The Dissertation Committee for Jeongmin Hyun Certifies that this is the approved version of the following dissertation:

Design and Engineering of Epitope Specific Antibodies

Committee:

Jennifer A. Maynard, Supervisor

David W. Hoffman

Adrian T. Keatinge-Clay

Lauren J. Webb

Jeffrey E. Barrick

Design and Engineering of Epitope Specific Antibodies

by

Jeongmin Hyun, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2016

Dedication

I would like to dedicate this work

To my loving husband, Donghyi Koh

To my family, Moo Kwan Hyun, Hye Kyung Lee and Jungmoon Hyun

Acknowledgements

First, I would like to thank my professor, Dr. Jennifer A. Maynard for encouraging me to pursue every project that I am interested in. With her support, advice, and patient I was able to make this dissertation possible.

I would also like to thank my committee members: Dr. David Hoffman, Dr. Adrian Keatinge-Clay, Dr. Lauren Webb, and Dr. Jeffrey Barrick for their guidance and comments on my dissertation.

I'd like to thank all of my current and former group members for their help and advice: Dr. Jen Pai and Dr. Kevin Entzminger who mentored me when I started my Ph.D. project; Dr. Raquel Lieberman, Dr. Jennifer Johnson, and Sibel Kalyoncu who performed all crystallization works; Dr. Xianzhe Wang for collaboration of single domain antibody project against RTX; Dr. Edith Acquaye for boosting up my confidence; other Maynard group members for their valuable advice and discussion. I also appreciate some undergraduate students, Taegyu Kim and Hyun Jae Lee.

Special thanks goes out to all my friends in Korea and in Austin, Yoomin Han, Sujin Hong, Yerim Yang, Hyunji Lee, Hyokyung Kim, Minhye Shin, Seunghee Cho and Eundeok Kim. With them, I always enjoyed my graduate life.

Last, but most important, much love goes to my family: my husband who supported me in so many ways, my parents who always believed and encouraged me in my choice, my sister who is my best and perfect friend in my life. There is no way I could complete my Ph.D. without their endless support and love.

Design and Engineering of Epitope Specific Antibodies

Jeongmin Hyun, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Jennifer A. Maynard

The knowledge of three-dimensional structures of membrane proteins aids in structure-based drug design, since about 60% of approved drug targets are known as membrane proteins. To date, chaperone-assisted protein co-crystallization that bypass the need for animal immunization is becoming an attractive method to elucidate structures of recalcitrant targets such as proteins with intrinsically disordered domains as found in membrane proteins. Here we describe antibody-engineering strategies for developing crystallization chaperones. Toward this goal, we (1) engineered EE peptide-specific single-chain variable fragment (scFv) to improve biophysical characteristics, (2) constructed synthetic single domain antibody library to be specific for targets by phage display, and (3) *de novo* designed FLAG peptide-specific antibodies using a novel computational method. In the first study, we converted peptide-specific scFv to antigen-binding fragment (Fab), which is the most successful format of antibody-based crystallization chaperones for integral membrane proteins so far. The larger size of Fab/EE increased the overall stability without disruption of binding affinity and extended

crystal contact areas those are favorable characteristics for use as a crystal chaperone. In the second study, a 10^6 synthetic phage display single domain antibody (sdAb) library was constructed and used to identify sdAbs binding the repeat in toxin domain of *B. pertussis* adenylate cyclase toxin (ACT). This protein is an intrinsically disordered calcium binding protein with no homology to any known protein structure and is a candidate vaccine antigen. From phage-based screening, we isolated three sdAbs to be used for further characterization. In the last study, we utilized an *in silico* approach to the design the antibodies using OptCDR that is a general computational method that employs *de novo* design of complementarity determining regions (CDRs) to engineer antibody-antigen interactions. Using this method, we designed CDRs binding the minimal FLAG peptide (sequence: DYKD) and isolated four antibodies with high specificity and nanomolar affinity for the DYKD. The result demonstrates that antibody specificity based on *in silico* design method can guide future engineering of antibody-based crystallization chaperone. Taken together, we have identified antibodies with improved binding properties and biophysical characteristics for using as crystallization chaperones without animal immunization to help guide future antibody chaperone engineering for the structural investigation of diverse target proteins.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Background

1.1 Antibody engineering and its applications

1.1.1 Antibody structures and recombinant antibodies

IgG antibody molecules are roughly Y-shaped molecules composed of four polypeptide chains, approximately 50kDa of two identical heavy chains and 25kDa of two light chains resulting in two identical antigen-binding sites. By disulfide bonds, each heavy chains are connected as well as each heavy chain and a light chain is linked together to maintain the Y-shape (Figure 1.1a). Another important feature of antibody is differentiated by the variability of amino acid sequences. The amino (N)-terminal regions of heavy or light chain are more variable (V) domains and the remaining regions are highly conserved, called constant (C) domain. Each variable regions contains three hypervariable loops with highly variable sequence and length, called complementarity-determining regions (CDRs) where the regions of antigen recognition and interaction. Because each CDRs are connected to the Framework Regions (FR), represented as the mostly conserved sequences and structure, β-sheet. The constant domain of the heavy chain (Fc) is able to interact with cell surface receptor (Fc receptor) and allow to elicit effector functions by activating the immune system.

For traditional antibodies, with proteolytic enzymes (protease) which cleaves the disulfide bonds, full IgG can be dissected with smaller fragments (Figure 1.1a), Fragment antigen binding (Fab) having the antigen-binding activity and Fragment crystallizable (Fc) interacting with effector molecules. The Fab fragments consist of variable light (V_L)

and heavy (V_H) chain and constant light (C_L) and heavy (C_H1) chain. In addition to the conventional antibodies, antibodies derived from camelidae have significantly different features, consist of only two heavy chains (Figure 1.1b). Even the lack of the light chains, the camelidae antibody can bind and behave effectively to their targets. The functional target binding site, called V_HH is similar as conventional V_H domain in their size (12 – 15 kDa) and activity. But in contrast to the V_H, the length of CDR3 region is generally longer resulting to access to epitope in clefts of folded proteins. Moreover, due to its small size, the V_HH has advantages over the V_H in biophysical characteristics such as stability, expression level in microorganisms, or solubility.

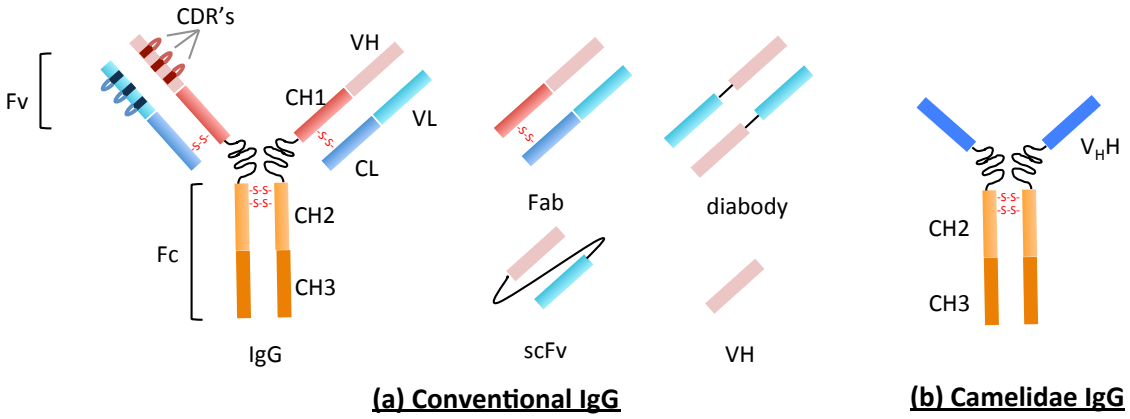


Figure 1.1: Antibody Structure: Conventional and full IgG to smallest VH and Camelidae IgG.

(a) IgG is a homodimer glycoprotein composed of two identical heavy and light chains. The variable regions, represented as Fv have antigen binding activity recognized by CDRs. The Fc regions contain the second and third constant domains serve as effector activation functions. Two heavy chains and each heavy and light chain are linked by disulfide bond. There are several formats of antibody fragments. A Fab fragment consists of the V_H + C_H1 and the V_L + C_L and the smaller scFv is made up the V_H and the V_L connected by peptide linker for flexibility. The smallest format of antibody is Variable only domain (V_H) retaining antigen recognition characteristics. (b) Camelidae IgG is

100kDa of heavy chain only antibodies originated from camels or llamas. Like the traditional antibodies, the N-terminal region of antibody, V_HH, retain antigen recognition activity without an association of the light chain.

In addition to the antibody fragments described above, additional useful antibody formats have been available with recombinant DNA technology. The single chain fragment antibody (scFv) is made of the V_H and V_L, which are connected by flexible glycine-serine linker. The scFv can retain the binding affinity to the targets as whole IgG, but is known to express easier and be produced with higher yield in bacterial system that amenable to handle compared to the IgG.

1.1.2 Antibodies in co-crystallization

To date, antibody-based therapeutics have become a mainstream of antibody engineering and application, and the use of antibodies as research tools is still growing to develop. Antibody fragments as used as crystallization chaperones for structural determination of un-crystallizable protein targets is one way of antibody applications in research [1]. Crystallization chaperone is an auxiliary protein that is known to bind and increase probability of crystallization of the target of interest and monoclonal antibody and its fragments such as Fab, scFv or V_HH are the favorable molecules to be used [2,3]. So far, high value structural biology targets for crystallization are membrane proteins like G-protein coupled receptor (GPCR) [4], intrinsically disordered proteins, which are flexible multi-domain proteins, or hydrophobic proteins (amyloidgenic proteins). Structural determination of these targets are especially important for getting knowledge of the mechanism of protein functions that provide us the insight of vaccine development and drug design. But despite extensive efforts and strategies, it is challenging to crystallize due to their hydrophobic and amphiphilic characteristics. The crystallization chaperone

can aid to assemble crystals by reducing conformational heterogeneity, shielding counterproductive surface and providing extended hydrophilic surfaces (Figure 1.2).

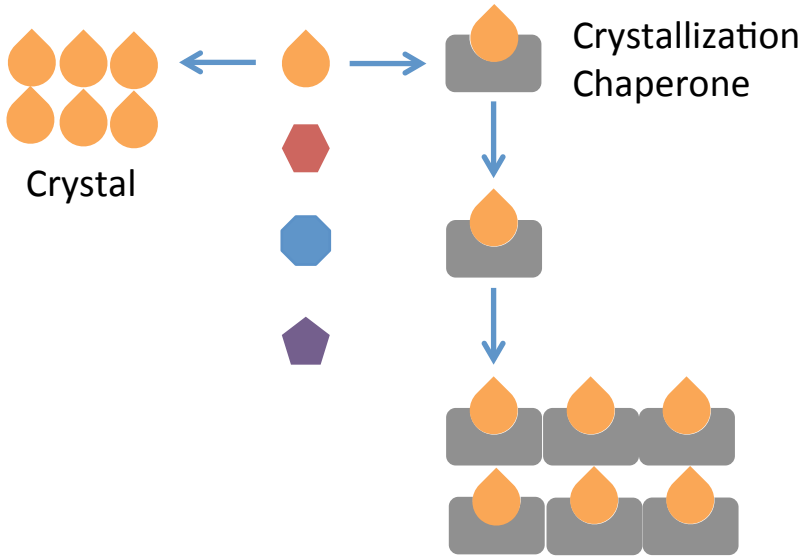


Figure 1.2: The strategy to increase the probability of obtaining crystals.

The ideas to obtain well-ordered crystals of structurally recalcitrant proteins are minimizing the conformational heterogeneity to reduce the conformational entropy for lattice formation. The other strategy is supplying additional hydrophilic surface that can facilitate a stable contacts between molecules in the crystal lattice [2].

The advantages of antibody mediated crystallization over other chaperone molecules; first, using antibody fragment, the size of the chaperone (12–50kDa) can be amenable to match to the specific targets which is important factor to consider for the quality of crystallization [3]. Second, the antibodies have intrinsic capacity for self-assembly due to their inherent β-sheet-rich structure. This conformation can help nucleation and provide additional level of symmetry [2]. Third, antibodies as chaperones can lead to trap proteins in specific conformations meaning reduce heterogeneity or

flexibility [5]. And additionally, co-crystallization with an antibody fragment can be used for providing model-based phasing information [1].

In general, antibody-mediated co-crystallization is to engineer a target protein specific antibody so identification a novel chaperone should be required for each protein of interest. Recent studies, in order to develop a generalizable and ideal format of antibody chaperone, a short peptide or tag binding antibody fragment has been engineered. This recombinant antibody can be applied in co-crystallization for any target proteins if the short peptide is inserted at specific sites, usually in flexible loop that does not interfere any activity or function of the protein of interest [6–8].

1.2 DIFFERENT APPROACHES TO GENERATE ANTIBODIES FOR CHAPERONE

1.2.1 Immune repertoire libraries

To construct antibody libraries, the genetic source of diversity for recombinant antibodies is the natural immune repertoires, generally from B cells of living organisms [9]. There are two types of immune libraries: Biased immune libraries or Naïve libraries. The immune libraries are the most common sources to generate high quality, affinity and antigen specific antibodies because the V_H and V_L chains from B cells are isolated from immunized animals which were repeatedly challenged by target antigen [10]. Because the genetic variety for antigen binding undergoes somatic hypermutation and affinity maturation during prime and boosts [11], there are higher probabilities to isolate high affinity binders. However, even though the number of relevant sequences may be higher than that of non-biased immune libraries, a range of target is limited compared to Naïve

libraries, meaning different targets will require the construction of different libraries. In contrast to the biased immune libraries, the naïve libraries often can be applicable for a broad range of target antigens including toxin or unstable antigens even though the number of chance to isolate specific and high affinity clones is low since the immune system has not been challenged with any targets. Therefore, additional *in vitro* affinity maturation and optimization processes are required to isolate high specific clones from the un-stimulated libraries. After isolation and selection of candidate clones, pooling and cloning the variable gene are easily accessible because the flanking regions of variable gene sequences are conserved [12]. But despite their many advantages described above, the shortcomings such as antibody stability, expression, and diversity have led to alternative sources of library systems.

1.2.2 Synthetic libraries by random mutagenesis or computational design

Synthetic antibody libraries have been successfully developed to generate stable antibody fragments with high affinity and specificity for various antigens without animal immunization. Generally, the synthetic antibody libraries can be constructed by introducing diversity into CDRs of well-behaved antibody frameworks, meaning highly stable and well-expressed scaffolds and resulting in isolation of more functional clones. The main advantage of synthetic libraries is that they are suitable for the broad range of antigens as Naïve libraries. Also, using a human scaffold for the framework can be resulted in avoiding human anti-mouse antibody response [13]. All possible codons or limited variations can be applied diversity in every, or specific position of CDRs. One

disadvantage of the synthetic libraries is it can be possible to isolate misfolded or aggregated-prone antibodies due to unusual combination of amino acids. Therefore, in order to prevent this limitation, it is important to understand the structure and function of scaffold antibodies. Structural information implies more details about antigen/antibody binding interface, side-chain of amino acids in CDRs that is involved in binding so it can be used to determine how to introduce mutation and tailor the specific sites in the CDRs [14–16]. Computational approaches are attractive approaches to create the synthetic libraries. The approaches can be used for improving biophysical properties of existing antibodies [17,18] or predicting antibody structure based on their amino acid sequences [16,19]. And even more, the computational method is useful for *de novo* designing of antibodies to acquire binding affinity and specificity to new targets [20]. The method is known as Optimal Complementarity Determining Regions (OptCDR) and the main idea of this approach is selecting amino acid sequences in CDR regions based on the canonical structures that are favorable to bind the target antigens (Figure 1.3). Therefore, the computational approaches are one of feasible methods to improve the overall quality of synthetic antibody libraries.

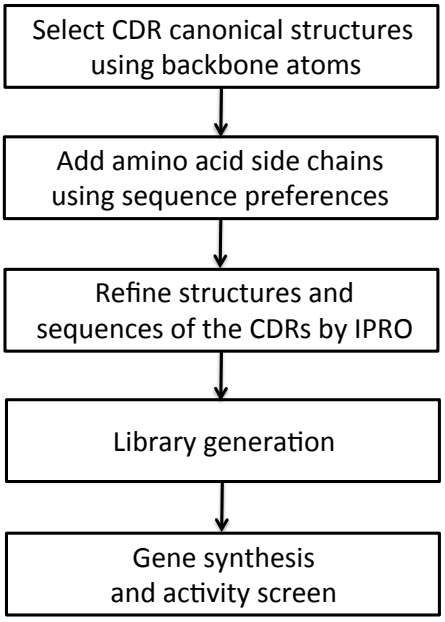


Figure 1.3: Antibody design by CDR modeling.

OptCDR is a general computational method to design the CDRs of antibodies to have high specificity and affinity to target [20]. Step 1: CDR canonical structure identification and selection. Step 2: Initialization of CDR amino acid sequences of the lowest-energy combination set. Step 3: Structure and sequence optimization with an iterative computational protein library redesign and optimization procedure (IPRO). Step 4: Antibody library generation to isolate novel antibodies to bind any specified antigen.

1.2.3 Antibody library display systems and selection methods

To create and screen constructed antibody libraries, some critical factors such as library's diversity and size should be considered to maximize a number of functional clones. Even the large size of libraries is available, maximum ~10^{10} total variants can be screened with high throughput selection methods [21]. In addition to balancing the library's diversity and size, affinity maturation and biophysically optimization of antibody libraries are also crucial to success the antibody engineering. By continuously exposing to the antigens, higher binding variants are selectively dominated and isolated.

Recently, along with the production of antibody libraries, a number of powerful antibody display systems have been developed to select high affinity antibodies through linking genotype to phenotype against various targets [22]. In contrast to the conventional animal immunization followed by hybridoma technology, the molecular display platforms, including phage display [23,24], bacteria [25] or yeast [26] cell based display, and cell-free translational based system [27], led to easily accessible to efficiently enriched isolated antibodies by displaying highly various and large size of antibody libraries. Of them, the phage display platform is the most frequently used system in molecular biology labs for biased and naïve immune libraries and synthetic libraries [28,29] due to its readily handling and managing even the loss of functions of coated antibody can be an issue. Expressed antibodies were displayed on the surface of phage M13 as a fusion with phage surface protein, pIII. Antibody library genes were encoded in phagemid vector and the phagemid DNA is transformed into the appropriate host *E.coli*, resulting in a large antibody library which can be efficiently generated and be a good source of further application [30]. The cells transformed with phagemid are infected with M13 helper phage to acquire all of the genes required for replication and packaging into the phage virion, followed by being induced antibody-pIII fusions. After construction of antibody libraries in phagemid vector, strategies to select and screen is essential to enrich desired antibodies. This biopanning method has been well established as follows: (a) Antibodies displayed on phage surface are allowed to bind a interested target protein, (b) wash out non or weakly bound phage from the target, (c) isolate strongly bound phage by eluting with acidic buffer followed by adding basic buffer to neutralize, (d) do the same manner for more

times to isolate tightly bound antibodies, and enrichment is monitored using phage ELISA in each round, (e) Once enrichment reaches an acceptable level, individual clones are screened with colony screening ELISA and picks clones showing strong signal, and finally selected antibodies are expressed in *E.coli* and apply for the further characterization (Figure 1.4).

