonal antibodies targeting matrix metalloproteinase-14 by motif grafting and CDR optimization *Protein Enigneering Design and Selection* **2017,** *30* (2), 113-118.
49. Devy, L.; Dransfield, D., New strategies for the next generation of matrix-metalloproteinase inhibitors: selectively targeting membrane-anchored MMPs with therapeutic antibodies. *Biochemistry Research International* **2011**.
50. Nam, D.; Rodriguez, C.; Remacle, A.; Strongin, A.; Ge, X., Active-site MMP-selective antibody inhibitors discovered from convex paratope syntehtic libraries *PNAS* **2016,** *113* (52), 14970-14975.
51. Arabi-Ghahroudi, M., Cameldi single-domain antibodies: historical perspective and future outlook. *Fronteirs in Immunology* **2017,** *8* (1589).
52. Nagase, H.; Visse, R.; Murphy, G., Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular Research* **2006,** *69* (3), 562-573.
53. Sharbi, O.; Shirian, J.; Grossman, M.; Ledendiker, M.; Sagi, I.; Shifman, J., Affinity and specificity enhancing mutations are frequent in multispecific interactions between TIMP2 and MMPs *PLOS One* **2014** *9*(4).
54. Arkadash, V.; Yosef, G.; Shirian, J.; Cohen, I.; Horev, Y.; Grossman, M.; Sagi, I.; Radisky, J.; Papo, N., Development of high affinity and high specificity inhibitors of matrix metalloproteinase 14 through computaitonal design and directed evolution *The Journal of Biological Chemistry* **2017,** *292* (8), 3481-3495.
55. Gossage, D.; Cieslarova, B.; AP, S.; Zheng, H.; Xin, Y.; Lal, P.; Chen, G.; Smith, V.; Sundy, J., Phase 1b Study of the Safety, Pharmacokinetics, and Disease-related Outcomes of the Matrix Metalloproteinase-9 Inhibitor Andexcaliximab in Patients With Rheumatoid Arthritis. *Therapeutics* **2018,** *40* (1), 156-165.
56. Fleishman, S.; Whitehead, T.; Ekiert, D.; Dreyfus, C.; Corn, J.; Strauch, E.; Wilson, I.; Baker, D., Computational design of protiens targeting the conserved stem region of influenza hemagglutinin. *Science* **2011,** *332* (6031), 816-821.
57. Silwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E., Computational Methods in Drug Discovery. *Pharmacological Reviews* **2014,** *66* (1), 334-395.
58. A, C.; al., e., Massively parallel de novo protein design for targeted therapeutics. *Nature* **2017,** *550* (7674), 74-79.

59. Huang, P.; Boyken, S.; Baker, D., The coming of age of de novo protein design *Nature* **2016,** *537* (7620), 320-327.
60. Brekke, O.; Sadlie, I., Therapeutic antibodies for human diseases at the dawn of the twenty first century *Nature Reviews Drug Discovery* **2002,** *2* (1), 52-56.
61. Adams, J.; Sidhu, S., Synthetic antibody technologies *Current Opinion in Structural Biology* **2014,** *24*, 1-9.
62. Sidhu, S.; Fellouse, F., Synthetic therapeutic antibodies *Nature Chemical Biology* **2006,** *2* (12), 682-688.
63. Liposvek, D.; Mena, M.; Lippow, S.; Basu, S.; Baynes, B., Library Construction for Protein Engineering. In *Protein Engineering and Design*, Park, S.; Cochran, J., Eds. CRC Press: Boca Raton, FL, 2010; pp 83-102.
64. Mahon, C.; Lambert, M.; Glanville, J.; Wade, J.; Fennell, B.; Krebs, M.; Armellino, D.; Yang, S.; Liu, X.; O'Sullivan, C.; Autin, B.; Oficjalska, K.; Bloom, L.; Paulsen, J.; Gill, D.; Damelin, M.; Cunningham, O.; Finlay, W., Comprehensive interrogation of a minimalist synthetic CDR-H3 library and its ability to generate antibodies with therapeutic potential. *Journal of Molecular Biology* **2013,** *425* (10), 1712-1730.
65. Vaughan, T.; Williams, A.; Pritchard, K.; Osbourn, J.; Pope, A.; Earnshaw, J.; McCafferty, J.; Hodits, R.; Wilton, J.; Johnson, K., Human antibodies with sub-nanomolar affinity isolated from a large non-immunized phage display library *Nature Biotechnology* **1995,** *14* (3), 309-314.
66. Feldhaus, M.; Siegel, R.; Opresko, L.; Coleman, J.; Weaver, J.; Yeung, Y.; Cochran, J.; Heinzelman, P.; Colby, D.; Swers, J.; Graff, C.; WIley, S.; Wittrup, K., Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces Cerevisiae surface display library *Nature Biotechnology* **2003,** *21* (2), 163-170.
67. Willuda, J.; Honeggar, A.; Waibel, R.; Schubiger, A.; Stahekl, R.; Zangemeister-Witte, U.; Pluckthun, A., High thermal stability is essential for tumor targeting antibody fragments *Cancer Research* **1999,** *59* (22).
68. Ewert, S.; Huber, T.; Honegger, A.; Pluckthun, A., Biophysical properties of human antibody varriable domains *Journal of Molecular Biology* **2003,** *325* (3), 531-553.
69. Worn, A.; Pluckthun, A., Identification, classification and improvement by protein engineering *Biochemistry* **1999,** *38* (27), 8739-8750.
70. Persson, H.; Ye, W.; Wernimount, A.; Adams, J.; Koide, A.; Koide, S.; Lam, R.; Sidhu, S., CDR-H3 diversity is not required for antigen recognition by synthetic antibodies *Journal of Molecular Biology* **2012** *435* (4), 803-811.
71. Miersch, S.; Maruthachalam, B.; Geyer, R.; Sidhu, S., Structure-directed ant tailored diversity synthetic antibdoy libraries yield novel anti-EGFR antagonists *ACS chemical biology* **2017,** *12* (5), 1381-1389.
72. Koerber, J.; Thomsen, N.; Hannigan, B.; Degrado, W.; Wells, J., Nature-inspired design of motif-specific antibody scaffolds *Narure Biotechnolgy* **2013,** *31* (10), 916-921.
73. Bowley, D.; Labrijin, A.; Zwick, M.; Burton, D., Antigen selection from an HIV-1 immune antibody library displayed on yeast yeilds many novel antibodies compared to selection from the same library displayed on phage *Protein Engineering Design and Selection* **2007,** *20* (2), 81-90.
74. Bradbury, A.; Sidhu, S.; Dubel, S.; McCafferty, J., Beyond natural antibodies: the power of in vitro display technologies *Nature Biotechnology* **2011,** *19* (3), 245-254.
75. McCafferty, J.; Griffiths, A.; Winter, G.; Chiswell, D., Phage antibodies: filamentous phage displaying antibody varriable domains. *Nature* **1990,** *348* (6301), 552-554.
76. Hanes, J.; Pluckthun, A., In vitro selection and evolution of functional proteins by using ribosome display *PNAS* **1997,** *94* (10), 4937-4942.
77. Boder, E.; Wittrup, K., Yeast surface display for screening combinatorial polypeptide libraries *Nature Biotechnology* **1997,** *15* (6), 553-557.

78. Ho, M.; Pastan, I., Mammalian Cell Display for Antibody Engineering In *Therapeutic Antibodies Methods and Protocols q*, Dimitrov, A. S., Ed. Springer New York, New York, 2009; Vol. 525, pp 337-352.
79. Angelini, A.; Chen, T.; de Picciotto, S.; Yang, N.; Tzeng, A.; Santos, M.; Van Deventer, J.; Traxlmayr, M.; Wittrup, K., Protein engineering and selection using yeast display In *Yeast Surface Display* 2015; Vol. 1319, pp 3-36.
80. Van Deventer, J.; Le, D.; Zhao, J.; Kehoe, H.; Kelly, R., A platform for constructing, evaluating and screening bioconjugates on the yeast surface *Protein Engineering Design and Selection* **2016,** *29* (11), 485-493.
81. Van Deventer, J.; Kelly, R.; Rajan, S.; Wittrup, K.; Sidhu, S., A switchable yeast dispaly/secretion system *Protein Engineering Design and Selection* **2015,** *28* (10), 317-325.
82. Zorniak, M.; Clark, P.; Umlauf, B.; Cho, Y.; Shista, E.; Kuo, J., Yeast display biopanning identifies human antibodies targeting glioblastoma stem-like cells *Scientific Reports* **2017,** *7* (1).
83. Aziuteu, M.; Correia, B.; Ban, A.; Carrico, C.; Kalyuzhiny, O.; Chen, L.; Schroeter, A.; Huang, P.; McLellan, J.; Kwong, P.; Baker, D.; Strong, R.; Scheif, W., Computation-Guided Backbone Grafting of a Discontinous Moteif onto a Protein Scaffold *Science* **2011,** *334* (6054), 373-376.
84. Moroncini, G.; Kanu, N.; Solforosi, L.; Abalos, G.; Telling, G.; Head, M.; Ironside, J.; Brockes, J.; Burton, D.; Williamson, A., Motof-grafted antibodies containing the replicative interface of cellular PrP are specific for PrPSC *PNSA* **2004,** *101* (28), 10404-10409.
85. Horiya, S.; Bailey, J.; Temme, S.; Guillen, Y.; Krauss, I., Directed evolution of multivalent glycopeptides tightly recognized by HIV antibody 2G12. *Journal of the Amrican Chemical Society* **2015,** *136* (14), 5407-5415.
86. Yang, A.; ha, S.; Ahn, J.; Kim, R.; Kim, S.; Lee, Y.; Kim, J.; Soll, D.; Lee, H.; Park, H., A chemical biology route to site-specific authentic protein modifiecations *Science* **2016,** *354* (6312), 623-625.
87. JY, A.; Bajjuri, K.; Ritland, M.; Hutchins, B.; Kim, C.; Kazane, S.; Hadler, R.; Forsyth, J.; Santidrian, A.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A.; Broc, S.; Szydlik, A.; Pinkstaff, J.; Tian, F.; Sinha, S.; Fedling-Habermann, B.; Smider, V.; Schultz, P., Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *PNAS* **2012** *109* (40), 16101-16106.
88. Flentke, G.; NMunoz, E.; Huber, B.; Plaut, A.; Kettner, C.; Bachovchin, W., Inhibition of dipeptidyl aminopeptidase IV (DP-IV) by Xaa-boroPro dipeptides and use of these inhibitors to examine the role of DP-IV in T-cell function. *PNAS* **1991,** *88* (4), 1556-1559.
89. Meyer, S.; Shomin, C.; Gaj, T.; Gosh, I., Teathering small molecules to a phage dispaly library: discovery of a selective bivalent inhibitor of protien kinase A *Journal of the Amrican Chemical Society* **2007,** *129* (45), 13812-13813.
90. Holliger, P.; Hudson, P., Engineered antibody fragments and the rise of single domains *Nature Biotechnology* **2005,** *23* (9), 1126-1136.
91. Tonegawa, S., Somatic generation of antibody diversity *Nature* **1983,** *302* (5909), 1079-1087.
92. Xu, J.; Davis, M., Diversity in the CDR3 region of VH is sufficient for most antibody specificities *Immunity* **200,** *13* (1), 37-45.
93. Kunik, V.; Ofran, Y., The indistinguishability of epitopes from protein surface is explained by the distinct binding preferences of each of the six antigen-binding loops. *Protein Engineering Design and Selection* **2013,** *26* (10), 599-609.
94. Kwong, P.; Wyatt, R.; Robinson, J.; Sweet, R.; Sordroski, J.; Hendrickson, W., Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **1998,** *393* (6686), 648-659.
95. Huh, J.; White, A.; Brych, S.; Franey, H.; Matsumura, M., The identification of free cystine residues within antibodies and a potential role for free cystine residues in covalent aggregation because of agitation stress *Journal of Pharmocological Sciences* **2013,** *102* (6).
96. Kelly, R.; Le, D.; Zhao, J.; Wittrup, K., Reduction of nonspecificity motifsi n synthetic antibody libraries *Journal of Molecular Biology* **2018,** *430* (1), 119-130.