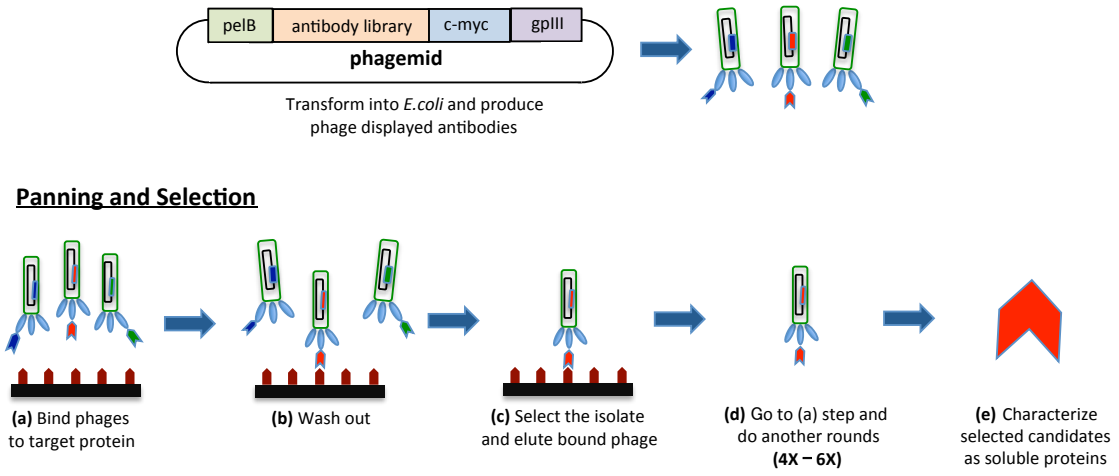


Figure 1.4: In vitro selection using phage display.

Antibody library is constructed in phagemid vector and cells carrying this vector produce antibody on phage surface protein like pIII. The library is allowed to incubate with immobilized target protein, followed by removing weak binders. After three to six rounds of binding, washing, eluting of enriched library, only strongly bound antibodies can be isolated from the selection strategy. Using this technique, tons of clones encoded antibodies can be readily screened for any purpose, and selected candidates are able to be optimized for the further characteristics [31].

Another selection technology is cell-free (Ribosome or mRNA) antibody display by *in vitro* transcription and translation [27,32]. In contrast to the phage-displayed library, the

library size and diversity does not depend on the transformation efficiency of bacterial cells, so more than 10^{13} variants can be screened with this method [33]. Moreover, random mutation for library diversity is easily introduced using non-proofreading polymerases [34]. In order to screen and select antibodies from the libraries, antibody-ribosome-mRNA complex should be accomplished and stabilized during the process. One of the shortcomings is caused from the complex stability because the endogenous nucleases and proteases contained in working system degrade mRNA and also conditions required to stabilize the link are not always ideal conditions for antibody-antigen interactions. Recently, to overcome the shortcomings, the PURE (protein synthesis using recombinant elements) system or mRNA display system has been used [35]. The PURE system only contains a mixture of purified elements for protein synthesis, so there's no effect of nucleases or proteases. And in the mRNA display system, mRNA is able to covalently bonded with its encoded protein resulting in highly stable for *in vitro* selection and screen.

Cell based display is also frequently applied method during selections. Among microbial cells, the bacterium *E.coli* and the yeast *Saccharomyces cerevisiae* are the most commonly used for displaying proteins [36–38]. To screen the cell-based libraries, flow cytometry is used which is the method to measure the amount of bound fluorescently labeled antigen to displayed antibody directly [39]. Compared to previous two types of display technology the critical advantage of cell-based display is that there's higher chance to isolate the antibodies behaving better because the number of bound antigens

per cells depends on the affinity, the expression level of the antibody, and proper folded antibody.

1.3 OBJECTIVES

In this study, we report whether epitope-specific antibodies could be designed and engineered for potential use as generalizable crystallization chaperones. In the first study (Chapter 2), we converted EE-specific scFv to Fab format, which is currently dominant antibody crystallization chaperone format. Its better biophysical properties and additional epitopes providing to mediate crystal contacts have made Fab format having advantages to larger client proteins such as membrane proteins. By subcloning V_L and V_H of previously crystallized EE-specific scFv, we successfully generated Fab/EE retaining binding affinity and specificity. Two model membrane proteins containing EE-peptide were used to form a stable complex and gain crystals by co-crystallization.

In the second study (Chapter 3), we used CDR-specific random mutagenesis to generate single-domain antibody (sdAb) library. The repeat-in toxin domain of *B. pertussis* ACT was used as a target in the library construction, which is an intrinsically disordered calcium binding protein with no homology to any known protein structure and is a candidate vaccine antigen. Using optimized Kunkel mutagenesis method, we constructed 10^6 synthetic phage display single domain antibody (sdAb) library and used to identify sdAbs binding the RTX of *B. pertussis* adenylate cyclase toxin [40]. After the initial screening, we isolated three leads those are specific to sdAbs and this could be used for initial step of construction of synthetic phage display sdAb library for using as

crystallization chaperones that can save enormous time and resources compared to conventional camel immunization.

In the third study (Chapter 4), we designed FLAG-tag specific single-domain variable fragment antibody with a novel computational method. Using an anti-EE peptide specific scFv as an acceptor, the CDRs were designed to specific FLAG peptide. As EE-peptide specific antibodies, the anti-FLAG antibodies can be a potential general crystallization chaperone and engineering antibodies using *in silico* method can save enormous time and resources compared to conventional synthetic library selection strategy. We successfully identified four designed antibodies that possessed minimal FLAG peptide (sequence: DYKD) binding affinity and the binding affinity and specificity of these designed antibodies were experimentally confirmed using phage displayed-scFvs and soluble Fabs. This is the first experimental activity test of OptCDR – generated antibodies and the results proved that this *in silico* method could be used to generate antibody libraries with affinity to diverse antigens as well as other protein engineering methods [41].

In the forth study (Chapter 5), we proposed to develop a number of monoclonal antibodies with high specificity and affinity for A and B subunits of Ricin or Abrin toxins. Moreover, because they share significant homology in both structure and function, we expected to isolate any neutralizing antibodies by binding key catalytic or binding epitopes. To rapidly produce the neutralization antibodies, we generated immune libraries as well as computationally designed antibodies for A and B subunits of both toxins.

Using both approaches, we successfully isolated and developed specific antibodies for each targets.

Chapter 2: Engineering peptide-specific Fab for common crystallization chaperone

2.1 Chapter Summary

Protein co-crystallization is an emerging technique in use of crystallization chaperones to elucidate structures of difficult targets such as membrane proteins. While strategies to date have typically employed protein-specific chaperones, peptide-specific chaperone to crystallize multiple client proteins containing peptide epitope is developing that bypass some of challenges with associated with current target specific chaperone-production step. Based on previously engineered single-chain variable (scFv) antibody fragment that is specific to the peptide sequence EYMPME (scFv/EE), here, we engineered an epitope-specific Fab fragment (Fab/EE). Surface plasmon resonance, an enzyme-linked immunosorbent assay, and size-exclusion chromatography were used to assess Fab/EE binding to EE-tagged soluble and membrane proteins. The addition of constant chains increased the overall stability and had a negligible impact on the antigen binding activity. The 2.0Å resolution crystal structure of Fab/EE reveals contacts with larger surface areas than those of scFv/EE. These results demonstrate that this hypercrystallizable antibody fragment has shown potential as a co-crystallization chaperone for the structural investigation of diverse target proteins.

2.2 Introduction

Crystallization of membrane proteins remain the bottleneck to structure determination. Various strategies in each stage of the pipeline have developed to enhance the probability of the crystallization of difficult proteins [42]. In molecular biology and protein expression stage, hydrophobicity of membrane proteins causes the aggregation hampering the crystallization. To overcome this, bacterial orthologs, which are more

stable in bacterial expression system but representing very similar structures can be used. Or Using the stabilizing mutations that do not interfere conformational change but promote crystal growth is alternative ways to increase chances of crystallization [31]. In protein purification steps, selection of appropriate detergent for protein solubilization and crystallization is a critical factor to optimize. In crystallization stage, there are much more options to consider to improve crystallization. First, even though a good ligand (agonist or antagonist) is unknown in many cases, add ligand or other additive is a good way to make less conformational heterogeneity. Another strategy is replacing flexible or less stable loop of membrane protein with more compact three-dimensional soluble protein such as T_4 bacteriophage lysozyme resulting in increased hydrophilicity and diffraction quality crystal formation [43,44]. Lastly, to date, the crystallization chaperone approach has been increasingly applied for elucidating structures of membrane proteins. The desirable biophysical properties of chaperones include an increase in hydrophilic residues available for forming crystal contacts, resulting in the improvement of obtaining well-ordered crystals of the chaperone-target membrane protein complex. Covalent or non-covalent chaperones have been utilized to crystallize several G protein-coupled receptors (GPCRs) [45,46]. Previously, we proposed the use of peptide-specific antibody fragments as a potentially general crystallization chaperone to overcome limitations of protein-specific chaperones [31]. To the best of our knowledge, the only successful example of a target-independent non-covalent chaperone involved the use of a Fab fragment that recognizes a portable small structural RNA element to crystallize a ribozyme [47,48]. Co-crystallization with client membrane proteins has been attempted using our engineered peptide-specific

scFvs exhibiting hypercrystallizable characteristics and forming a stable complex with EE peptide-tagged proteins [6,49] and using commercially available FLAG-binding Fab fragments, but successful co-crystallization has not been achieved to date. A great number of co-crystal structures of membrane protein deposited in the Protein Data Bank (PDB) were solved using Fab format of antibody crystallization chaperone [31]. Due to its larger size (~54kDa) compared to scFv (~28kDa), the Fab format can provide additional solvent exposed hydrophilic surfaces for lattice-forming crystal contacts and confer improved stability and uniformity into the crystal lattice [50]. Here, we generated a Fab fragment with nanomolar affinity for the EE peptide (Fab/EE) that is shown to be a potential crystallization chaperone [51]. The Fab/EE could form a successful complex with a variety of proteins containing EE peptides, including EE-tagged maltose binding proteins (MBP-EE or MBP-KEE), the α-helix membrane protein human adenosine A_2a G proteins-coupled receptor (A2aR-GFP-EE), and the *E.coli* β-barrel membrane protein intimin (intimin-EE1 and intimin-EE2). We tried co-crystallization of the complex of Fab/EE and EE-contacting target protein after isolating from SEC, but unfortunately, none has resulted in diffracting crystals yet. The results indicate that even though the EE peptides that we placed into target proteins does not interfere with protein expression and purification, unintended removal of native contacts may be happened to prevent crystallization, therefore, more efforts to determine optimal place and loop length for complexation and crystallization should be required for especially for target proteins of unknown structure.

2.3 MATERIALS AND METHODS

2.3.1 Conversion to Fab version of peptide specific antibody

To convert 3D5/EE_scFv to the Fab format, the V_L and V_H domains were amplified by PCR using gene-specific primers, 5'pAkpel and 3'NotI into 48 for V_L and 5' NheI into 48 and 3'HindIII into 48 for V_H. Each PCR product was subcloned into NcoI-NotI and NheI-HindIII sites of the pFab vector and resulted in Fab/EE. Sequence was identified from the University of Texas at Austin Core Facility.

2.3.2 Expression and purification of Fab/EE and peptide carrier proteins

Fab/EE was expressed in *E.coli* BL21 cells. A fresh single colony of Fab/EE was inoculated in 2ml Terrific Broth (TB, Fisher Scientific) with 200μg/ml ampicillin and incubated for 6~8 h at 37°C. The starter culture was diluted in 250ml TB in a 1L baffled flask and grown for overnight at 37°C. Cells were harvested and resuspended using same volume of fresh TB and incubate for 1 h at 25°C before inducing protein with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Fisher Scientific) for 5 h. The cells were pelleted and resuspended with 10ml of resuspension buffer (0.1M Tris pH8.0, 0.75M sucrose). To fractionate the cells, osmotic shock was carried out by adding 10ml 1mM EDTA and lysozyme (10mg/ml), stirring for 45min at 4°C, then adding 0.5ml 0.5M $MgCl_2$, and stirring additional 45min. After centrifugation for 20min, the supernatant was pooled and dialyzed in dialysis buffer (10mM Tris pH8.0, 0.5M NaCl) for overnight at 4°C. Ni^{2+}-based immobilized metal affinity chromatography (IMAC) resins were

subjected to the dialyzed supernatant and protein was eluted using 1ml of elution buffer (0.1M EDTA, 20mM Tris pH8.0, 0.5M NaCl). Eluted protein was further purified by size exclusion chromatography (SEC) using an ÄKTA FPLC system (GE healthcare) and a Superdex S75 (GE healthcare) column with a HBS (pH7.4) running buffer. A maltose binding protein (MBP) was used as a client protein to carry the EE peptide. Both MBP-EE that EE-peptide was appended to the C-terminus and MBP-KEE that a surface loop of MBP (residue 170-175) was replaced with KEE peptide were used for test proteins and MBP-$HIS_{6,}$ was used as a negative control. These proteins were expressed and purified via IMAC as described for Fab/EE. EE peptide-carrier membrane proteins, A_2aR and intimin, were used as target proteins for crystallization. All molecular works expression, and purification for these proteins were performed by Dr. Kevin Entzminger in Maynard lab (University of Texas at Austin) and Dr. Jennifer Johnson in Lieberman lab (Georgia Tech). Briefly, an EE-tagged variants, pITy-A_2aR-GFP-EE provided by Dr. Anne Robinson (University of Tulane), was constructed by adding GS-EE-GS epitope after Lys209 residue in the third intracellular loop (ICL3) using site-directed mutagenesis. The WT and EE-containing A_2aR-GFP plasmids were transformed into *Saccharomyces cerevisiae* BJ5464 by electroporation and after screening the expression level of GFP, the highest expressing clones were used for the rest of experiments as described previously [52]. For *E.coli* intimin, the plasmid was generously provided by Dr. Susan Buchanan (NIH). Two variants of intimin were generated: EE peptide was inserted into an extramembraneous loop in WT intimin, named intimin-EE1 [49] and EE epitope flanked by

AA on each side was incorporated for intimin-EE2. The intimin variants were expressed described as previously as wild-type intimin [53].

2.3.3 Biophysical characterization of Fab/EE

For Fab/EE, protein purity and size were determined by reducing or non-reducing 12% SDS-PAGE analysis and protein concentration was estimated by BCA assay (Pierce). Monomeric fractions of Fab/EE were pooled for further experiments. Protein solubility was determined by measuring the concentration of remaining soluble protein after 4 days incubation of concentrated Fab/EE (20mg/ml) at 4°C. Thermal stability was measured by thermal unfolding after mixing 20µl of 200mM Fab/EE or HBS buffer blank with SYPRO Orange (final: 15X, Molecular Probes) in a real-time PCR machine (ViiA7, Applied Biosystems) in increments of 0.96°C/min from 25°C to 90°C. Analysis to determine the melting temperature (T_m) was performed with the ViiA 7 software.

2.3.4 Determination of binding specificity by ELISA and SEC

To measure the binding affinity of Fab/EE to EE peptide, two soluble protein ELISA was performed. Firstly, EE peptide containing carrier proteins, MBP-EE or MBP-KEE, and MBP-His_6 , which is control protein were coated (need to add concentration) on high-binding ELISA plate wells (Costar™) overnight at 4°C and blocked with blocking solution (PBST with 5% milk) for 1 h at room temperature. Purified Fab/EE was serially diluted in blocking solution and allowed to bind with coated proteins with 1 h incubation. Bound Fab/EE was detected using anti-human Cκ-hrp. After final washing step, TMB

substrate (Thermo Scientific) was added to develop the signal and the reaction was quenched with 1N HCl. Absorbance was read at 450nm on a SpectraMax M5 microplate reader (Molecular Devices). Data was analyzed using *Graphpad Prism 6*. Inverse version of protein ELISA was also performed. In here, Fab/EE was coated on the wells and serially diluted MBP-EE, MBP-KEE and MBP-HIS_6 were added to measure the affinity. And then, anti-MBP-HRP was used for detection antibody.

Fab/EE – Client protein interactions were further evaluated by SEC. Equimolar of Fab/EE and MBP-KEE were incubated for 90min either together or separately at room temperature prior to fractionation on a Superdex S200 column (GE healthcare). Elution fractions for each peak were analyzed by 12% reducing SDS-PAGE

Kinetic binding assays were performed by surface plasmon resonance (SPR) with a BIAcore 3000 (GE Healthcare) instrument using immobilized bovine albumin (BSA) or Fab/EE coupled to CM5 chips via NHS-EDC chemistry to a level of ~1200 RU as bait for ligand proteins. Responses owing to sample refractive-index changes and nonspecific binding were corrected using the signal from a flow cell coupled with BSA. Purified MBP-KEE, MBP-EE, scFv-EE_1 or control MBP-HIS_6 were injected in a duplicate dilution series from 2 to 0.125μM at a flow rate of 50μl/min to minimize mass-transport effects in HBS running buffer supplemented with 0.005% Tween-20. Surface regeneration was performed after each run with a single 30s injection of 2M $MgCl_2$. The association rate constant (k_{on}), dissociation rate constant (k_{off}) and equilibrium dissociation constant (K_d; $K_d = k_{off}/k_{on}$) were calculated assuming a Lanmuir 1:1 binding model with the BIAevaluation software. Only data sets with $\chi^2 < 0.5$ were used.

2.3.5 Protein crystallization, date collection, structure determination and refinement

All crystallization-related works were done by Dr. Jennifer Johnson in Lieberman group (Georgia Tech). 6.5 mg/ml of Fab/EE was crystallized at room temperature by the sitting drop vapor-diffusion method. Crystal conditions were screened with Wizard I and II (Emerald Bio) and optimized based on solution G4 consisting of 20% polyethylene glycol (PEG) 8000, 100mM MES (pH 6.0), 200mM calcium acetate. Crystals used for structure determination were grown from a reservoir solution including 0.1M HEPES (pH 7.5), 100mM calcium acetate, 20-26% (w/v) PEG 8000, and 3% 1-propanol. The crystals were harvested and cryocooled in reservoir solution supplemented with 15% glycerol. Crystallographic data were collected and the structure was solved as described previously [51]. The final refined structure was deposited in the PDB as entry 4x0k. Fab/EE crystal contacts and critical amino acids in crystal contacts were analyzed by surface area and energy. After excluding the native heavy-light chain interface within the Fab/EE monomer, the top three interfaces were identified as major crystal contacts and used in further analysis.

2.4 RESULTS

2.4.1 Rationale and characterization of Fab/EE

Previously, we successfully engineered peptide-specific scFvs that is specific to EYMPME or hexa-histidine for using as crystallization chaperones [49,54]. The scFv/EE proved to have higher affinity and to be more preferable than anti-hexa-histidine scFv due to its chemical diversity, its insensitivity to pH especially at near-physiological

conditions and its greater compatibility with a variety of peptide-insert location (C-terminal and internal). Even though scFv/EE recognized EE peptide carrier proteins, successful co-crystallization has not been reported yet. In protein co-crystallization, a majority of crystal structures solved with assistance from conformation-specific Fab antibody fragments due to their better biophysical properties and larger size than scFv[55]. 54kDa of Fab would provide additional epitopes to mediate crystal contacts, especially advantageous for larger client membrane proteins. Therefore, we converted the EE-peptide specific scFv to Fab format by sub-cloning the variable regions into the pFab vector[56], and the Fab/EE was produced in *E.coli* strain BL21 (Figure 2.1). The level of total and monomeric Fab/EE after purification was similar as scFv/EE (2.4 mg/L culture versus 2.1 mg/L culture; 87% versus 81% monomeric, respectively), but thermal stability determined by melting temperature was significantly enhanced as compared to scFv/EE (Table 2.1). Taken together, biophysical properties, which are strongly correlated with success rates of crystallization[57] demonstrate that generated Fab/EE is more promising than the parent scFv/EE for use in large-scale co-crystallisation trials.

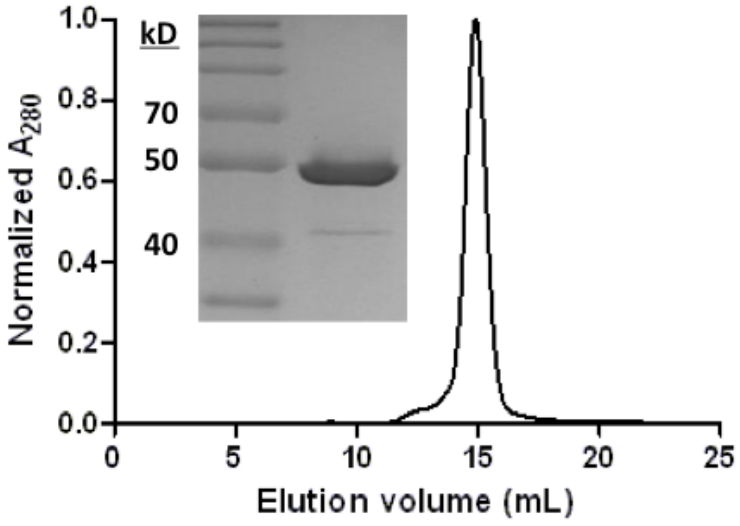


Figure 2.1: Fab/EE production and purification.

Fab/EE was purified from the *E.coli* periplasm by IMAC followed by SEC and eluted using an analytical gel filtration column as a single peak at the expected elution volume. *Inset*, non-reducing SDS-PAGE gel of a fraction taken from the major monomer peak and stained with coomassie dye predominantly shows a single dominant band at ~50kDa and >95% purity.

Table 2.1: Biophysical characteristics of EE peptide-binding antibody fragments.

	3D5/EE_48 scFv[a]	**3D5/EE_48 Fab (Fab/EE)**
Expression level (mg/L culture)	2.1	2.4
Solubility (mg/ml)	12.8	9.4
Melting temperature (°C)	47 ± 0.3	59.8 ± 0.1
% Monomeric protein	81	87
Affinity (MBP-EE-His)	389 nM	308 nM

[a]As reported in Pai, J.P. *et al.* (2011) *Protein Eng. Des. Sel.*

The Fab format shows comparable expression levels but lower overall solubility as compared to the scFv format. Both melting temperature and % monomeric fraction of the Fab are improved compared to the scFv format

2.4.2 Fab/EE structural characterization

Crystals of Fab/EE belong to space group P1 and its 2.0Å resolution structure was solved by molecular replacement. Most residues were successfully modeled except Ser128-Ser134 in chain H, the FLAG tag at the C-terminus of the heavy chains (chains H and A), the linker residues between the C-terminal ends of the light chain (chains L and B) and the last four histidine residues in 10 X histidine-tag. The lattice of Fab/EE demonstrates the variety of crystal contacts available to aid in the crystallization of client proteins. Overall, the crystal contacts of Fab/EE had larger interface surface areas (1099 $Å^2$ at most) than those of scFv/EE (Figure 2.2b). The only residue that participates in crystal contacts in both Fab/EE and scFv/EE is Lys107 (Figure 2.2b) even though the two interfaces are not similar. Unexpected interaction was shown that a portion of the decahistidine tag on the C-terminus of the light chain of Fab/EE forms hydrogen-bonding and salt-bridge interactions with the CDRs of the heavy chain. Since the negative control proteins for complexation and co-crystallization studies contain histidine tag, we assume that this is an artifact of crystallization. Finally, the *P*1lattice lacks solvent channels to accommodate a client EE-tagged protein (Figure 2.2d) and the CDRs are being used in the crystal contacts. Upon binding to target protein, Fab/EE would likely utilize other residues that are available for forming crystal contacts, perhaps including the crystal contact areas seen in scFv/EEs as represented previously [49].

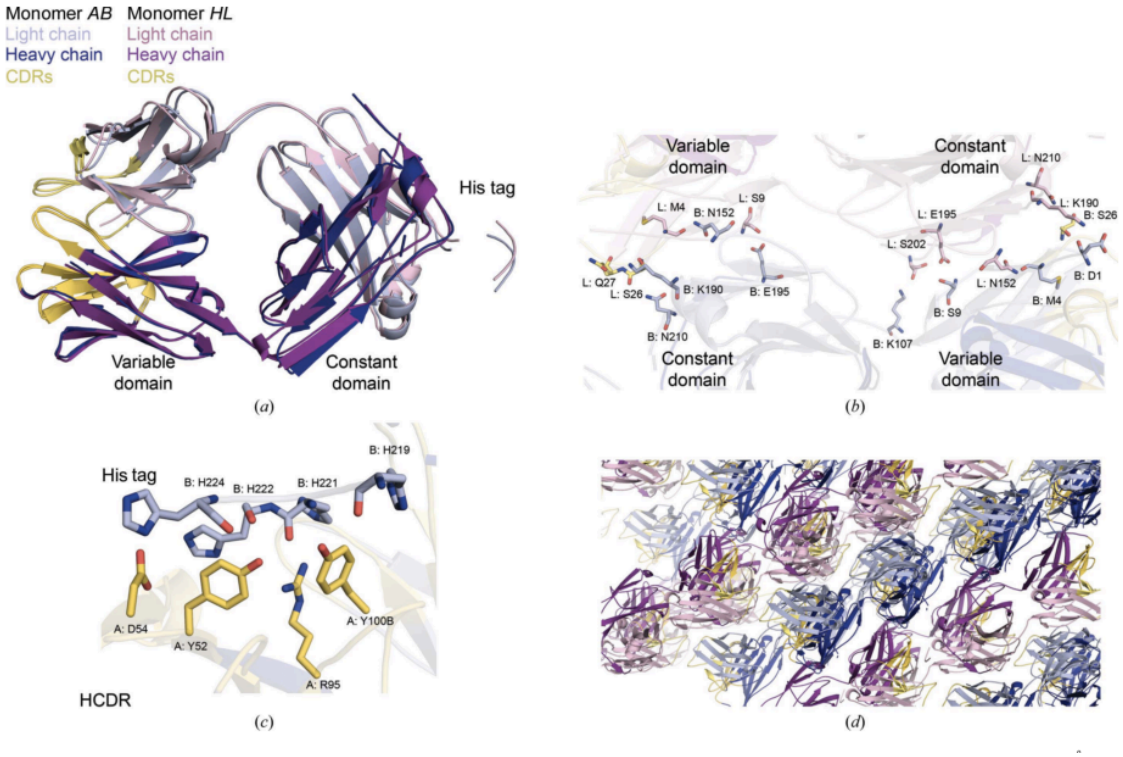


Figure 2.2: Structure of Fab/EE.

(a) Overlay of two Fab/EE molecules in the asymmetric unit. (b) Fab/EE crystal contact ID2. The highest interface is 1109 Å and includes nine residues from each L and B chain. (c) Crystal contact ID5. The figure represents interaction between the modeled decahistidine tag from chain L and CDR of chain H of different molecule. (d) Fab/EE lattice showing extended crystal contact areas and the lack of a channel that could accommodate a membrane protein.

2.4.3 Fab/EE forms stable complexes with soluble EE-tagged client proteins

The binding affinity of Fab/EE to EE peptide was accessed by examining interactions with EE-peptide containing soluble proteins using SPR. Fab/EE was shown to bind to EE peptide at c-terminus (MBP-EE; Table 2.1), internal (MBP-KEE), or flexible linker region of scFv (scFv-EE_1) in the nanomolar range but no binding was observed to hexa-histidine containing protein (MBP-His_6), the negative control (Figure

2.3). The results indicate that Fab from the conversion of the scFv format retained binding affinity to the target, as previously reported [58]. Next, we tested binding affinity with similar condition as in solution or crystallization drop by assessing complex stability. SEC was performed to separate MBP-KEE, Fab/EE alone or the complex, and a clear shift in elution volume corresponding to higher molecular mass was observed for the complex. The size of complexation was confirmed by SDS-PAGE (Figure 2.4).

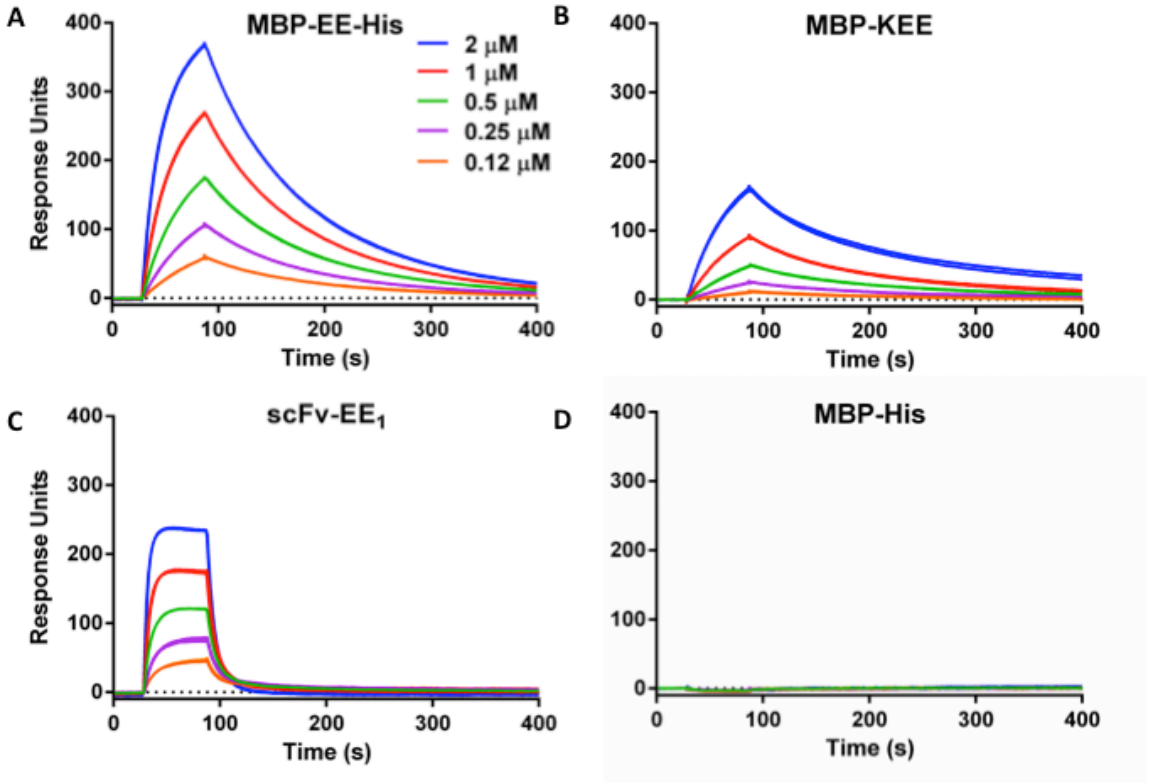


Figure 2.3: SPR reveals mid nM affinity of Fab/EE for EE peptide-carrier proteins.

Peptide-containing ligands were injected in duplicate for each concentration tested and binding to immobilized Fab monitored by SPR. Both duplicates traces are shown and demonstrate Fab binding to (A) MBP appended with a C-terminal EE peptide followed by a His_6 peptide, (B) MBP containing the EE peptide inserted into a surface-exposed

loop, and (B) a scFv containing the EE peptide inserted into the inter-domain linker. (D) No binding was observed for MBP appended with on a His_6 peptide.

Table 2.2: Characterization of EE peptide ligand binding kinetics by SPR

	k_{on}, $M^{-1} \cdot S^{-1}$	k_{off}, S^{-1}	k_D, nM
MBP-EE-His	$3.39 \pm 1.23 \times 10^{4}$	$9.22 \pm 0.15 \times 10^{-3}$	308 ± 117
MBP-KEE	$1.25 \pm 0.30 \times 10^{4}$	$7.88 \pm 0.91 \times 10^{-3}$	612 ± 95
scFv-EE$_1$	$2.95 \pm 1.47 \times 10^{5}$	$5.56 \pm 1.66 \times 10^{-2}$	224 ± 160

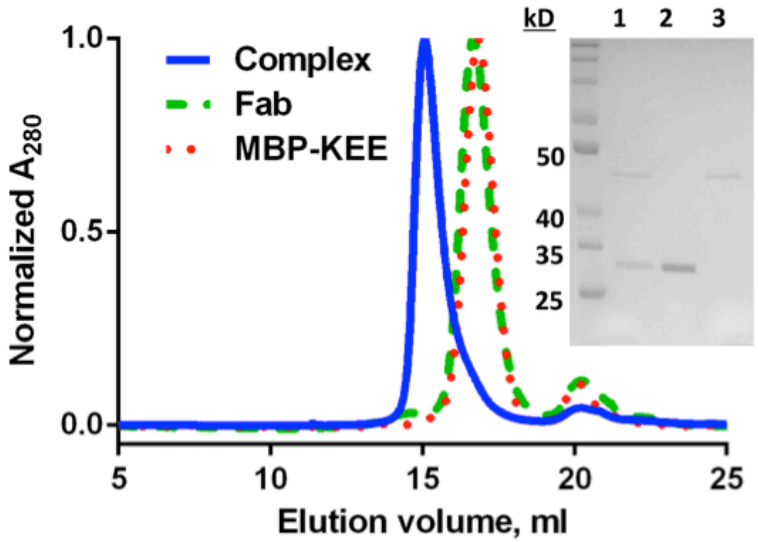


Figure 2.4: Fab/EE forms a stable complex with EE peptide-containing proteins in solution.

Equimolar amounts of MBP-KEE and Fab/EE were incubated at room temperature separately or together prior to separation by analytical gel filtration. The sample containing both proteins eluted as a single peak with a smaller elution volume than either individual protein. *Inset,* fractions from each major peak were analyzed by reducing SDS-PAGE, Lane 1, peak fraction from the putative complex peak. Lane 2, peak fraction from the Fab/EE peak. Lane 3, peak fraction from the MBP-KEE peak.

2.5 DISCUSSION

Previously, we reported the successful engineering scFv/EE for potential use as crystallization chaperones [6]. We proved that the scFv/EE exhibited higher affinity and preferable over scFv/His due to its chemical diversity, insensitivity to pH, and accessibility of peptide insertion site (C-terminus or internal). But despite this advantage, no co-crystallization has been successful with scFv/EE yet. Here, we sought to convert this EE-specific scFv to Fab format because so far, the majority of crystallization chaperone proteins in solved structures have been represented as Fab format. Fab platform is known to exhibit better biophysical properties than scFv format and its larger size may provide additional surfaces to mediate crystal contacts, especially for larger client proteins. Therefore, Fab/EE is thought to be more promising molecule than the scFv/EE for use in co-crystallization.

The Fab/EE was crystallized in space group *P*1 and CDR residues were involved in crystal contacts unlike the crystal lattice of scFv 3D5. Because most residues of scFv/EE except linker region are also in Fab/EE, all residues that form crystal contacts in scFv/EE are also available in Fab/EE and other residues may be available for forming crystal contact when Fab/EE serve as a crystallization chaperone.

Using SPR and SEC, we proved that Fab/EE could bind to several EE-tagged client proteins when the target proteins are soluble. Fab/EE also could bind to EE epitope that is placed in the middle or c-terminus of client protein (MBP) so that the use of Fab/EE as a general crystallization chaperone is promising. In order to prove whether engineered Fab/EE is able to be used for crystallization chaperone and show better

behavior over scFv/EE, we used Fab/EE:EE-tagged MBP complex set up for crystallization trial. Even when placement of the EE peptide dose not interfere with protein expression and purification, an unintended consequence may be the removal of native contacts and thus an increase in conformational heterogeneity that is detrimental to crystallographic efforts. Such conformational changes can be challenging to predict, especially in case of target proteins of unknown structure. If an inflexible region of the target protein is not known, a reduction in loop flexibility is to be achievable by shortening the EE epitope-containing loop to a minimum number of residues that can still be complexed with Fab/EE.

Chapter 3: Design synthetic single domain antibody library by phage display

3.1 Chapter Summary

The single domain antibody fragment (sdAb) is generating major interest as a crystallization chaperone to elucidate the structures of recalcitrant targets such as intrinsically disordered proteins or membrane proteins. Given that current protocols require immunizing camels, which are not widely accessible to academic laboratories, we created a synthetic phage display sdAb library as a source of sdAb binding potentially any antigen. A synthetic library is expected to save significant time and resources compared to the conventional camel-based method. To demonstrate proof-of-concept of the library's utility, we identified sdAbs binding two model antigens, maltose binding protein (MBP) and a bacterial toxin. The RTX domain of *B. pertussis* adenylate cyclase toxin (ACT) was chosen since it is an intrinsically disordered calcium binding protein with no homology to known structures and a protective antigen in whopping cough. After standard panning, three sdAb with moderate affinity for MBP and RTX were identified. The RTX-binding sdAbs were subjected to affinity maturation and further panning rounds. One RTX-specific sdAb exhibited weak binding affinity yet high solubility, while the other two sdAbs exhibited higher binding affinity, yet poor solubility when expressed in *E.coli*. For those sdAbs, increased affinity correlated with increased hydrophobicity of the binding site and a corresponding decrease in solubility, suggesting molecules may be best suited for hydrophilic targets. Even with this limitation, the sdAbs described here are the first to be isolated targeting a highly challenging protein without using camel immunization. This approach should be generalizable to other disordered or membrane protein targets for screening sdAb co-crystallization chaperones. With improvements in solubility, these innovative antibody fragments have the potential to be

powerful tools to further our understanding of the structure of recalcitrant proteins, and ultimately facilitate future structure-based drug discovery.

3.2 INTRODUCTION

The single domain antibody platform (sdAb) has become an attractive platform for developing novel antibodies for therapeutic applications and as research tools. In contrast to conventional antibodies that consist of light and heavy chains, these heavy chain-only antibodies initially discovered in the Camelidae family (V_HH), retain functionality and have evolved to be autonomously stable without a light chain partner [59,60]. Key characteristics of sdAbs include their small size (11 – 15kDa), long CDR3 and high thermal stability. Due to their small size and corresponding rapid *in vivo* clearance, they have generated interest to target imaging agents to specific cells for diagnosis and therapy [61–63]. Their relatively long CDR3 provides access to recessed active sites or cavities [62]. Moreover, the V_HH exhibit high thermal stabilities and can often be fully refolded upon denaturation [64]. Because V_HH molecules exhibit characteristic antibody-antigen binding traits, they may be used as building blocks for multivalent or bispecific antibodies [61].

Along with the V_HH, several studies have discovered autonomously V_H domains [65,66]. Unlike V_HH molecules, the stability and solubility of the autonomous V_H are not dependent on CDR3 regions, which is the major contributor to the antigen-binding sites. Therefore, using the autonomous V_H domain as a scaffold and altering CDRs in antibody engineering for novel antibodies is more promising approach than humanization of V_HH

or camelization of V_H. Previously, an autonomous V_H domain (4D5.B1ag) was engineered from an human conventional antibody, the Herceptin-binding 4D5 molecule [67]. The 4D5.B1ag includes just six amino acid changes at the heavy chain-light chain interface (H35G, Q39R, L45E, W47L, S93A, and W103S), which allow for soluble expression of the V_H domain without a light chain partner [65,68]. Notably, these residues do not occur in human antibodies or camelid V_HH repertoires. Because 4D5.B1ag stability and solubility are independent of the CDR3 loop sequence, it may be an attractive scaffold for generating synthetic sdAb libraries.

Here, we chose the single domain antibody platforms to develop as crystallization chaperones for obtaining structural information on highly challenging proteins. In addition to the advantages described above, in the context of protein crystallization, their beta-sheet rich structure that can promote nucleation and crystal lattice formation, while their binding can mask hydrophobic regions of client proteins, and more hydrophilic surface areas available to form a crystal lattice [2]. Although the sdAbs have been successfully used as membrane protein crystallization chaperones [69–71], notably for crystallization of GPCRs in the active state, identification of these molecules remains limited by the need for camel or shark immunization [70,72–77].