97. Fellouse, F. L., B; Compaan, D.; Peden, A.; Hymowitz, S.; Sidhu, S., Molecular recognition by a binary code *Journal of Molecular Biology* **2005,** *348* (5), 1153-1162.
98. Pierce Biotechnology Instructions: EZ-Link NHS-biotin reagents. https://www.thermofisher.com/order/catalog/product/21336 (accessed April 23, 2018).
99. Thermo Fisher Scientific Zeba Desalting Products https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-purification-isolation/protein-dialysis-desalting-concentration/zeba-desalting-products.html.
100. Pierce Biotechnology Instructions: Pierce Biotin Quantification Kit. https://www.thermofisher.com/order/catalog/product/28005.
101. Thermo Fisher Scientific HABA Calculator https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-labeling-crosslinking/protein-labeling/biotinylation/biotin-quantitation-kits/haba-calculator.html.
102. Zymo Research Zymoprep Yeast Plasmid Miniprep https://www.zymoresearch.com/zymoprep.
103. Kay, B.; Thai, S.; Volgina, V., High-throughput biotinylation of protiens In *High throughput protein expression and purification* 2011; Vol. 498, pp 185-198.
104. Fairhead, M.; Howarth, M., Site-specific biotinylation of purified proteins using BirA In *Site-Specific Protein Labeling*, 2014; Vol. 1266, pp 171-184.
105. Azim-Zadeh, O.; Hillebrecht, A.; Linne, U.; Marahiel, M.; Klebe, G.; Lingebach, K.; Nyalwidhe, J., Use of biotin derivatives to probe conformational changes in proteins *Journal of Biological Chemistry* **2007,** *282* (30), 21609-21617.
106. Vermeer, A.; Norde, W., The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein. *Biophysical Journal* **2000,** *78* (1), 394-404.
107. de Haard, H.; Van Neer, N.; Reurs, A.; Hufton, S.; Roovers, R.; Henderikx, P.; De Bruine, A.; Arends, J.; Hoogenboom, H., A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *Journal of Biological Chemistry* **1999,** *274* (26), 18218-18230.
108. Abbas, A.; Lichtman, A.; Pillai, S., *Cellular and Molecualr Immunology* 2018.
109. Wang, F.; Eldert, D.; Ahmad, I.; Yu, W.; Zhang, Y.; Bazugan, O.; Torkamani, A.; Raudsepp, T.; Mwangi, W.; Criscifello, M.; Wilson, I.; Schultz, P.; Smider, V., Reshaping antibody diversity *Cell* **2013,** *153* (6), 1379-1393.
110. Fanning, S.; Walter, R.; Horn, R., Structural basis of an engineered dual-specific antibody: conformational diversity leads to a hypervariable loop metal-binding site. *Protien engineering design and selection* **2014,** *27* (10), 391-397.
111. Yan, J.; Li, G.; Hu, Y.; Wan, Y., Construction of a synthetic phage-dispalyed nanobody library with CDR3 regions randomized by trinucleotide cassettes for diagnostic applications *Journal of Translational Medicine* **2014,** *12* (343).
112. Gram, H.; Marconi, A.; Barbas, C.; Collet, T.; Lerner, R.; Kang, A., In vitro selection and affinity maturation of antibodies from a naive combinatorial immunoglobin library *PNAS* **1992,** *89* (8), 3576-3580.
113. Fantini, M.; Pandolfini, L.; Lisi, S.; Chirichella, M.; Arisi, I.; Terrigno, M.; Goracci, M.; Cremisi, F.; Cattaneo, A., Assessment of antibody library diversity through next generation sequewncing and tehnical error compensation. *PLOS One* **2017,** *12* (5).
114. Gerorgiou, G.; Ippolito, G.; Beausang, J.; Busse, C.; Wardemann, H.; QUake, S., The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nature Biotechnology* **2014,** *32* (2), 158-168.
115. Yang, W.; Yoon, A.; Lee, S.; Kim, S.; Han, J.; Chung, J., Next-genreation sequencing enables the discovery of more diverse postitive clones from a phage-displayed antibody library *Experimental and Molecular Medicine* **2017,** *49* (308).
116. Bogan, A.; Thron, K., Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* **1998,** *280* (1), 1-9.