To provide a source of sdAb, which is broadly accessible to most lab with molecular biology expertise, we have develop a synthetic sdAb phage display library based on the 4D5.B1ag sdAb. From this library, we isolated sdAb binding two antigens, the maltose-binding protein (MBP) and the repeat-in-toxin (RTX) domain [67] an intrinsically disordered calcium-binding protein and protective antigen with no homology

to endogenous proteins. A total of seven positive clones for RTX and 14 positive clones for MBP were identified after initial screening and the top three ranked positive sdAb clones from each library were chosen for further characterization. Among them, one of the isolated RTX-specific sdAb having weak binding affinity exhibited high solubility, and two that bind great, exhibited poor solubility when express *E.coli*. While the synthesized sdAb library has identified some potential sdAbs, top scoring anti-RTX sdAbs are still being tested for improved biophysical characteristics. The results suggest that the engineered phage-displayed sdAb libraries may provide a readily accessible source of crystallization chaperones for a wide range of challenging protein targets.

3.3 MATERIALS AND METHODS

3.3.1. Library design and construction

Synthetic phage-displayed sdAb libraries were designed and constructed using previously described methods based on Kunkel mutagenesis of the 4D5.B1ag sdAb in the pMoPac24 phagemid [66,79,80]. To introduce genetic diversity into the CDRs of the parent sdAb 4D5.B1ag, NNS or KMT degenerate codons (N=A/C/G/T, S=C/G, K=G/T, M=A/C) were used. However, a few residues in CDRs, which are known to be highly reserved, were retained as the same sequence with the 4D5.B1ag (Table I). To generate synthetic libraries, a uracil-containing, single-stranded plasmid DNA (dU-ssDNA) of pMoPac24_4D5.B1ag containing stop codon and *Sac*II restriction enzyme recognition site was prepared as previously described [40]. The prepared dU-ssDNA annealed with phosphorylated mutagenic oligonucleotides for synthesis of covalently, closed, circular double-stranded DNA (CCC-dsDNA). CCC-dsDNA was digested with SacII to minimize the presence of wild-type sequence and then transformed into *E.coli XL1-Blue*. The

diversity of the resulting library was measured by plasmids sequencing. After expression of the library as phage-display sdAbs [12], the library phage were allowed to bind immobilized anti-cMyc tag (Sigma 9E10), protein A, and target antigen, RTX_{751} or MBP for four rounds in order to remove stop codon containing conformational unfavorable sdAbs and to select target specific antibodies [67]. In each round of panning, bound phages were eluted with 0.1M HCl (pH2.0), immediately neutralized with 2M Tris base, and re-amplified for the next round of panning or selection to target.

3.3.2 Screening and selection by phage ELISA

A sufficient number of colonies from the library, 144 colonies from each library, were screened to confirm the synthesis of library and binding activity after panning. High-binding 96-well plates (Costar™) were coated and incubated with a target antigen, RTX_{751}, MBP (diluted in PBS, pH7.4, at 4µl/µg), or anti-c-Myc antibody (1µl/µg) at 4°C for overnight. The wells were blocked with blocking buffer (5% non-fat milk in PBS-0.05% Tween, PBST) followed by adding phage (1:1 diluted in blocking buffer); those were produced from a single colony. After 1-hour incubation and washing with PBST, anti-M13-HRP (GE healthcare, 1:3000 in 2.5% non-fat milk PBST) was applied and incubated an additional one hour. After four washings, signals were developed using 50µl of TMB solution (Thermo Scientific ™ Pierce ™,) and quenched with the same volume of 1N HCl. Signals were determined by measuring the absorbance at 450nm using a SpectraMax M5 microplate reader (Molecular Devices). Clones showed high absorbance with anti-c-Myc or $RTX_{751,}$ indicating that the level of antibody displays on phage or target binding activity were selected for further characterization. To verify the binding

activity and specificity, mono-phage ELISA was performed in similar ways but with enriching phages in 30ml of 2X YT media containing 0.1% glucose and antibiotics, serially diluted with blocking buffer (1:√10), and added to the wells.

3.3.3. Protein production and purification

The lead sdAb variants selected from the colony screening were sub-cloned into the pAK400 vector for soluble expression of sdAb proteins. Periplasmically expressed antibodies collected by osmotic shock, purified by IMAC and size exclusion chromatography (SEC, S75), as previously described [6]. Positive control anti-RTX antibodies (M1H5 or M2B10) antibodies in the scAb format were produced and purified as previously described [78]. A starter culture of scAb was inoculated in 100ml of Terrific Broth (TB) media containing ampicillin (200µl/µg) and 1% glucose, then grown at 25°C overnight. The next day, cells were pelleted by centrifugation, then resuspended with the same volume of fresh TB media without glucose and grown at 25°C for 1 h. Antibodies were then induced for an additional 4 hours by adding 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Fisher Scientific). After harvesting the cells, osmotic shock was used to collect produced proteins. Purification was conducted as described above.

The RTX domain (residues 751-1706) was expressed in *E.coli* strain BL21(DE3) and purified as previously described [78]. Briefly, a starter culture of the RTX domain was inoculated in 250ml TB media, grown at 37°C until OD_{600} = 0.5, and induced by adding 0.4mM IPTG for 4 h at 25°C. The cells were pelleted and resuspended with 10ml of resuspension buffer (50mM HEPES, 250mM NaCl, 2mM $CaCl_2$, 40mM imidazole, pH

8.0). A French press (Thermo Scientific) was performed to lyse the cells, followed by 20-min centrifugation at 20,000 rpm (JA-20 rotor). The supernatant was applied to a HisTrap column (GE Healthcare) and the RTX domain was eluted with a linear gradient of elution buffer (resuspension buffer + 500mM imidazole).

3.3.4. Biochemical Assays

To measure sdAb-RTX_{751} binding activity, a soluble protein ELISA was performed using purified sdAbs. Ninety-six (96) high-binding wells were coated with soluble wild type V_H or sdAb variants (10μg/ml) and blocked with blocking solution as above. After three washes with PBST, biotinylated RTX_{751} was added with serial dilution and incubated for 1 h at room temperature. Biotin labeling followed manufacturer's instructions (Thermo Scientific). Bound RTX_{751} was detected using HRP-conjugated streptavidin (BD CELL ANALYSIS).

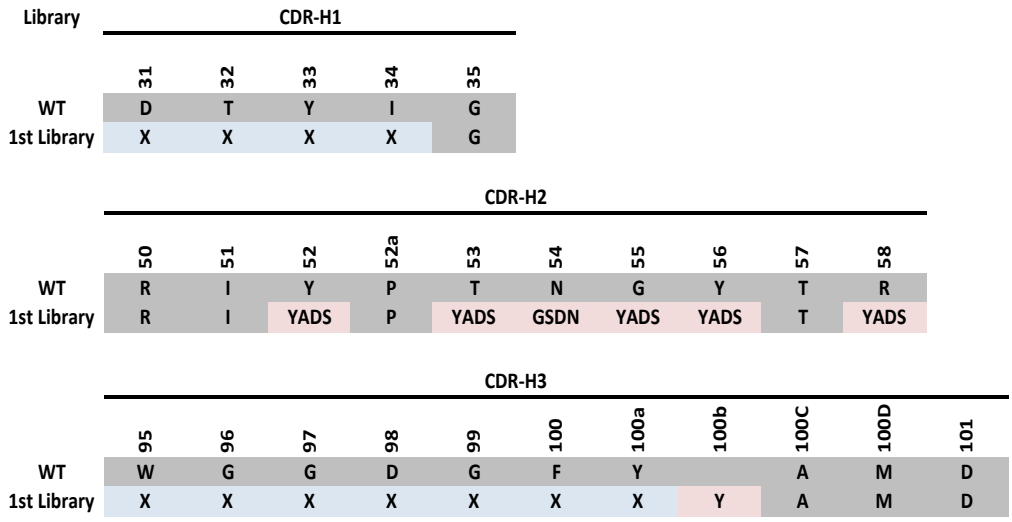
To test the binding specificity of sdAb to different species of RTX, monoclonal phage ELISA was performed as described above with a few modifications. Wells were coated with the RTX_{751} of *Bordetella pertussis (Bp), Bordetella bronchiseptica (Br),* or *Bordetella parapertussis (Bpp),* and blocked. After adding serial diluted phages, bound phages were detected using anti-M13-HRP (1:5000 diluted in blocking solution). To access a binding epitope, completion ELISA was performed similarly to phage ELISA but with more modifications. Serially, 1 to 2 diluted soluble control antibodies (M1H5 or M2B10) were added to the RTX_{751} coated and blocked wells. After 1 h pre-incubation of

M1H5 or M2B10, a monoclonal antibody-displayed phage was added, followed by additional 1 h incubation. Bound phages were detected as described above.

3.4 RESULTS

3.4.1. Design and construction of 1st generation of synthetic sdAb library

In order to engineer sdAbs that are specific to desired antigens (RTX domain or MBP), 4D5.B1ag was chosen as an acceptor due to its well-known autonomous behavior. We designed the sdAb library by introducing sequence diversities into all three CDRs of 4D5.B1ag using previously described methods [66]. In CDR1, positions 31 – 34 were completely randomized and residue 35 was fixed with G because it is highly conserved residue in most of camel nanobodies as well as the human V_H domain. In CDR2, positions 52, 53, 55, 56 and 58 were replaced with 4 amino acid YADS, which are known to play a predominant role in antigen recognition[81]. Another set of 4 amino acids, GSDN, was applied into the residue 54 because those four amino acids are abundant in that region. In CDR3, residues 95 – 100a were fully randomized to acquire unique target specificity (Figure 3.1). We successfully constructed both libraries and randomly picked ten clones from each library, proving that the libraries were generated with high quality (90% sequence diversity).



Library	CDR-H1				
	31	32	33	34	35
WT	D	T	Y	I	G
1st Library	X	X	X	X	G

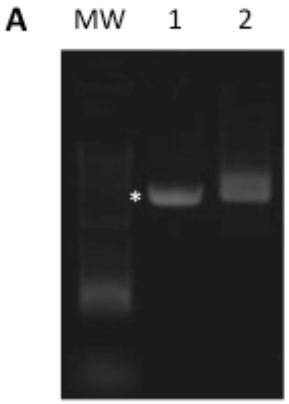
	CDR-H2									
	50	51	52	52a	53	54	55	56	57	58
WT	R	I	Y	P	T	N	G	Y	T	R
1st Library	R	I	YADS	P	YADS	GSDN	YADS	YADS	T	YADS

	CDR-H3										
	95	96	97	98	99	100	100a	100b	100C	100D	101
WT	W	G	G	D	G	F	Y		A	M	D
1st Library	X	X	X	X	X	X	X	Y	A	M	D

Figure 3.1: Rationale for synthetic sdAb library design.

Design of CDRs for constructing synthetic sdAb library. Residues that were fixed as the acceptor 4D5.B1ag are shaded gray. X indicates residues that were replaced by all possible amino acids (blue). The others that were represented as single-letter code, only restricted amino acids could be applied (pink).

Only one out of ten clones contained stop-codon after randomization, which represented the functional capacity of the library.[82] No frame-shift mutations were detected as expected. The overall library size, which is important to obtain high affinity clones, was determined as approximately 10^6 (Figure 3.2).



B Library size based on degeneration codon

Library	Degeneration codon	# primers	Diversity	% recombinant
Random	NNS	3	$1.15X\ 10^6$	90

Figure 3.2: Generation of ccc-DNA using designed oligonucleotides

(A) Random library using three designed oligonucleotides. Lane 1: dU-ssDNA of 4D5.B1ag with STOP codon and restriction enzyme site in CDR3 region, lane 2: ccc-dsDNA product. MW indicates 1kb plus DNA ladder (Invitrogen), * (white; expected size of dU-ssDNA and black; ccc-DNA)

3.4.2. Selection and characterization of RTX specific sdAb variants

In order to select sdAbs that acquired specificity to the targets, amplified phages from the library were pooled and incubated with RTX_{751} as a structurally recalcitrant target or Maltose Binding Protein (MBP) as a soluble protein control. Prior to panning with target antigens, the library phages were allowed to bind anti-cMyc and protein A to filter out any stop codon-containing and misfolded sdAbs, respectively[66]. After three or four cycles with targets, enough colonies were screened to identify positive clones that demonstrated matured affinity to targets by phage ELISA (Figure 3.3).

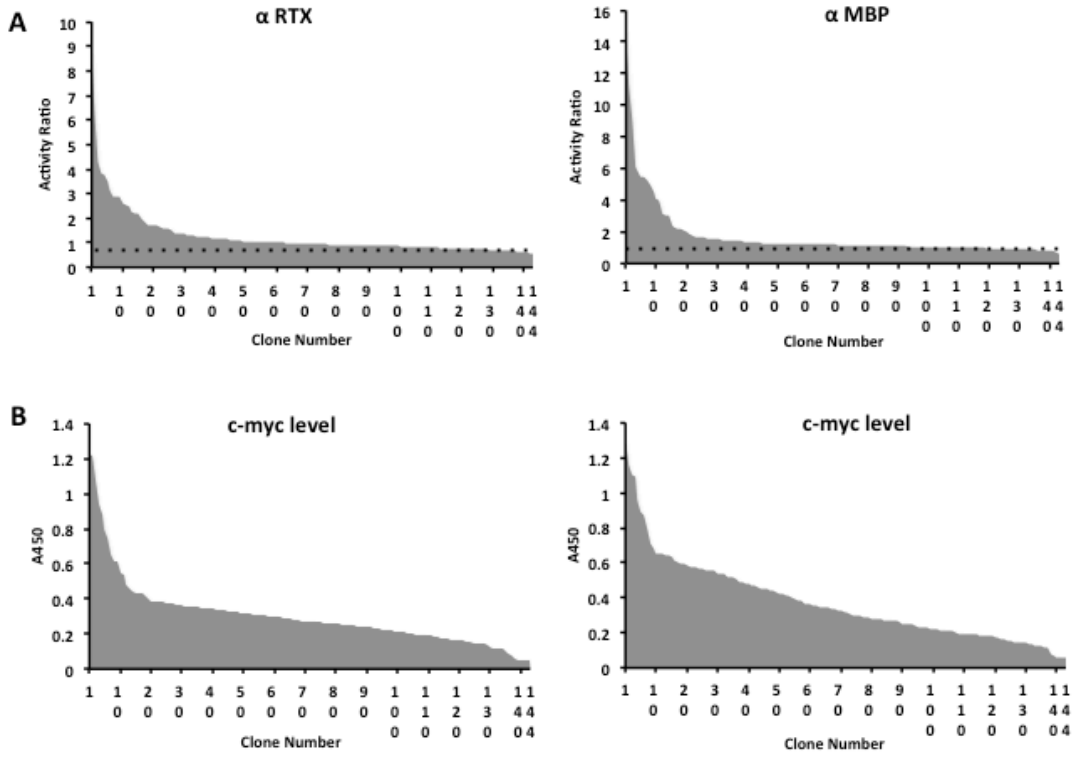


Figure 3.3: Phage ELISA screen of constructed sdAb libraries.

Binding affinity and selectivity of selected sdAb for two antigens was tested by phage ELISA screen after five rounds of panning. Enough number of sdAb-displayed phage was produced from single colonies and activity was measured by A450nm. (A) The activity ratio selected sdAb is shown as the signal from an antigen coated well divide by the signal from an uncoated control well (left; RTX and right; MBP). The average signal of unrelated control sdAbs is shown as a dashed line. (B) The relative sdAb display level on phage. sdAbs are appended with cMyc peptide tag at the C-terminus and antibody display level of each library was detected by anti-cMyc antibody.

Applying the binding activity ratio (determined by dividing the signal from an RTX-coated well by the signal from an uncoated well), a result of more than 3 represents

positive clones. A total of 7 positive clones for RTX_{751} and 14 positive clones for MBP were identified using this ratio (Table 3.1).

Table 3.1: Amino acid sequence alignment in CDRs after initial screening.

VH	CDR1					CDR2										CDR3											
	31	32	33	34	35	50	51	52	52A	53	54	55	56	57	58	95	96	97	98	99	100	100A	100B	100C	100D	101
WT	D	T	Y	I	G	R	I	Y	P	T	N	G	Y	T	R	W	G	G	D	G	F	Y		A	M	D
anti-RTX																										
V1	T	V	M	A	G	R	I	Y	P	Y	N	A	Y	T	D	L	Y	S	H	S	P	V	Y	A	M	D
V2	S	V	M	Y	G	R	I	Y	P	Y	G	S	A	T	Y	R	V	K	H	S	G	H	Y	A	M	D
V3	G	P	V	S	G	R	I	Y	P	T	N	G	Y	T	R	S	R	V	L	S	T	D	Y	A	M	D
V4	K	H	E	Y	G	R	I	A	P	A	S	A	Y	T	A	W	V	E	N	S	W	W	Y	A	M	D
V5	T	H	M	S	G	R	I	S	P	S	N	S	S	T	R	W	R	Q	H	S	T	I	Y	A	M	D
anti-MBP																										
V1	T	D	M	A	G	R	I	S	P	A	N	Y	A	T	S	G	R	S	R	S	W	K	Y	A	M	D
V2	S	D	M	S	G	R	I	Y	P	D	G	S	Y	T	A	T	L	R	H	S	I	K	Y	A	M	D
V3	S	A	K	S	G	R	I	Y	P	A	G	A	Y	T	S	L	K	V	I	S	H	L	Y	A	M	D
V4	D	D	L	Y	G	R	I	Y	P	A	N	A	Y	T	A	A	G	G	S	Y	S	Y	Y	A	M	D
V5	R	L	V	A	G	R	I	Y	P	Y	D	S	S	T	S	Y	G	G	A	Y	A	Y	Y	A	M	D

CDR sequences were selected from anti-RTX or anti-MBP library after 5 rounds affinity maturation. Residues that were fixed as acceptor sequence are shaded grey.

The top three ranked positive clones from each library were chosen for further characterization. First, we verified the binding affinity for three anti-RTX or anti-MBP variants by performing a monoclonal phage ELISA. All sdAbs and control antibody fragments were well displayed on the phage surface (data not shown). Isolated sdAbs were shown to acquire the target specificity after the bio panning (Figure 3.4). The signal of anti-RTX_V2 did not reach a plateau due to the limitation of phage concentration, but the trend after the highest concentration showed increment as others.

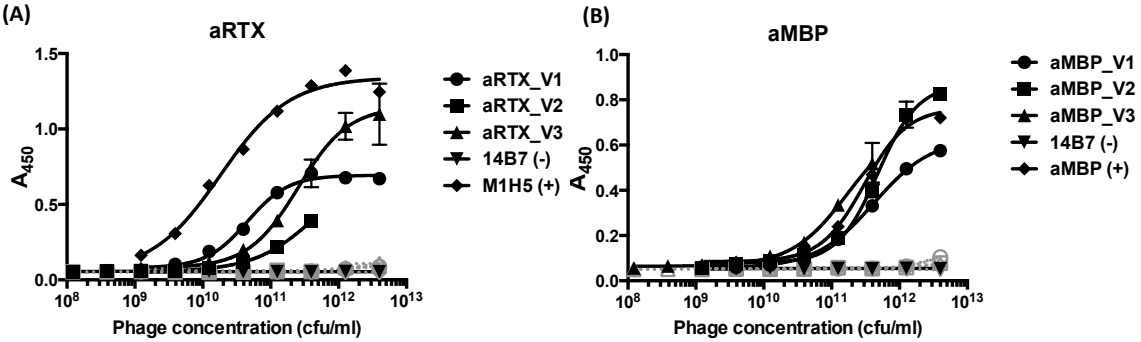


Figure 3.4: Phage displayed sdAb confirms target binding by monoclonal phage ELISA. (A) ELISA demonstrated RTX specificity of isolated three sdAb variants (solid lines; anti-RTX_V1 to V3). M1H5 scAb (diamond) that is known to bind RTX and 14B7 scFv, which is unrelated antibody (inverted triangle) were used as positive and negative control, respectively. No binding was observed to uncoated, blocked wells (grey, dashed lines). (B) Binding affinity and specificity to MBP was tested by ELISA and selected anti-MBP sdAbs showed increased binding for MBP.

In order to be utilized in protein co-crystallization, isolated sdAbs should express well as soluble proteins and exhibit high stability. To assess the solubility, wild type acceptor V_H (4D5.B1ag) and the three lead variants from each library were sub-cloned into the protein expression vector pAK400. The sdAbs were produced in *E.coli* periplasm, purified using IMAC, and followed by size exclusion chromatography (SEC) on superdex S75 column (GE healthcare) as previously described[6]. Monomer species of WT (Figure 3.5A) and anti-RTX sdAb_V2 were eluted at the expected volume. Two other variants of anti-RTX (V1 and V3) behaved as aggregation prone (Figure 3.5B). Anti-MBP sdAb variants also showed aggregation prone characteristics during antibody production and purification, including the anti-MBP_V2 that was predominantly eluted at

the volume of dimer size (Figure 3.5C). After SEC, only anti-RTX sdAbs were concentrated for further characterization. The size and purity of anti-RTX sdAbs were analyzed by SDS-PAGE gel (Figure 3.5D). Even though a C-terminal hexa-histidine tag fell off the sdAbs, the SDS-PAGE demonstrated that the size of purified sdAbs is close to the WT (14kDa and 12kDa, with or without hexa-histidine tags respectively) and the purity of antibody fragments reached about 90%. Even though we applied protein A selection strategy that is for obtaining autonomous behaved sdAb as WT [67], only one out of six identified sdAbs retained its original stability after applying diversity.

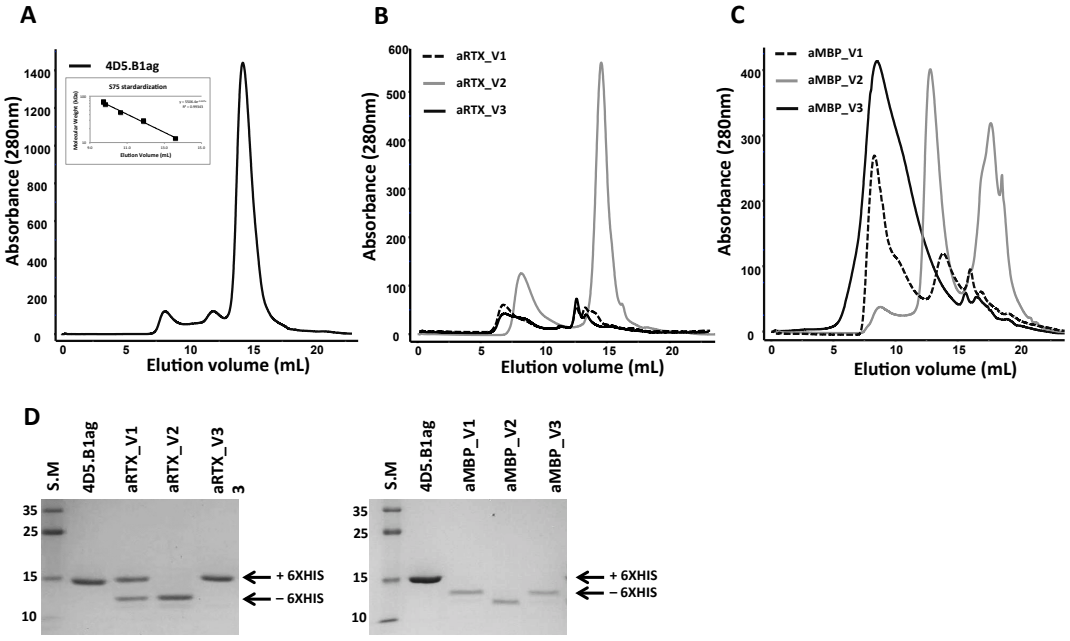


Figure 3.5: Expression and purification of selected sdAbs.

(A) The wild type sdAb (WT 4D5.B1ag, ~14kDa) was expressed in *E.coli* and purified by size exclusion chromatography on superdex 75 column. Monomer form of WT eluted at the expected volume. (B) anti-RTX sdAb variants were expressed and purified as previously described. Only the anti-RTX_V2 highly expressed as monomer, but V1 and V3 did not produced well in *E.coli*. (C) anti-MBP sdAb variants behaved as aggregation

prone. V1 and V3 were highly aggregated during production and purification, and V2 mostly eluted at the volume of dimers. (D) Size and purity of same amount of soluble anti-RTX and anti-MBP sdAbs were analyzed by SDS-PAGE gel. The hexa-histidine was fused at C-terminus of WT or three variants.

Concentrated anti-RTX sdAbs were used in an indirect ELISA to detect the binding specificity after the purification. Accessible sdAb variants (V2 and V3) as well as controls (the acceptor WT V_H and M1H5 scAb) were coated on the ELISA plate and biotinylated-RTX_{751} was applied. Streptavidin-HRP was added to detect specific interaction between the isolated sdAbs and RTX. Different from the no coat or acceptor V_H negative control, the binding ratio increased with the increased concentrations of anti-RTX sdAbs (Figure 3.6).

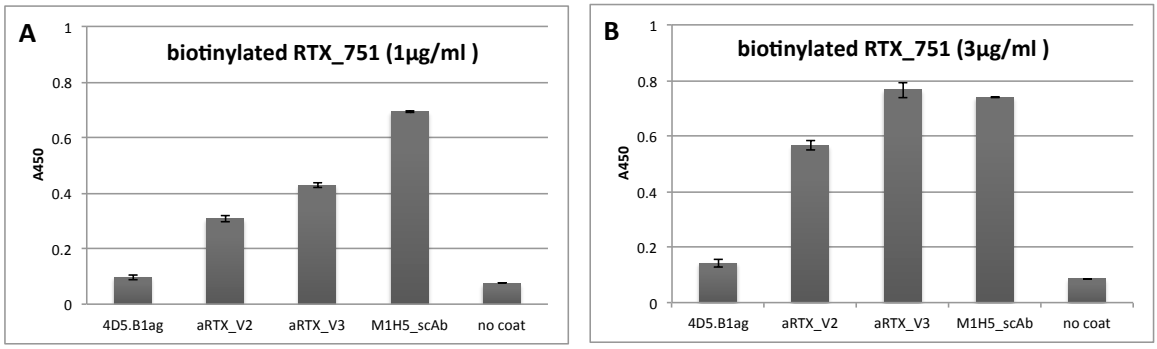


Figure 3.6: RTX binding affinity was confirmed using soluble sdAb ELISA

Soluble protein ELISA was performed to confirm RTX binding affinity with isolated sdAbs. Purified antibodies including aRTX sdAb variants, 4D5.B1ag and M1H5 as negative and positive control antibody respectively, were coated on wells followed by adding biotinylated RTX (A, 1µg/ml or B, 3µg/ml) as baits. The binding affinity was detected using HRP conjugated streptavidin.

A melting temperature (T_m) was measured to detect thermostability by monitoring proteins unfolding in the presence of a fluorescence dye that is able to bind with exposed hydrophobic regions during the heating. After the analysis, the T_m values of all isolated sdAbs decreased up to 30 degrees C (Figure 3.7 and Table 3.2), perhaps explaining the loss of solubility. Therefore, the 2nd generation library was constructed to further improve the biophysical characteristics of isolated sdAbs.

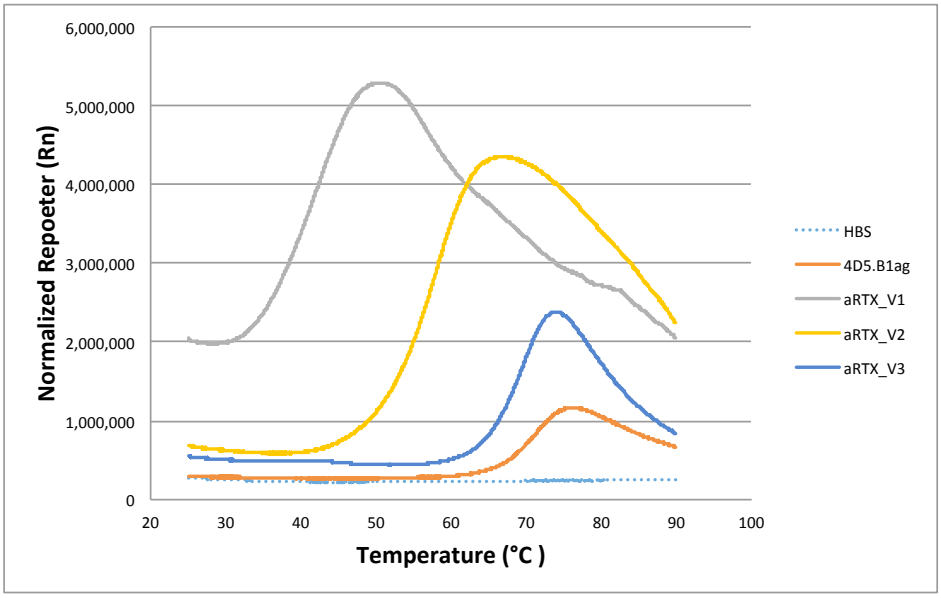


Figure 3.7: Three isolated anti-RTX sdAbs exhibit decreased thermal stability

Thermal stability of identified anti-RTX sdAbs from constructed library was determined by measuring apparent T_m. All sdAbs denatured at lower temperature compared to WT. Triplicate of each samples was performed and data was shown as an average.

Table 3.2: Biophysical characterization of lead sdAb candidates.

Table 3.2. Biophysical characteristics of nAb variants

	4D5.B1ag (WT)	aRTX_V1	aRTX_V2	aRTX_V3
Expression level (mg/l culture)	8.7	ND*	8.845 ± 0.08	ND*
Melting temperature (°C)	71.06 ± 0.30	41.68 ± 0.67	58.04 ± 0.57	69.54 ± 0.60

Expression level of anti-RTX_V2 is almost identical and purified as predominantly monomeric states. Thermal melting temperature of each variant was measured and T_m of variants decreased compared to WT, indicating less thermally stable. ND indicates not determined.

3.4.3. Improved sdAbs were isolated from 2nd generation of sdAb library

To improve the solubility or target affinity of 1st generation anti-RTXs without disturbing their behaviors, error prone PCR was performed based on the sequence of lead sdAbs using Mutazyme II DNA polymerase (Stratagene), following the manufacturer's instructions. To introduce mutations into the 1st generation of anti-RTX sdAbs, reaction samples (including primers named 5'pAkpel and 3'scforlong) were heated at 95°C for 2min, 30 cycles of 95°C for 30sec, 41.9°C for 30sec and 72°C for 1 min, then final synthesis for 4 min at 72°C. The mutation rate and diversity of libraries were validated by plasmid DNA sequencing, resulting in a 2nd generation library. The library size of 2nd generation was up to 10^7 CFU/ml, which is higher than that of the 1st generation, but we only got about 50% of sequence diversity because we adjusted the amount of template sequence for the error prone PCR not to lose the clones behaving the binding activity (Figure 3.8).

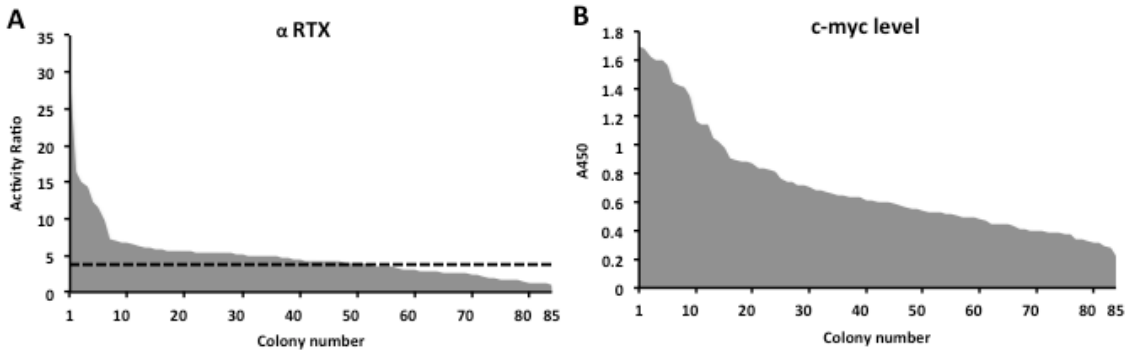


Figure 3.8: Construction of 2nd generation anti-RTX sdAbs library

2nd generation of sdAb library was constructed by error prone mutagenesis using the lead sdAbs as templates. (A) Enough number of colonies was screened by Phage ELISA to select biophysically improved sdAb isolates. (B) Levels of displayed c-Myc peptide tag that is appended to C-terminus of sdAb were detected using anti-c-Myc antibody.

The 2nd generation phage-displayed library was pooled and bio-panned with the target protein as previously described. After colony screening, five clones exhibiting high activity ratios were identified for sequencing and further characterization. Interestingly, the sequence of positive clones related to anti-RTX_V1 (V1a to V1e), and obtained only one or two amino acids replacement in the frameworks or CDRs before framework 3 regions. The mutations in residues 29 and 47 were recognized twice and the other regions were unique. After the sequence analysis, the biochemical or biophysical improvement of isolated sdAbs was verified by monoclonal phage ELISA. The binding activity of all of the 2nd generation sdAbs were moderately improved compared to that of anti-RTX_V1 by replacing only one or two amino acids. However, even though a decent increase of

binding affinity to RTX_{751} was detected, the improvement of solubility of sdAbs was negligible (Figure 3.9).

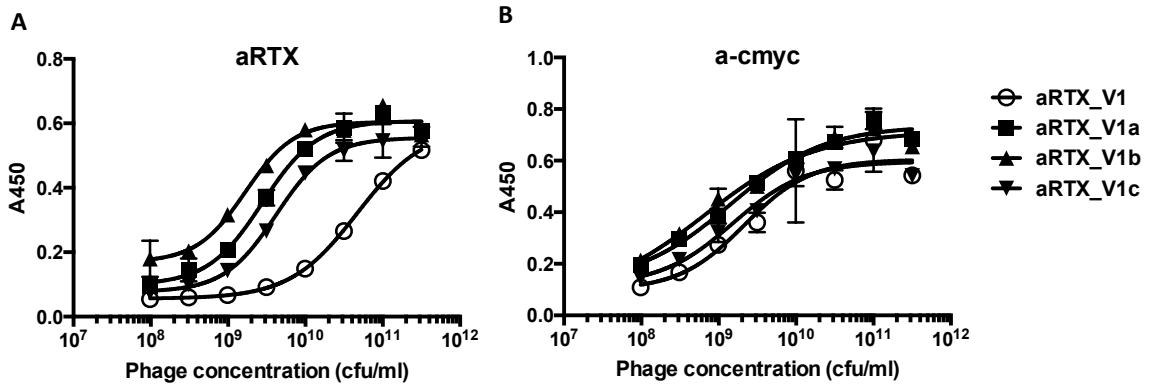


Figure 3.9: 2nd generation anti-RTX sdAbs retain antigen binding.

(A) Isolated sdAbs from 2nd generation library were all related with anti-RTX_V1 (V1a: L47F, V1b: I29T, V1c: S25P and I29T, V1d: G35S and D58E, and V1e: L47M; data for V1d and V1e not shown). Monoclonal phage ELISA demonstrated that selected 2nd generation sdAbs retained specificity (Solid lines) and improved antigen-binding affinity compared to the acceptor (anti-RTX_V1, open circle). (B) The display level of each sdAb on phage surface was shown to be comparable.

3.4.4. Identified anti-RTX sdAbs target to novel epitope

As a further characterization of all isolated 1st and 2nd generation sdAbs, we were interested in identifying whether previous sdAbs can recognize a unique epitope or not. The assessment of sdAbs to distinguish epitopes was determined by competitive binding ELISA using two well-known RTX-specific antibodies (M1H5 and M2B10)[78]. To test this, serially diluted soluble M1H5 or M2B10 scAb was mixed with a fixed concentration of monoclonal control antibody or sdAb-displayed phage and added to the RTX_{751} coated well. Then, the amount of bound phages was determined using the detection antibody

anti-M13-HRP. Decreased A_{450} signal indicated that two antibodies were competing with each other to the same overlapping or shared epitope, while a similar level of signal demonstrated that both antibodies could assess to a distinct epitope. According to the results shown in Figure 3.10, V2 and V3 seemed to slightly compete with M1H5 but not M2B10, implying that V2 and V3 could share the epitope that is close but not fully overlapped with M1H5. Also, V1 and all of the 2nd generation sdAbs did not compete with any control antibodies, so V1 and its related sdAb variants could access a distinct epitope from that of M1H5 and M2B10.

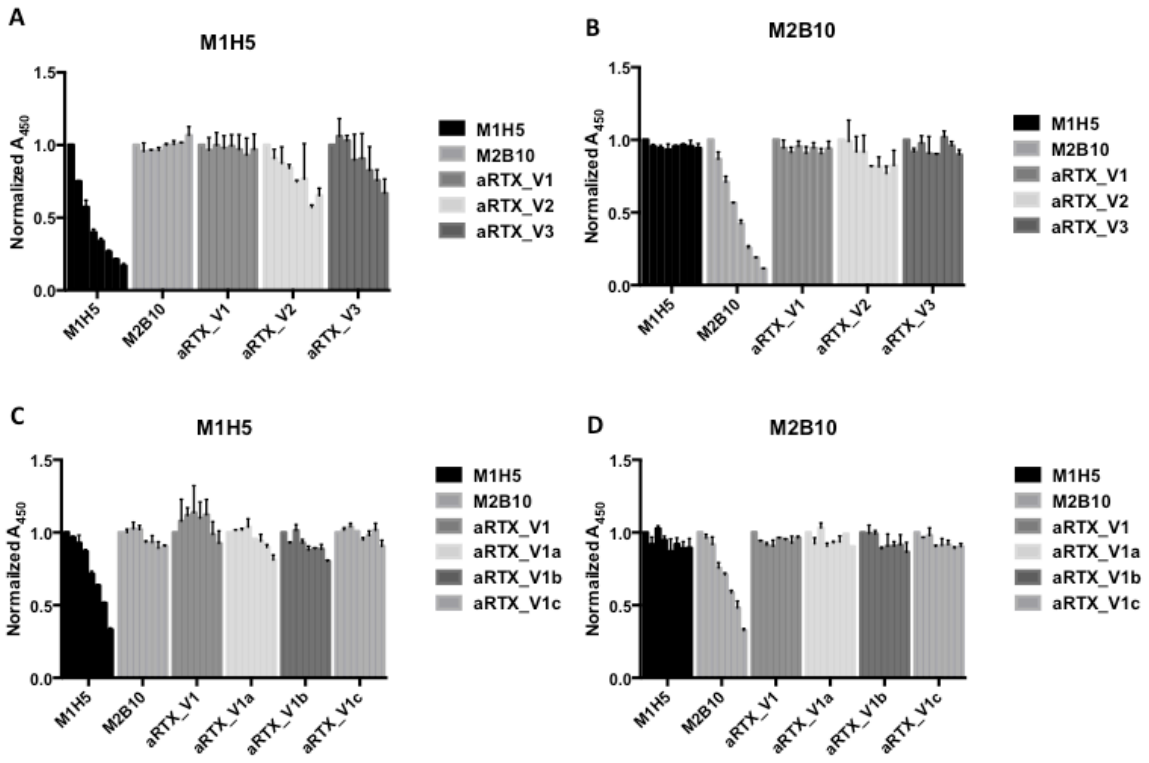


Figure 3.10: Identified sdAbs bind to a novel epitope.

Competition ELISA determined that anti-RTX_V1 and related 2nd generation sdAbs bound distinct epitope from that of M1H5 or M2B10. The level of bound phage displayed

anti-RTX_V2 and V3 slightly decreased when incubated with M1H5 indicating that they were competed with same or overlapped binding epitope (A). Plate was coated with soluble RTX and blocked. Serially diluted soluble M1H5 (A, C) or M2B10 (B, D) scAb was incubated for 30min prior to be mixed with sdAb-displayed phage. Bound phage was detected and the absorbance was normalized.