117. Moreira, I.; Fernandes, P.; Ramos, M., Hot spots: A review of the protein-protein interface determinant amino-acid residues *Protiens* **2006,** *68* (4), 803-8112.
118. Peng, H.; Lee, K.; Jian, J.; Yang, A., Origins of specificty and affinity in antibody-protein interactions *PNAS* **2014,** *111* (26), 2656-2665.
119. Li, X.; Keskin, O.; Ma, B.; Nussinov, R.; Liang, J., Protien-protien interactions: Hot spots and structurally consered residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking *Journal of Molecular Biology* **2004,** *344* (3), 781-795.
120. Xu, Y.; Roach, W.; Sun, T.; Jain, T.; Prinz, B.; Yu, T.-Y.; Torrey, J.; Thomas, J.; Bobrowicz, P.; WVasquez, M.; Wittrup, K.; Krauland, E., Addressing polyspecificty of antibodies selected from an in vitro yeast presentation system: a FACS-bbased high-throughput selection and analytical tool *Protein Engineering Design and Selection* **2013,** *26* (10), 663-670.
121. Hanes, J.; Schaffitzel, C.; Knappik, A.; A, P., Picomolar affinity antibodies from a fully sunthetic naive library selected and evolved by riboosme display. *Natrue Biotechnology* **2000,** *18* (12), 1287-1292.
122. Knappik, A.; Ge, L.; Honegger, A.; Pack, P.; Fischher, M.; Wellnhofer, G.; Hoess, A.; Wolle, J.; Pluckthun, A.; Virnekas, B., Fully synthetic human combinatorial antibody libraries based on modular consensus frameworks and CDRs randomized with trinucleotides. *Journal of Molecular Biology* **2000,** *296* (1), 57-86.
123. Chen, L.; Kwon, Y.; Wu, X.; O'Dell, S.; Cavacini, L.; Hessell, A.; Pancera, M.; Tang, M.; Xu, L.; UYang, Z.; Zhang, M.; Arthos, J.; Burton, D.; Dimitov, D.; Naabel, G.; Posner, M.; Sordoski, J.; Wyatt, R.; Mascola, J.; Kwong, P., Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science* **2009,** *326* (5956), 1123-1127.
124. Tominson, I.; Walter, G.; Marks, J.; Llewelyn, M.; Winter, G., The repertoire of human germline VH sequences reveals about fifty groups of VH segments with different hypervariable loops *Journal of Molecular Biology* **1992,** *227* (3), 776-798.

Conclusions and Future Work

5.1: Conclusions

The overall goal of this thesis work was to assess whether our yeast displayed CDRH3 scFv library could serve as a tool in engineering protein-small molecule hybrids. We accomplished this goal by validating the baseline scFv library itself, which will eventually act as a scaffold into which additional small molecule diversity can be incorporated. We validated the library by screening it against a panel of model antigens, including IgGs and enzymes. The resulting IgG binding clones that we isolated had triple-digit nanomolar affinities for their targets. While the affinity of isolated clones was relatively low, it is important to note that eventual synergies with incorporated small molecules will enable high bivalent affinities to be achieved. Moreover, isolated clones showed little to no off-target binding with other IgGs. The fact that the library was able to yield binders capable of discriminating between such structurally similar antigens suggests that the library could eventually be used to produce binders capable of discriminating between individual enzyme of multiple-enzyme families, such as the MMPs.