Another interesting approach to characterize sdAbs was to determine the binding specificity of isolated variants to the RTX of different species (Figure 3.11). Three different species of RTX, *Bordetella pertussis* (Bp), (B) *Bordetella bronchiseptica* (Br), and *Bordetella parapertussis* (Bpp) were tested by monoclonal phage ELISA (Figure 3.12). The level of normalized A_{450} slightly decreased when the anti-RTX_V2 and V3 simultaneously added with M1H5 but not M2B10, implying that V2 and V3 were shown to share the epitope only with M1H5 and could recognize distinct epitope from M2B10. The anti-RTX_V1 and all of the 2^{nd} generation sdAbs did not compete with any control antibodies so that V1 and its related sdAb variants could access to different epitope from M1H5 and M2B10.

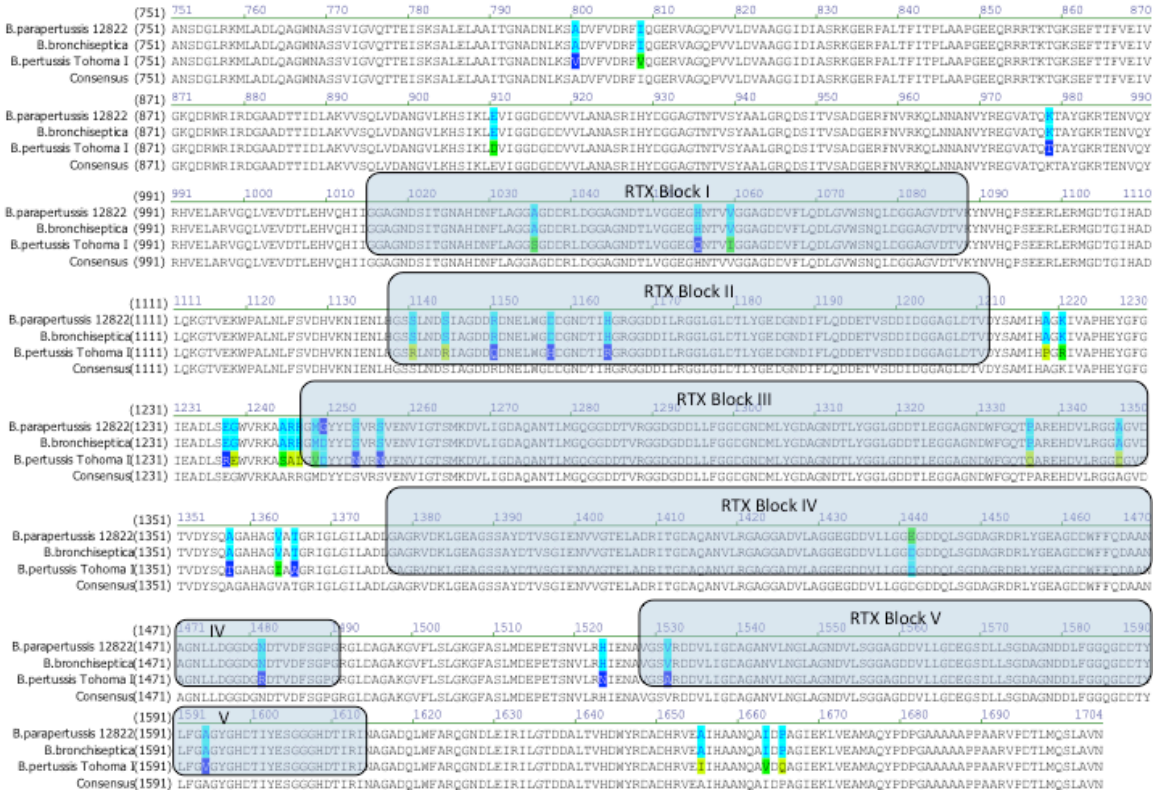


Figure 3.11: Amino acid sequence alignment of RTX* from three closely related *Bordetella* species.

The RTX from *Bordetella parapertussis* and *bronchiseptica* are only different by three residues, both highly homologous to *B.pertussis* RTX. The shaded boxes are the five RTX blocks, each containing multiple nine-residue repeats. *provided by Dr. Xianzhe Wang.

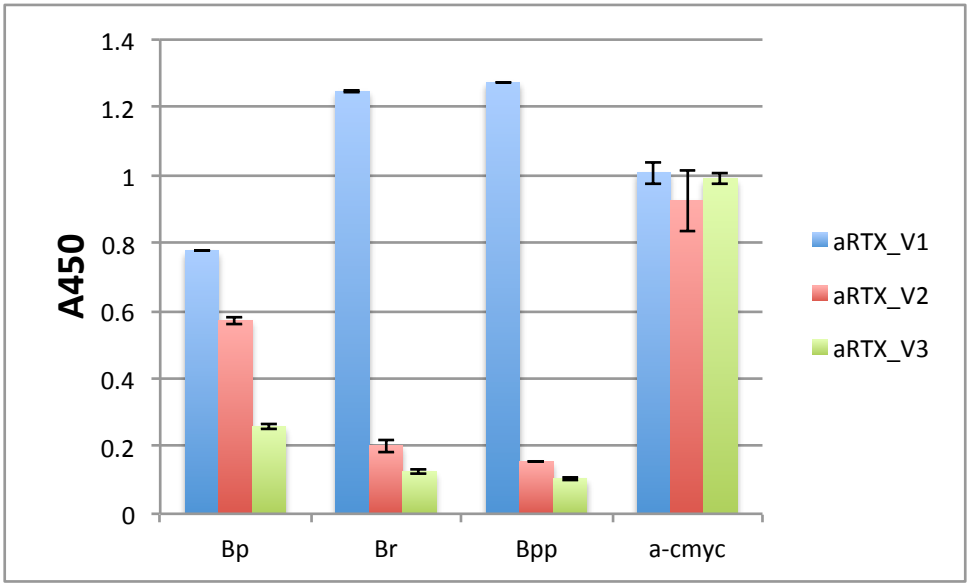


Figure 3.12: anti-RTX sdAb variants are able to bind different epitopes.

Binding specificity of variants from 1st or 2nd generation of library to different species of RTX, *Bordetella pertussis* (Bp) *Bordetella bronchiseptica* (Br) and *Bordetella parapertussis* (Bpp) was tested by phage ELISA when phage concentration is 5E11 (cfu/ml). Two control antibodies (data not shown) were also tested and only M1H5 showed specificity to Bp. The result demonstrated that unlike other two variants, anti-RTX_V1 was observed to bind to RTX of three species (blue bar), but V2 (red) and V3 (green) behaved as M1H5. All 2nd generation sdAbs were also tested and as expected, they exhibited as similar as anti-RTX_V1 thereof (data not shown)

3.5 DISCUSSION

Even though isolation and purification of *in vivo*-matured single domain antibodies by camel immunization is the general protocol for structural biology, the successful generation of sdAbs still takes few months and depends largely on the quality of antigens[83]. And only a number of limited antigens can be available to immunize for camel due to their highly toxicity, pathogenicity, and non-immunogenic traits. Recently, construction of a synthetic phage-displayed nanobody (V_HH) library against human

prealbumin (PA) and neutrophil gelatinase-associated lipocalin (NGAL) was reported [82]. Without *in vivo* affinity maturation steps, it was successful to obtain high affinity Nbs by randomizing CDR3 region. Taken together and to overcome obstacles of animal immunization, herein, we have developed a synthetic phage-displayed single domain antibody library with novel specificity, the Repeat-in-toxin domain (RTX), by employing random sequences into CDRs of the ideal autonomous V_H scaffold. Like camelid V_HH, 4D5.B1ag has proved to behave in a highly thermally stable manner, with good expression yield in *E.coli,* which are preconditions for crystallization chaperones. And unlike V_HH, its autonomous nature is evolved as completely independent of the CDR3 sequences[67] even the formal light chain interface is highly tolerant to mutation [66].

In order to generate efficient and diverse libraries that can acquire the unique specificity to RTX, we introduced completely random sequences in all three CDR regions using degenerate codons. The current size of our library is about 10^6 CFU/ml and the constructed library was validated as diverse enough to obtain high affinity clones after maturation with targets. From the library, we successfully isolated three promising sdAbs specific to RTX with high affinity. Interestingly, the sdAbs isolated from the library seemed to attempt to trade off binding affinity against solubility in the bacterial system. Even all selected sdAbs displayed and folded properly on the phage surface, sdAbs behaved high binding to the RTX domain have been dramatically lost the solubility in *E.coli*. But the variant, which expressed highly in *E.coli,* the binding affinity to the target protein was lower than we expected. Therefore, despite the successful identification of sdAbs against RTX, further optimization should be required to improve antibody

biophysical characteristics to be usable as a crystallization chaperone. Among the candidate sdAbs, two phage-displayed sdAbs can bind tightly, but behaved poorly for *E.coli* expression. The other candidate that was expressed at a similar level as WT showed a weaker binding affinity with the RTX domain. Previous reports demonstrated that inserting negatively charged residues at the edge of CDR regions of near-neutrally or negatively charged V_H scaffolds improved sdAb solubility and prevented aggregation without disrupting the binding affinity [84–86]. This approach can be applied to the isolated anti-RTX sdAbs because overall pH and all CDR regions, especially in V1, are recognized as neutral. Thus, introducing charged mutations can impact and increase the solubility.

To our knowledge, the library constructed here is the first experimental study to generate sdAbs *in vitro* specifically for intrinsically disordered proteins that are not crystallized yet. Compared to previous immunization strategies, the proposed technology to generate a phage-displayed single domain antibody library is a more time- and effort-saving way to develop antigen specific antibodies even though further optimization stage is required. Recently, plausible 3D structures and functions of RTX fragments have been evaluated by homology-based modeling and their capability for interaction with target membrane proteins have been suggested [87]. But the experimental characterization of its function is still remaining to demonstrate due to the lack of crystal structure of RTX. As a crystallization-friendly scaffold [88], the advance of synthetic sdAb phage-displayed library will become a faster and economical way to address the structures and functions of target proteins that are hard to crystallize. In conclusion, we have developed a synthetic phage-

displayed single domain antibody library, identifying and isolating sdAbs against an intrinsically disordered protein target. While the conversion from phage-displayed version to soluble protein did not retain solubility or stability, we anticipate that the synthetic single domain antibody library protocol will increasingly contribute to determining the protein crystal structures for describing many structurally hampered target proteins.

Chapter 4: De novo design of peptide specific antibody

4.1 Chapter Summary

Computational antibody engineering to date has been used almost exclusively to improve binding affinity or biophysical characteristics. *In silico* design of antibodies with novel binding specificity could save enormous time and resources compared to conventional immunization or synthetic library selection strategies. One such method, OptCDR, employs *de novo* complementarity determining region (CDR) design to engineer antibody-antigen interactions but has not yet been experimentally tested. Toward this aim, OptCDR was used to design CDRs binding the minimal FLAG peptide (sequence: DYKD), which were grafted onto a previously crystallized single-chain variable fragment (scFv) acceptor framework. Fifty designs, employing two different CDR modeling strategies, were chosen based on similarity of predicted antibody-peptide interactions to experimentally characterized antibody-peptide complexes. Genes were synthesized and screened for peptide binding by phage ELISA. Four antibodies, designed entirely *in silico*, possess high specificity and good affinity for the DYKD sequence in phage-displayed scFv and soluble Fab formats. These results demonstrate that antibody specificity design based on CDR modeling is a viable engineering strategy.

4.2 Introduction

Antibodies are one of the most important protein classes, widely used in commercial diagnostics and therapeutics, garnering global annual sales of $75 billion in 2013 [89]. They have conventionally been developed through experimental approaches, such as animal immunization followed by hybridoma generation and, more recently,

screening of synthetic libraries. These approaches are limited for targets that harbor poorly accessible epitopes or require precise molecular engagement to achieve the desired biological effects. Moreover, since sequence diversity expands at a rate of 20^n, where n is the number of randomized amino acids, synthetic library sizes rapidly approach the limit of what can be reasonably screened using display technologies. Computational approaches have the potential to dramatically reduce the resources required for antibody discovery while increasing success rates for challenging targets. The growing utility of *de novo* protein design is demonstrated by a number of recent successes with therapeutic potential, including the design of anti-HIV [90] and antimicrobial peptides [91], epitope mimics for vaccination [92] and influenza inhibitors [93].

Antibody-antigen interactions are dominated by the complementarity determining regions (CDRs), three on each of the heavy and light variable domains. Typically, CDRs bind antigens by forming a shape-complementary pocket with favorable interactions distributed throughout the CDRs. Although there are exceptions to this typical binding mode, such as the VRC01 class of broadly neutralizing anti-HIV antibodies [94], they are rare. Soon after the first antibody structures were solved, it was recognized that CDR backbones cluster into distinct groups of canonical structures [95] with unique amino acid sequence preferences. This has facilitated the development of several methods for predicting antibody tertiary structures from their amino acid sequences [96]. However, despite successes in *de novo* protein design in general, computational antibody engineering to date has typically focused on improving characteristics of existing antibodies rather than designing novel specificities. Examples include identifying charged

mutations that confer thermo-resistance [97], guiding affinity maturation [98], improving association rates [99] and identifying aggregation prone regions [100]. A recent exception to this trend was the development of a method to rationally design antibodies to bind epitopes in disordered portions of the antigen [101]. While effective, this method uses an atypical binding mode and is limited to epitopes in disordered regions of protein structure.

To address these limitations, we previously reported development of a computational method, Optimal CDR (OptCDR) for *de novo* design of antibody binding interfaces complementary to specific three-dimensional epitopes [20]. This approach is unique in that it selects canonical CDR structures, then appends them with amino acid residues, using energy minimization to refine the structure and maximize predicted interactions with the target epitope. Here, we used this approach to design CDRs binding the minimal FLAG peptide (sequence: DYKD) [102]. The FLAG peptide is widely used for protein detection and affinity purification in conjunction with commercially available antibodies binding the extended form of the peptide. Additionally, the FLAG peptide has been incorporated into external protein loops without altering core structure [103], allowing for its use in combination with an anti-FLAG crystallization chaperone to guide high-throughput structural biology efforts [31]. As a target sequence, its short length restricts the range of possible conformations while its residues are capable of forming nonpolar, hydrogen-bonding and cation-π interactions. Indeed, the FLAG peptide has been observed to form specific structures [104] and thus is a candidate to assess our ability to design antibodies binding small, conformational epitopes.

Two libraries were created by full design of all six CDRs or only the three heavy chain CDRs. Fifty designs were synthesized in single chain variable fragment (scFv) format and screened for activity by phage ELISA. Four antibodies, two from each library, exhibited FLAG peptide binding which was maintained after conversion to soluble antigen binding fragment (Fab) format. These antibodies bind the minimal FLAG peptide with half maximal effective concentrations (EC_{50}) ranging from 4 - 50 nM and are extremely specific, as even conservative mutations result in complete loss of binding. As Fabs, they expressed well, were predominantly monomeric and retained binding after several weeks of storage. These results demonstrate that *de novo* CDR design to target specific epitopes is a viable approach.

4.3 Materials and Methods

4.3.1 Computational antibody design

A model for the complete FLAG peptide (amino acid sequence: DYKDDDDK) ligand was created using the build function in PyMOL (DeLano Scientific) and docked to the Fab structure of a partially refined commercial FLAG peptide-binding antibody [103] using the ClusPro docking server [105] on antibody mode to orient the peptide in a position likely to be compatible with binding. Two glycine residues were then appended at both ends to mask terminal carboxyl and amine groups. This antigen model was used for OptCDR-guided design as previously described [41] with some updated modifications, including the use of an updated database of canonical structures [106] and allowing for the presence of framework residues during calculations, which were never modified in any

way. A peptide-binding scFv (PDB ID 3NN8)[6] served as the framework structure upon which CDRs were built. Separate approaches were used to create two unique libraries: a full design in which all six CDRs were optimized (termed EEf) and a heavy chain-only design (EEh) that constrained light chain CDR sequences as wild-type. Briefly, the first step of OptCDR is the selection of CDR canonical structure backbones that are most likely to allow favorable interactions with the antigen. Canonical structures were previously identified by clustering peptide backbones isolated from a set of non-redundant antibody structures from the Protein Data Bank (PDB, www.rcsb.org). Each cluster member possesses a backbone atom RMSD of no more than 1.5 Å from the canonical structure, and clusters possessing only one or two members were discarded. All clusters were retained for CDR H3 regardless of cluster size. During the second step, the amino acid side chains are initialized one-at-a-time onto the chosen canonical structure using a rotamer library [107], energy functions, and a mixed-integer linear programming optimization formulation. Additional sequence-based constraints are included. For canonical structures from clusters of at least 30 members, the amino acid at each position is constrained to only those previously observed within the cluster. If the chosen canonical structure is derived from a cluster with fewer than 30 members, then amino acids of the same type, for example only charged residues (D, E, K, R and H), as those previously observed within the cluster are allowed at each position. This is followed by several thousand iterations of a modified version of the Iterative Protein Redesign & Optimization (IPRO) [108] procedure, which simultaneously refines the CDR backbone structures and amino acid sequences. The refined parent models, appended with a '.0'

label, each possess a unique structural solution. To generate libraries, mutants were selected by tracing backwards through the computational affinity maturation process for each parent model and extracting unique CDR sequences that differ by 2-11 amino acids. These derivative designs are appended with a ranked '.1-.6' label, with a higher number denoting further distance in design space from the final refined parent model and thus predicted to have less optimal antigen interactions. This returned about 300 designs based on 30 parent models total, each of which was visually inspected by PyMOL for agreement with known antibody-peptide interactions. Ten unique designs were selected from the EEf library, based on three parent models, and 40 designs selected from the EEh library, from 12 parent models. For each parent model, the depth of the FLAG peptide into the V_L/V_H interface was measured as the z-axis distance from the plane made from the N- and C-terminal residues of the heavy and light CDR3s. This analysis was initially run with the total distance from the centroid of the four residues, but was considered less informative than z-axis depth due to the variable peptide orientations.

4.3.2 Library synthesis

Genes for all 50 designs were synthesized by protein fabrication automation as described previously [109]. Briefly, amino acid sequences for V_L and V_H were reverse-translated using an *E. coli* class II codon table and combined in the scFv format in the orientation V_L-$(GGGGS)_4$-V_H. Bidirectional *Sfi*I cloning sites were appended to allow cloning into the pAK vector series [110]. Synthesized products for each library were pooled and cloned into the phage display vector pMopac24 [111], which expresses the scFv fused to

a truncated *C*-terminal pIII protein to facilitate M13 phage display and co-expresses the protein chaperone Skp to increase active scFv yield. Colony PCR confirmed >95% cloning efficiency, and ten colonies for each library were sequenced using the primer 5'pAKpel [54] to identify the proportion of correctly assembled genes. All oligonucleotides were purchased from Integrated DNA Technologies and all DNA sequencing was performed at the University of Texas at Austin Core Facility.

4.3.3 Construction of FLAG variants

All peptide variants were constructed by site-directed mutagenesis of the pAK400/14B7 plasmid that contains a minimal DYKD epitope at the protein n-terminus and employs the irrelevant scFv 14B7 as a carrier protein[112]. Mutagenic primers (5'14B7_Flag_mut, 3'14B7_Flag_mut1 to mut6, 5'14B7_Flag_c-term, and 3'14B7_Flag_c-term) were used to alter or eliminate the n-terminal DYKD and to introduce the full FLAG peptide and the n-and c-terminal locations. Briefly, the parental plasmid was methylated using CpG Methyltransferase (*M.sssI*, NEB) for 1.5 h at 37°C and then used as template in PCR-based site-directed mutagenesis. The PCR reaction followed the commercial instructions, after which *Dpn*I was added to digest methylated template plasmid. The products were transformed into *E. coli* strain XL1-Blue and sequenced at the University of Texas at Austin Core Facility. The previously described MBP-KEE[49] ligand was used to detect binding by the parental a-EE antibody. Production and purification of all peptide ligands was performed as previously described[49,112].

4.3.4 Phage production and ELISA

24 individual colonies from each library were screened using phage ELISA as previously described to test binding activity to short flag peptide (DYKD) at the N-terminus. Briefly, high-binding 96-well plate (Costar) was coated with 14B7-Flag (N, DYKD) or 14B7-His (4μg/ml), or anti-c-Myc antibody (Sigma 9E10, 1-2μg/ml) overnight at 4°C. Phage, diluted in blocking buffer (5% non-fat milk in PBST) were added and incubated for 1 h at RT. After 3 times washing with PBST, anti-M13-HRP (1:3000 in 2.5% non-fat milk in PBST) was added and incubated for 1 h. After the final washing, bound phage was detected with TMB (sigma) by measuring absorbance at 450nm. Clones showed high absorbance that indicated high peptide binding activity, were chosen to further characterization.

Another phage ELISA was performed to test the binding activity and specificity of selected clones. The 14B7 FLAG peptide variants (described above, at 4μg/mL) or anti c-myc antibody (1μg/mL) were coated overnight at 4°C. After 1 hour-blocking wells with blocking solution (5% non-fat milk in PBS-0.05% Tween -20), serially diluted phage was added and incubated for 1hour at room temperature. After washing with PBS-0.05% Tween -20 for 4 times, anti-M13-HRP (1:3000, in blocking buffer) was allowed to incubate for 1 hour. Signal was developed using TMB and detected at 450nm after quenching with 1N HCl.

4.3.5 Fab protein purification and binding assay

For expression as soluble Fabs, the V_H or V_L domain was amplified and sub-cloned into pFabF vector using conventional cloning or Gibson assembly method. Then, pFab/EEh14.3, pFab/EEh13.6 pFab/EEf15.4, and pFabF/EEf15.4a were co-transformed into BL21 with the pBAD33_Skp, which is a chaperone to improve antibody solubility. A single colony was inoculated in 12ml 2xYT media containing ampicillin (200μg/ml), chloramphenicol (34μg/ml), and 2% glucose and incubated for overnight at 37°C. The starter culture was transfer to 250ml TB containing appropriate antibiotics and 0.1% glucose at an OD_{600} of 0.1 and cells were grown at 25°C until OD_{600} = 0.5. 0.2% L-arabinose was added to induce skp protein 30min before adding 1mM IPTG to express Fabs, and then antibodies were expressed for additional 4 h. After performing osmotic shock and dialysis as previously described, IMAC resins were applied to dialyzed supernatant, washed with IMAC wash buffer (25mM Tris, 10% glycerol, 100mM NaCl, 20mM imidazole, pH9.0), eluted with elution buffer (25mM Tris, 10% glycerol, 100mM NaCl, 0.5M imidazole, pH9.0). Antibodies were further purified by SEC and S200 with a HBS (pH7.4) running buffer. Protein production of FLAG peptide carrier proteins (14B7 and its variants) were expressed and purified as described in Chapter 2. To measure the activity and specificity of Fab, ELISA was performed similarly as Phage ELISA. Briefly, high-binding 96-well plates (Costar™) were coated with 14B7, 14B7-Flag and its peptide variants (10μg/mL) and incubated at 4°C overnight. All wells were blocked using blocking buffer (5% non-fat milk in PBS-0.05% Tween-20) for 2 h at RT prior to add serially diluted purified Fab/EEh14.3 or Fab/EE in blocking buffer. The binding was

detected by goat-anti-human kappa-HRP (1:1000, Southern Biotech, 20060-05) for 1 h at RT. Commercial monoclonal anti-Flag M2-HRP (Sigma-Aldrich, A8592) was used for control antibody.

4.4 RESULTS

4.4.1 *In silico* antibody design by OptCDR

The OptCDR method of computational antibody engineering designs antigen-binding CDRs using *de novo* loop modeling guided by rules derived from CDR canonical structure clustering [41]. OptCDR was previously validated using *in silico* methods by comparison to high-resolution structures of antibody-antigen complexes. However, OptCDR has yet to be experimentally tested. Toward this aim, we used OptCDR to generate FLAG peptide-binding antibody fragments. The FLAG peptide is widely used for protein detection and affinity purification; an anti-FLAG alternative to commercial sources could reduce cost or be used for unique applications requiring smaller fragments, such as scFv. Additionally, FLAG peptide epitopes have been incorporated into external protein loops without altering core structure [103], allowing for their use in combination with an anti-FLAG crystallization chaperone to guide high-throughput structural biology efforts [31].

To choose an acceptor framework onto which CDRs could be grafted, used an scFv we previously engineered and crystallized to bind an EE peptide (αEE scFv; EE sequence EYMPME) [6]. A model of the FLAG peptide was built as described in the methods section, and OptCDR was used to create two libraries: EEf, in which all six

CDRs were designed to fully test the modeling process, and EEh, in which only the three heavy chain CDRs were designed and the light chain sequence was constrained to wild-type. We hypothesized that grafting only three CDRs would be less likely to disrupt the framework structure and that heavy chain diversity would be sufficient to confer specificity, as observed with our other peptide-binding scFvs and in previous anti-peptide antibody engineering studies [113].

Thirty unique models were generated for each library, and the top 15 by calculated interaction energy were further affinity matured using the library design function. FLAG peptides were docked parallel to the V_L-V_H interface for a majority of EEh models. Those that deviated from this binding mode, which is typically observed for peptide-binding antibodies [114], were discarded. Selection was further refined by choosing models that targeted the tetrad residues DYKD, known to form the core epitope for commercially-available FLAG-binding antibodies [103]. Ten EEf library (Figure 4.1A) and 40 EEh library (Figure 4.1B) members were chosen for synthesis and testing. The models are labeled by their library ID (EEf or EEh), followed by model number (1-24) and derivative number (.1-.6). Derivative numbers represent CDR solutions that employ the same canonical structures as the parent model (.0) but differ in CDR sequence. Because derivatives were selected by tracing backwards through the computational affinity maturation process, as described in the methods section, only the parent model structures were available for comparison in the following analysis.

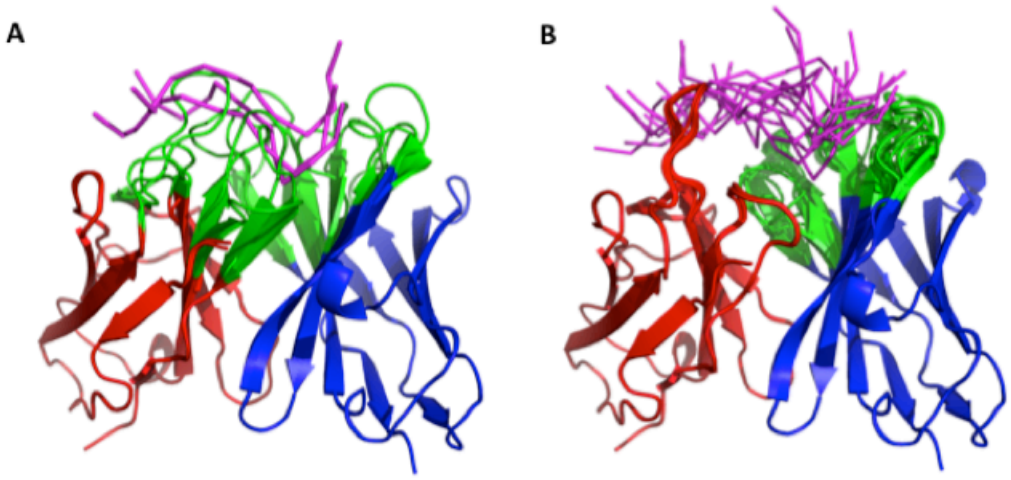


Figure 4.1: Alignment of OptCDR-generated models.

scFv light chain (red), heavy chain (blue) and designed CDRs (green) are shown docked to the FLAG peptide (magenta) in predicted conformations. Parent models of all designs chosen for synthesis are aligned and only vary in CDR conformations. (A)All six CDRs were optimized during EEf library modeling, while (B) the light chain was constrained as wild type for the EEh library.

Multiple metrics were used to compare the OptCDR-generated designs. Interaction energy, defined as the minimized energy of the complex minus the energy of the CDRs and the energy of the antigen individually, ranges from -191 to -609 kcal/mol for EEf designs and from -129 to -476 kcal/mol for EEh designs. Negative scores indicate favorable interactions; these scores are much lower than previous peptide-binding OptCDR designs, which ranged from -88 to -176 kcal/mol [41]. Contacts are defined as the number of CDR atoms within 3 Å of the antigen, and polar contacts were determined

using PyMOL. The number of total contacts for EEf designs ranged from 52-78, while EEh designs generally contributed fewer contacts, ranging from 21-60. The same trend continued when considering polar contacts only, where the EEf library ranged from 8-12 and EEh library from 1-10. Constraining the light chain sequence during EEh library modeling prevented *in silico* optimization of these CDRs, likely contributing to more unfavorable (less negative) predicted interaction energies and fewer CDR contacts compared to the EEf library. However, both overall and polar contacts were distributed nearly evenly between the heavy and light chains for the majority of the EEh designs, as both the average and median percentage of contacts contributed by the heavy chain was between 50-60%.

Because the FLAG peptide is highly charged, with five Asp and two Lys residues, we also calculated the net charge of the designed CDRs to ensure that OptCDR was not simply generating CDRs on the basis of charge complementarity. EEf overall net CDR charge was positive, though the distribution of positive and negative charges were observed among the individual CDRs. The overall net charge for EEh designs ranged from -1 to +4 (Table S4, only parent models shown), demonstrating that there was no strong bias for selection of charged CDRs. The OptCDR design procedure limits the proportion of all charged residues in a given CDR to within one standard deviation of the average percentage observed in antibodies deposited to the PDB. This constraint likely helped to prevent bias for optimization exclusively based on charge complementarity.

4.4.2 Identification of designed antibodies with peptide-binding

To test activity of the selected designs, scFv genes in V_L-linker-V_H orientation were constructed using automated protein fabrication. The synthesis products for each individual library were pooled, cloned *en masse* into a phagemid display vector and transformed into *E. coli*. Cloning errors were determined to be less than 5% based on colony PCR, while gene synthesis errors were found to be present in nearly 60% of clones by DNA sequencing, primarily localized to the repetitive GlySer linker region. Accounting for synthesis and cloning errors, sufficient colonies were screened to sample the library size at about three times coverage (72 clones for EEf and 368 for EEh). Phage from individual colonies were propagated and tested by ELISA for binding to two forms of the FLAG peptide: the intact DYKDDDDK peptide at the c-terminus and the minimal DYKD tag at the n-terminus of a carrier protein. The peptide binding activity ratio, defined as the A_{450} value from a peptide-coated well divided by the A_{450} value from an uncoated control well, is reported for each clone to account for non-specific binding. Both the EEh and EEf libraries included multiple clones binding the DYKD-protein with activity ratios up to seven (Figure 4.2A and B), while binding was much weaker for the highly charged protein-DYKDDDDK target (data not shown). The range of scFv display levels was similar for both libraries, as assessed by binding of a c-Myc peptide tag at the c-terminus of the scFv to an anti-cMyc antibody by ELISA (Figure 4.3C and D).

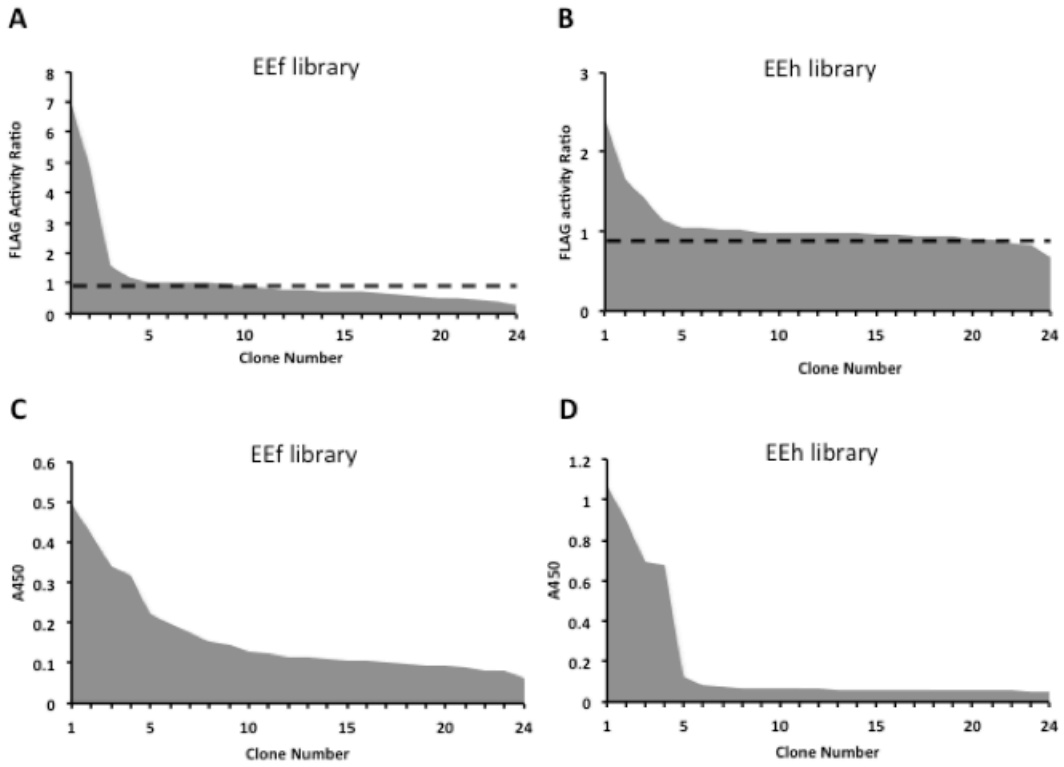


Figure 4.2: Phage ELISA screen of OptCDR designs in scFv format.

(A) EEf library gene synthesis products were pooled and cloned en masse to a phage display vector, scFv-bearing phage were produced from single colonies and activity tested by ELISA. Shown is the A_{450} ratio for a FLAG-coated well over a control well. The dashed line indicated the average signal for an unrelated control scFv. Enough colonies were tested ensure 3-5 times coverage of the library diversity, after accounting for synthesis and cloning errors. (B) EEh library clones were evaluated in a similar fashion. The level of c-Myc tag appended to *C-terminus* of scFv in (C) EEf library and (D) EEh library was monitored to test relative display level.

This demonstrates that display level likely did not contribute to observed differences in library performance. Clones with high activity ratios from each library were sequenced to

reveal four unique sequences with varying CDR lengths and sequences: EEh13.6, EEh14.3, EEf15.4 and EEf15.4a, the latter two differing only in CDR H1 (Figure 4.3).

	VL CDRs																																
	CDR1																	CDR2							CDR3									
scFv name	24	25	26	27	27A	27B	27C	27D	27E	28	29	30	31	32	33	34	50	51	52	53	54	55	56	89	90	91	92	93	94	95	95A	96B	96	97
Acceptor	R	S	S	Q	S	I	V	H	S	N	G	N	T	Y	L	E	K	V	S	N	R	F	S	F	Q	G	S	L	V	P	–	–	P	T
EEf15.4	R	S	S	N	–	–	–	–	–	A	R	S	G	S	L	E	D	G	N	N	R	F	S	S	A	F	D	Q	T	N	K	Y	V	G

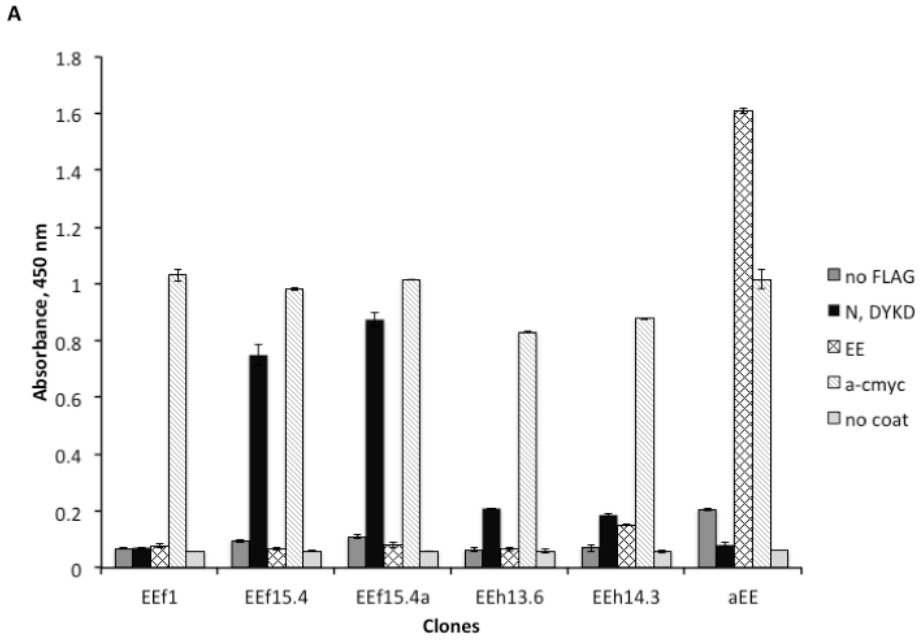
	VH CDRs																																		
	CDR1													CDR2												CDR3										
scFv name	26	27	28	29	30	31	32	33	34	35	35A	35B	50	51	52	52A	52B	52C	53	54	55	56	57	58	93	94	95	96	97	98	99	100	100A	100B	100C	100D	101	102
Acceptor	G	Y	S	L	S	T	S	G	M	G	V	N	H	I	Y	–	–	–	W	D	D	D	K	R	A	R	R	G	G	S	S	H	Y	Y	A	M	D	Y
EEf15.4	G	F	S	I	K	G	A	N	V	N	–	–	H	V	R	–	–	–	G	D	A	S	T	R	A	D	R	K	M	Y	S	F	Y	S	G	G	E	A
EEf15.4a	G	Y	I	G	S	S	H	T	V	N	–	–	H	V	R	–	–	–	G	D	A	S	T	R	A	D	R	K	M	Y	S	F	Y	S	G	G	E	A
EEh13.6	G	D	S	L	S	S	F	N	A	G	V	N	H	G	A	–	–	–	–	V	M	S	T	R	A	K	S	T	G	R	Y	–	–	–	–	–	D	F
EEh14.3	G	D	S	L	S	S	Y	N	A	G	V	N	H	M	A	–	–	–	–	G	V	S	T	R	V	R	N	E	W	S	G	–	–	–	–	–	A	F

Figure 4.3: Alignment of CDR sequences for EEf and EEh library models chosen for screening.

CDRs from the scFv used as the framework to accept OptCDR designs are also shown. CDRs are numbered Kabat standards, and CDRs are defined according to OptCDR's sequence-based rules.

To confirm peptide binding activity and specificity, monoclonal phage ELISAs were performed with these four clones and the parental anti-EE scaffold antibody. Each purified phage preparation was titrated on wells coated with a carrier protein presenting c-terminal DYKDDDK, n-terminal DYKD, or c-terminal or internal EE-peptides, in addition to a tag –free carrier protein and an anti-cMyc antibody to monitor scFv display level and uncoated wells to monitor non-specific binding. Each of the four unique clones identified in the screen bound specifically only to the n-terminal DYKD protein, while the parental anti-EE protein used for the framework only bound the EE peptide. A control library clone (EEf1), which exhibited a high activity ratio in the initial screen but

exhibited a high level of scFv display showed no binding to any peptide ligand (Figure 4.4A). To assess the role of each residue in the minimal DYKD peptide in antibody recognition, a series of peptides with conservative changes were generated and expressed at the carrier protein n-terminus: DFKD, DYRD and EYKE. A final variant was generated, ADYKD, to determine whether peptide position at the extreme n-terminus is essential for recognition (Figure 4.4B). Remarkably, binding to each of these variants was similar to background binding levels. This data demonstrates that these antibodies only the minimal FLAG peptide when present at the extreme N-terminus and that each residue in the minimal DYKD peptide is essential for binding.