There were some potential pitfalls of the CDRH3 library that surfaced during the validation. The main pitfall was that the library did not yield binders to enzymatic targets in screens performed here. We hypothesized that this was due to CDRH3 loop lengths being too short and therefor being hindered by the other antibody loops from forming binding interactions with the enzyme surface. While it would be simple to add additional CDRH3 loop diversity to the library, adding the small molecule diversity layer to the library will likely correct for this problem. The affinity of the small molecule for the enzyme active site may be able to pull the antibody close enough to the enzyme surface, such that the CDRH3 can form binding interactions.

In conclusion, the library described here and variations of it will be able to be used as a tool for the construction of protein-small molecule hybrids. The eventual multi-component hybrid library can be used to provide insights on the optimal positioning of small molecules and which small molecules to use in hybrid construction. The resulting hybrids would have the potential to be used both as therapeutics and as a molecular biology probes to evaluate the role of different proteases n the tumor microenvironment. In addition to protein-small molecule hybrids, this library could have broader applications in motif grafting. The fact that all CDRs besides one are held at the germline sequence would offer an increase in positions that can be explored for the grafting of peptide motifs.

5.2: Future Work

It is important for the CDRH3 library to be able to yield binders to different epitopes on a given antigen. Having clones that bind to multiple epitopes increases the chances that an scFv can be isolated that will bind close enough to the active site to be useful as a hybrid. We can analyze the epitope coverage of isolated clones by producing soluble antibody of the clones, using it to block epitopes and then observe how the binding of a different yeast displayed scFv against the same target changes using flow cytometry. The production of soluble antibody required for these experiments in it of itself can be used to verify that the library yields soluble and thermodynamically stable antibodies. This is important if the inhibitors are to be used as potential therapeutics. Soluble antibody can be secreted by the yeast display system with only minimal sub-cloning required. Finally, a proof of concept for the hybrids is still needed that shows that incorporating a small molecule to an scFv isolated from the library does not diminish binding. When these experiments are complete, there will be a more in-depth understanding of whether the library is viable for its intended purpose.

Appendix

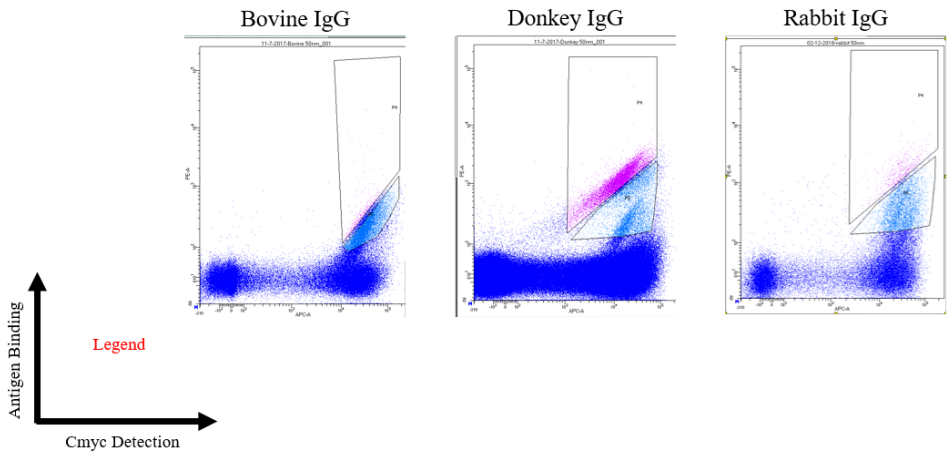

Figure A.1. Screen shots taken during FACS sorting for Bovine IgG, Donkey IgG and Rabbit IgG. FACS was done at the Tufts University Sackler School of Medicine using a BD FACS Aria cell sorter. Shown is sort round 5 in progress, so flow cytometry should look the same as the post-round 4 data. All cells were labeled with 50nM target IgG and anti-biotin PE secondary label. Cells were also labeled with chicken anti-Cmyc and the associated secondary label. Shown are two gates drawn: top and bottom. Gating was done to separate stronger and weaker clones based off the relative intensity of the antigen binding signal. 10,000 events collected from each gate.

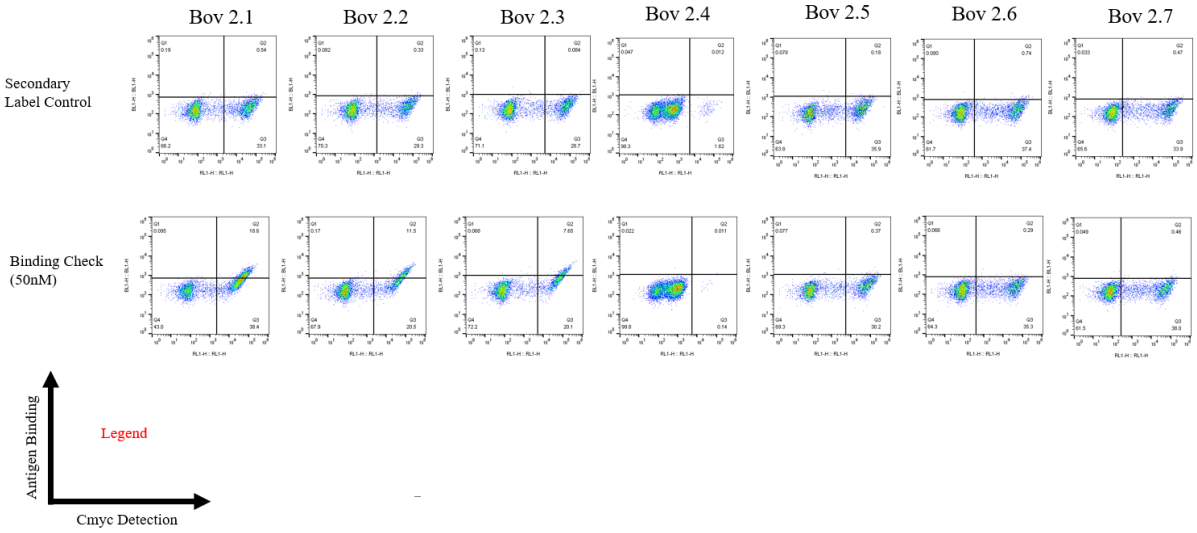

Figure A.2. 50nM Bovine IgG binding experiment for all sequences in Table A.1. Results shown as dot plot. All cells in the binding check condition were labeled with 50nM bovine IgG and a streptavidin 488 secondary label. Cells in the secondary control condition were only labeled with the streptavidin 488 secondary label. Cells in all conditions were labeled with Cmyc. Clones 1-3 have cells in quadrant 2 in the antigen binding condition but not in the control, and therefor are antigen binding. Clones 4-7 have no cells in quadrant 3 in either condition and are therefore not antigen binding and the result of carry over during screening. Clone 4 only has cells in quadrant 4 with minimal cells in quadrant 3 for both conditions. This makes sense considering this clone is truncated, meaning it did not have a full-length scFv on its surface and the Cmyc tag would therefore not be read through. The small number of cells in quadrant 3 are due to non-specific binding of the anti-Cmyc and associated secondary label to the yeast surface.

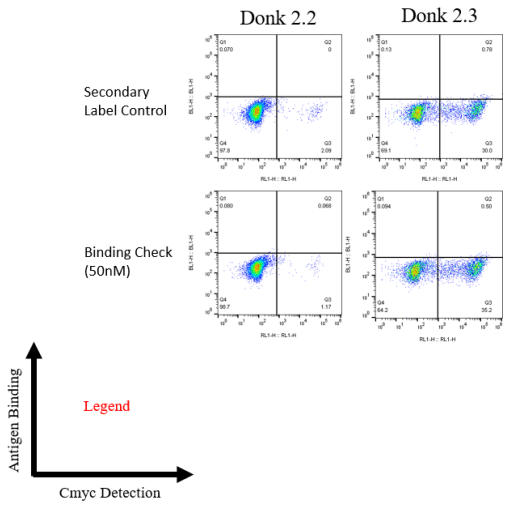

Figure A.3. 50nM Donkey IgG binding experiment for all sequences in Table A.1. Results shown as dot plot. All cells in the binding check condition were labeled with 50nM donkey IgG and a streptavidin 488 secondary label. Cells in the secondary control condition were only labeled with the streptavidin 488 secondary label. Cells in all conditions were labeled with Cmyc. No clones have cells in quadrant 2 for any condition, and therefore none are antigen binding. Clone 2 only has cells in quadrant 4 with minimal cells in quadrant 3 for both conditions. This makes sense considering this clone is truncated, meaning it did not have a full-length scFv on its surface and the Cmyc tag would therefore not be read through. The small number of cells in quadrant 3 are due to non-specific binding of the anti-Cmyc and associated secondary label to the yeast surface. Clone 1 is not shown here, because it had already been verified to be a genuine antigen binding clone as shown in Chapter 3.

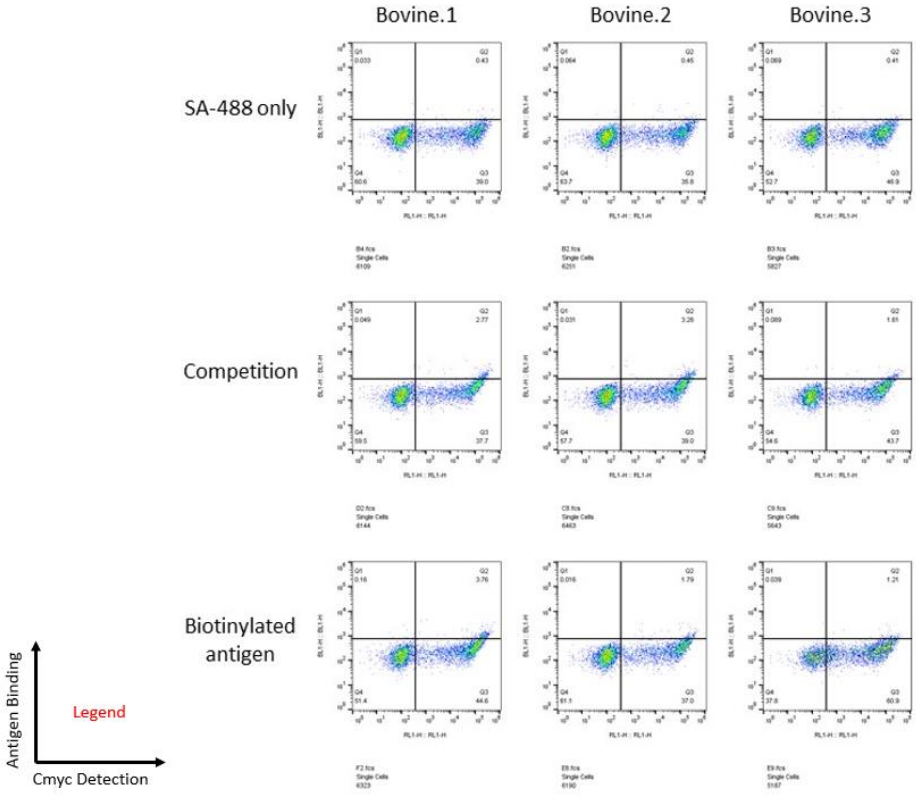

Figure A.4. Bovine IgG competition experiment with 50nM biotinylated antigen and 1uM non-biotinylated antigen for all bovine IgG binding clones isolated during the first round of CDRH3 library screening. All antigen was the same stock used in library screening. All clones have a comparable binding population size for the competition and biotin control conditions. This indicates that isolated clones bound preferentially to biotinylated antigen. Data collected by Haixing Kehoe.

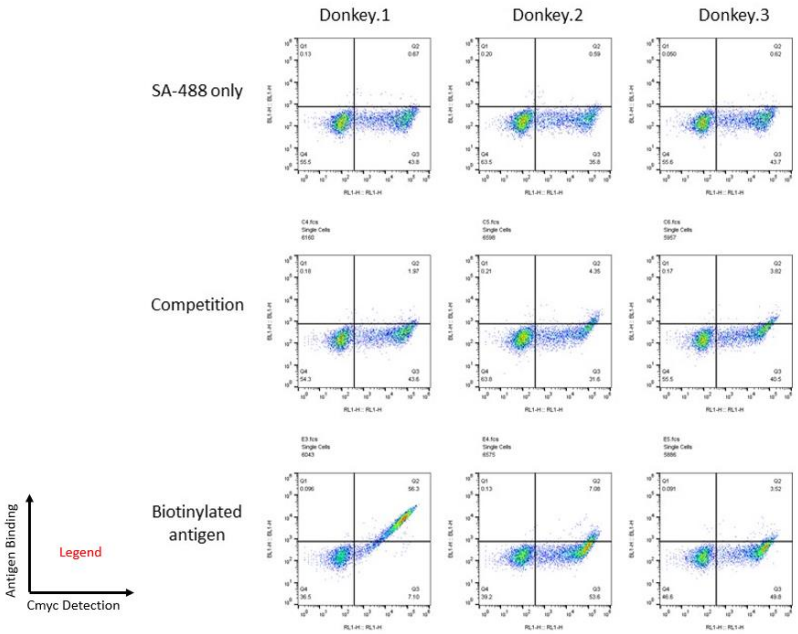

Figure A.4. Donkey IgG competition experiment with 50nM biotinylated antigen and 1uM non-biotinylated antigen for all donkey IgG binding clones isolated during the first round of CDRH3 library screening. All antigen was the same stock used in library screening. Donkey 2 and 3 clones have a comparable binding population size for the competition and biotin control conditions. This indicates that isolated clones bound preferentially to biotinylated antigen. Donkey 1 has a very small binding population for the competition condition, indicating this clone preferentially binds to an epitope on the antigen itself. Donkey 2 clone was shown in a subsequent repeat experiment to preferentially recognize the antigen. Data collected by Haixing Kehoe.

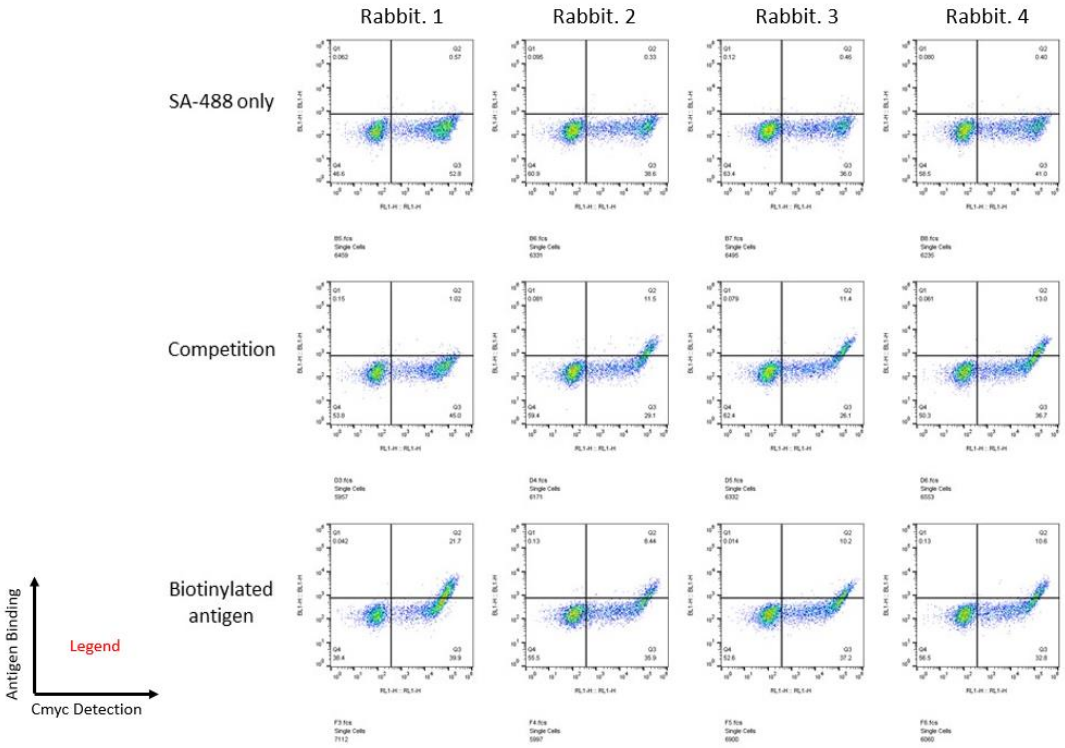

Figure A.5. Rabbit IgG competition experiment with 50nM biotinylated antigen and 1uM non-biotinylated antigen for all rabbit IgG binding clones isolated during the first round of CDRH3 library screening. All antigen was the same stock used in library screening. Rabbit 2-5 clones have a comparable binding population size for the competition and biotin control conditions. This indicates that isolated clones bound preferentially to biotinylated antigen. Rabbit 1 has a very small binding population for the competition condition, indicating this clone preferentially binds to an epitope on the antigen itself. Data collected by Haixing Kehoe.

A

Clone	CDRH3 Sequence	Loop Length
Bov 2.1	HYDPYDDYNFYFGFDY	16
Bov 2.2	TYYDNSYAYTALDY	14
Bov 2.3	SSIYDGYMYGLDY	13
Bov-NB-2.4	H*NNDKRPAIDY	1
Bov-NB-2.5	AFGDNANPNNAIDY	14
Bov-NB-2.6	YDVATNURALDY	12
Bov-NB-2.7	AKNHSFGFDY	10

B

Clone	CDRH3 Sequence	Loop Length
Donk 2.1	YDKTHHNPDYALDY	14
Donk-NB-2.2	*NAYSPAALDY	0
Donk-NB-2.3	AHCHHGLDY	9

Table A.1. Comprehensive sequencing results for CDRH3 library sort 2. A) All unique sequences from Bovine IgG top and bottom FACS populations. B) All unique sequences from Donkey IgG top population. Clones denoted with NB were determined not to bind to target antigen in binding assays. Asterix denotes a stop codon in the CDRH3, indicating a truncated clone. All CDRH3 loop lengths reported according to the Kabat scheme.

Clone	CDRH3 Sequence	Loop Length	K_D in nM (95% CI)
Bov 1.1	YLDHYYALDY	10	
Bov 1.2	YYDHHYYALDY	11	
Bov 1.3	YLDNHYYALDY	11	
Donk 1.1	YDKTHHNPDYALDY	14	4.012 (1.635-10.091)
Donk 1.2	YENSYALDY	9	421.8 (317.8-570.1)
Donk 1.3	YYDPDRYNYAMDY	13	
Rab 1.1	YNYHHPFSYDAFDY	14	1168 (1124-1215)
Rab 1.2	YYHTDYTAHNAYALDY	16	
Rab 1.3	PSHCSSSNNDIDTALD	16	
Rab 1.4	PSHCSSSNDIDTALD	15	
Rab 1.5	ASYPNTFYSIAGID	14	

Table A.2. List of all sequences determined to be binding after the first round of library screening. Data collected by Haxing Kehoe. Sequences denoted in red were later determined to be biotin binders. For all non-biotin binders, dissociation constant is listed in nanomolar along with the associated 95% confidence interval.