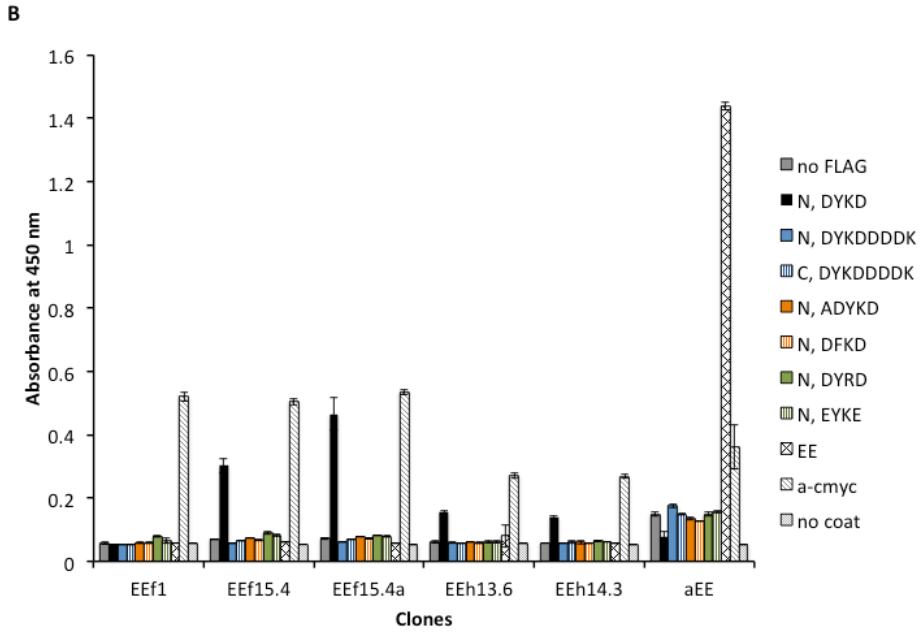


Figure 4.4: Initial screen of OptCDR designed molecules identifies several with peptide-binding activity.

(A) Phage displayed scFvs were analyzed by monoclonal phage ELISA for binding to proteins presenting various peptides. Shown is the raw absorbance at 450nm for wells containing 10^{11} pfu, which was in either the linear dose-response of background range for each titration curve. EEf clones 15.4 and 15.4a, EEh clones 13.6 and 14.3 showed high activity ratios in the initial screen; EEf_1 is a control which did not have a high activity ratio in the initial screen while anti-EE is the parental EE-peptide binding scFv which provided the framework to support the OptCDR designs. Ligands tested include carrier protein with no tags, with a n-terminal DYKD, EE tag, anti-c-Myc antibody to assess scFv expression level and no coat control. (B) Phage displaying scFvs were assessed for peptide binding specificity using a series of FLAG peptide variants, including carrier protein with no tags, with an n-terminal DYKD, n-terminal DYKDDDDK, c-terminal DYKDDDDK, n-ADYKD, n-DFKD, n-DYRD, and n-EYKE. Controls included EE peptide and anti-c-Myc antibody to assess scFv expression levels and no coat control.

4.4.3 Designed antibodies retain peptide specificity as soluble Fab

To test whether these antibodies retained DYKD peptide binding activity as purified proteins, the variable regions were cloned into bacterial periplasmic expression vector, followed by sequential IMAC and SEC purification steps. The EEh14.3 Fab eluted predominantly as monomer during the final SEC purification step. Analytical SEC of purified monomer showed Fab elution at the expected volume (10 mL), and Fab migrates at the expected sizes during SDS-PAGE under reducing and non-reducing conditions (Figure 4.5).

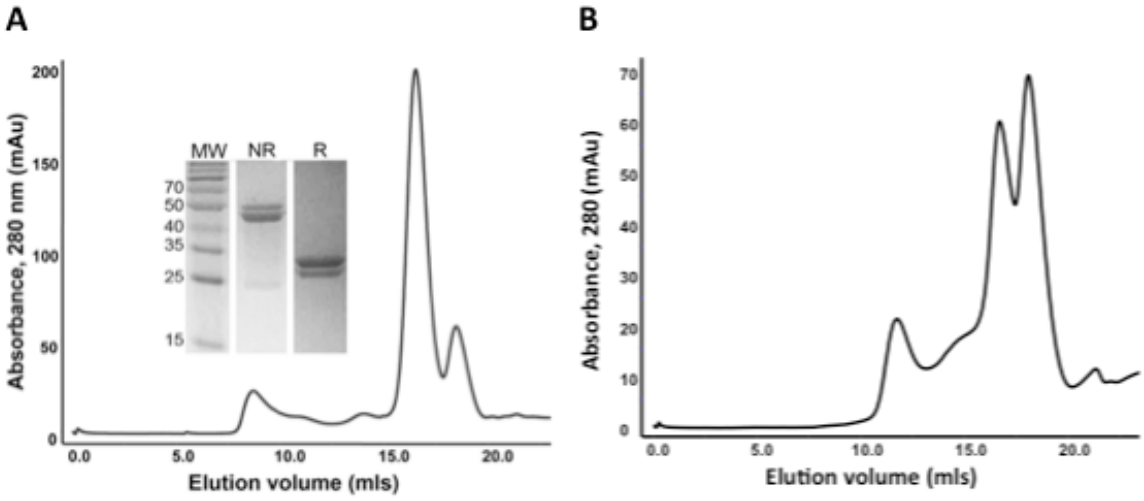


Figure 4.5: Expression and purification of anti-Flag Fab

(A) EEh14.3 was produced in *E.coli* and purified by SEC. Most of Fab eluted at the expected volume as monomer, though some aggregate species co-elute. *Inset,* size and purity of Fab was determined using SDS-PAGE under non-reducing or reducing condition. Doublet bands indicate hexa-histidine tagged or cleaved Fab. (B) The expression level of EEh13.6 was lower than EEh14.3. Two shoulders indicated that histidine-tag cleaved or uncleaved EEh13.6

To evaluate the peptide binding specificity with soluble Fab proteins, we repeated the peptide ELISAs. As was observed with the phage displayed scFv versions, strong binding was observed for the minimal DYKD peptide only when presented at the extreme n-terminus, with no binding observed for the intact FLAG peptide, the EE peptide or a control ligand lacking any peptide. The parental anti-EE framework used to generate libraries differs from EEh14.3 in only in the three heavy chain CDRs. When assessed for peptide binding in a similar set of ELISA, the anti-EE Fab exhibited strong binding to the EE peptide, but no binding to any FLAG peptide. Several commercial anti-FLAG peptide antibodies are available, with varying peptide specificities: M1 exhibits calcium-dependent binding to the FLAG peptide only when present at the extreme n-terminus, while M5 binds the peptide only at the n-terminus when proceeded by a methionine while the M2 antibody binds the intact peptide at any terminal or internal position. ELISA with the M2 antibody demonstrated clear binding to the intact FLAG peptide when presented at the n- or c-terminus but not the minimal DYKD peptide recognized by the OptCDR designs. Taken together, this demonstrates that the OptCDR antibodies recognize the DYKD peptide in a unique manner (*our antibodies are not calcium dependent, data not shown). To test activity, we repeated the ELISA similarly with Fab monomer. Significant binding above background was observed to only the minimal FLAG peptide-containing construct at the N-terminus. The acceptor framework used for CDR grafting was derived from an EE peptide-binding scFv; notably, EEh14.3 does not bind the EE-peptide, demonstrating that FLAG peptide specificity was conferred by the designed CDRs. To confirm that the framework itself does not confer low-level FLAG peptide binding, the

V_L and V_H domains for the EE peptide-binding scFv were converted to Fab and tested for FLAG peptide binding. This Fab differs from EEh14.3 in only the three heavy chain CDRs. Notably, no FLAG peptide binding was observed (Figure 5). A commercial anti FLAG peptide antibody (M2, Sigma-Aldrich, A8592) was also tested and unlike anti-FLAG peptide antibodies that we identified, it required full FLAG peptide to recognize.

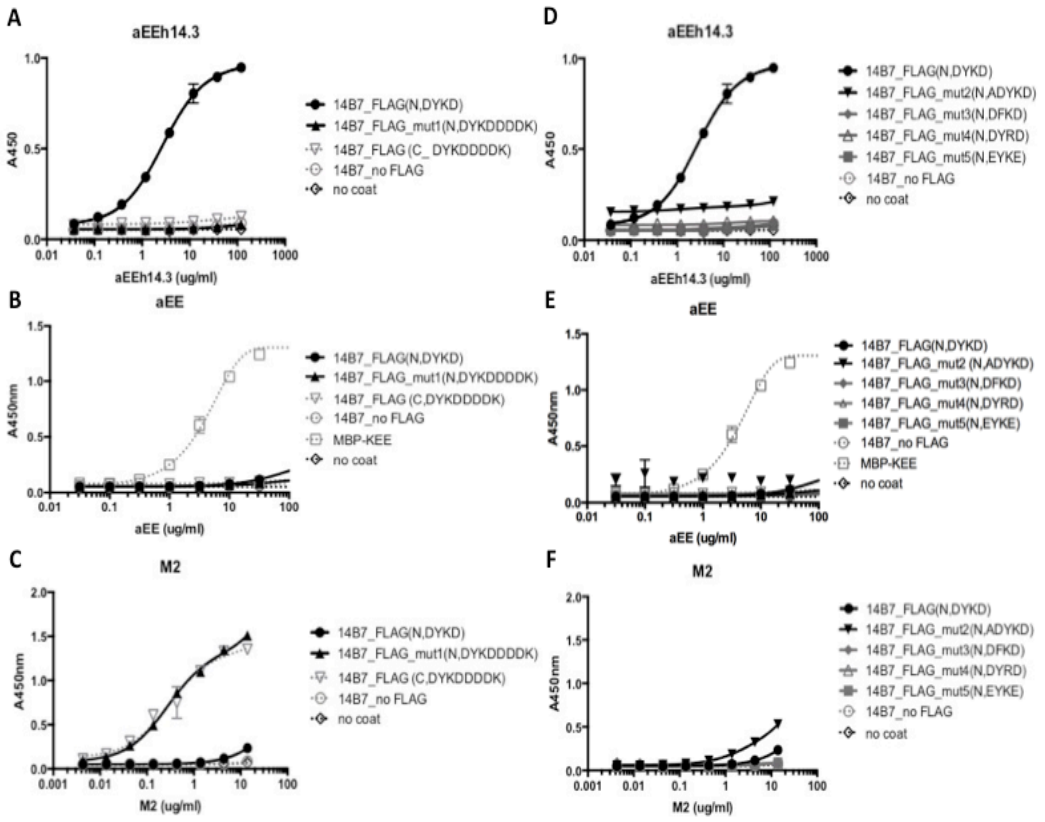


Figure 4.6: Peptide binding specificity of EEh14.3 Fab format.

OptCDR designs only bind the minimal FLAG tag at the protein n-terminus. Binding of purified Fab 14.3, (A) anti-EE Fab (B) and commercial M2 anti-FLAG lgG antibody (C) to carrier proteins presenting the minimal DYKD at the n-ternimus, the intact DYKDDDDK peptide at the n-or c-terminus or no peptide in addition to no coat control. Peptide specificity of the OptCDR clone EEh 14.3 was also tested. The same antibodies

were assessed for binding to variants of the n-DYKD peptide to evaluate the dependence of binding on each residue (D - F).

4.5 DISCUSSION

We have performed one of the first demonstrations of successful epitope-specific antibody design. Moreover, to our knowledge, this is the first example of successful *de novo* design of antibody specificities using the typical mode of antibody-antigen interactions. Using two different library approaches, we recovered four unique antibodies with different CDR sequences that each binds the same DYKD peptide conformation with exquisite specificity. While novel antibody specificities have been rationally designed before [101,115], these earlier examples introduced specificity by designing only the H3 CDR via introduction of hydrophobic or β-strand features that target linear epitopes. Although effective, this mode of binding is rare for natural antibodies. Additionally, the only antibodies designed with such methods have been single-domain antibodies (i.e. V_H without V_L). In contrast, we designed all or several CDRs to bind structural or linear epitopes by forming a pocket that complements the antigen shape with favorable interactions distributed throughout the CDRs.

The OptCDR predictions have a remarkably high success rate and notable efficacy. Although the rate of *de novo* protein design successes is increasing, it is common that tens or hundreds of designs must be screened before one shows the desired activity [116]. Here, fifty designs were screened, of which about half expressed well and four antigen binders were identified. At a structural level, it is even more remarkable that three of fifteen unique canonical structure combinations had antigen-binding variants.

Significantly, these successful designs were also highly efficacious. In the *de novo* design of novel binding interfaces, it is common for initial computational designs to have μM binding affinities [117] and require experimental affinity maturation to achieve nM binding affinities [118]. The four successful designs all exhibited good affinities, readily detectable at concentrations of 4-50 nM in ELISA. Beyond possessing good affinities, these designs are also highly specific. Our results demonstrate that the antibodies bound only the specified DYKD epitope, as even conservative mutations were entirely disruptive to binding (Figures 4.4). For antibodies to be viable experimental or therapeutic agents, they must possess good affinities and high specificities, and these antibodies meet those criteria.

Moreover, the DYKD peptide appears to present in a specific shape on this ligand due to interactions with the carrier protein. The modeled peptide structure used during antibody design has a very similar conformation as the KD in PDB 3ESV and the successful designs all interact with DYKD so that they do not clash with the carrier protein or disrupt its interactions with the tag. It is likely the antibodies are recognizing a specific conformation of the peptide, which explains why they only successfully bound n-terminal DYKD and did not bind the linear protein on a Western blot.

Although EEh13.6, EEh14.3 EEf15.4 and EEf15.4a bind DYKD specifically and with good affinities, our design process did not explicitly address protein folding or stability issues. First, in the EEh library, only four of the ten clones exhibited similar high levels of expression, two of which bound DYKD with high specificity. Second, grafting CDRs into frameworks can be structurally disruptive and/or require compensating

changes to framework residues [119]. Notably, PIGS [120] and RosettaAntibody [121] predicted main chain breaks in select OptCDR designs (data not shown), suggesting the structures may be poorly stable. This suggests that future OptCDR designs can be improved by screening CDR sequences for (i) reduced aggregation potential, (ii) compatibility with the acceptor framework and (iii) impact on the V_L-V_H interface packing angle, a factor known to be important to proper CDR positioning [122]. To address the latter two concerns, we recently developed a database of modular antibody parts [123] and an alternative design algorithm, OptMAVEn [124] for *de novo* design of complete light and heavy variable domains.

Antibodies are the leading class of therapeutics, due in large part to their ability to interact with distinct ligand conformations in order to elicit particular biological responses. However, discovery of new molecules binding specific epitopes is time consuming and challenging. We have addressed this challenge by engineering CDRs *de novo* to form a shape-complementary pocket around any specified epitope, here, a tetra-peptide in a specific kinked conformation. The resulting antibodies showed a high rate of successful affinity and specificity predictions. These results demonstrate the viability of specificity engineering by *de novo* CDR design.

Chapter 5: Developing novel anti-Ricin and anti-Abrin antibodies for toxin neutralization

5.1 Chapter Summary

Ricin and Abrin toxins derived from the castor beans and rosary pea seeds, respectively, share significant homology in both structure and function. Both consist of A and B chains which are connected by a single disulfide bond and both chains are required for toxicity because the enzymatic A chain serves as a catalytic depurinase that deactivate eukaryotic ribosome and the B subunit contains a galactose binding site for cell entry. Despite the toxicity and the history of attacks with Ricin and Abrin, only few high binding antibodies have been identified and produced for detection and neutralization. Here, we aimed to develop novel anti-Ricin and anti-Abrin monoclonal antibodies for therapeutics uses with two alternative strategies. We first immunized mice with purified A or B chains of Ricin and Abrin and screened antibodies using phage display technology and successfully construct scFv antibody libraries against each toxin chain. And in parallel, *in silico* design method was employed to to generate and identify high affinity, antigen-specific antibodies. Even though more optimization should be required to develop promising antibodies, these strategies can be used for the development of highly effective neutralization antibodies for toxin poisoning.

5.2 Introduction

Ricin toxin, derived from the castor bean *Ricinus communis*, and Abrin toxin, derived from rosary pea *Abrus precatorius*, are the member of the type 2 ribosome-inactivating proteins (RIP) family [125]. Due to their high lethality (LD_{50} of 2.7μg/kg body weight in mice for Ricin and 0.7μg/kg in mouse for Abrin)[126] and ease of purification and preparation, both toxins were classified as a Category B agent by the U.S. Centers for Disease Control and Prevention (CDC) and are considered as potent biological weapons.

Both toxins are composed of catalytically active A chain that inactivate ribosomes by depuronate 28S RNA and lectin B chain which can bind β-D-galactopyranoside to on the cell surface [127]. Despite the history of attacks with Ricin and Abrin, only a few high affinity antibodies have been produced for detection and neutralization. Some false positives (non-specific binding) or false negatives (low affinity in toxin mixture) reactions of polyclonal antibodies and their heterogeneous nature is known as one of the barriers to develop advanced diagnostics, therapeutics, and prophylactics reagents for Ricin and Abrin toxins. Moreover, even though several studies were conducted to develop both anti-Ricin and anti-Abrin neutralizing antibodies [128–133], no U.S. Food and Drug Administration (FDA) approved vaccines or treatment for both toxin poisoning have been reported [134]. To overcome the drawbacks posed by the used of polyclonal sera, the production of potent highly specific and high affinity monoclonal antibodies, which specifically block the biological active site, can add benefits to toxin neutralization.

In this study, we suggest and describe two alternative approaches, the immune library screening techniques by phage display system and the computational prediction, to generate and identify high affinity anti-Ricin and anti-Abrin monoclonal antibodies for toxin neutralization and diagnosis. Although we are still working on to isolate the most potent monoclonal antibodies exhibiting high affinity against toxin intoxication, our strategies may open the door the development of vaccines aimed at inducing a similar immune response.

5.3 MATERIALS AND METHODS

5.3.1 Murine Immunization

For the murine immunization, all protocols were approved by the University of Texas at Austin IACUC (protocol number 2015-00077), and all mice were handled in

accordance with IACUC guidelines. Three 6-week-old BALB/c mice were primed by subcutaneously with 1 to 1 mix of 5μg of recombinant Ricin A or B subunit (NR-853 or NR-853, BEI), or 1μg of Abrin A or B (NR-43945 or NR-43946, BEI) in PBS and TiterMax gold adjuvant (T2684-1ML, Sigma). Four weeks later, the mice serum were collected through tail vein and mice were boosted with same amount toxin subunits in a same way. Two weeks later, the mice were bled as previously and boosted again except one-time boost group. Rests of the mice were boosted for two more times in every two weeks. When the boosting was done, mice were sacrificed and spleens were removed. Spleens were immediately immersed in 1ml of RNAlater stabilization solution (AM7024, Thermo Fisher Scientific), and allow to soak overnight at 4 °C, and then stored at -80 °C. Serum ELISA was done to determine anti-toxin subunit antibody titers (described below). Serum titer defined as background titer when the A450 value is 0.2.

5.3.2 Phage Displayed scFv Library Construction

Total RNA was extracted from spleens with Purelink RNA mini kit (Invitrogen) according to the manufacturers' instructions. After measuring the quality and concentration of total RNA by Nanodrop 2000 (Thermo Scientific), first-strand cDNAs were synthesized using 5 μg of total RNA, random hexamer primers (Thermo Scientific), and Superscript III (Invitrogen). The synthesized cDNA were used as a template for amplifying the V_L and V_H repertoires with previously described primer sets [110]. 10 ng of each of purified V_L and V_H were combined to generate scFv (V_L –linker- V_H) by overlapping PCR. The gel purified scFv fragment were digested with SfiI, ligated into

pMoPac24 vector, and transformed into *E.Coli* XL1Blue by electroporation. The transformed cells were immediately recovered with 4ml of 2xYT media and transferred to 30ml 2xYT medium with 5 μg/ml tetracycline, 200 μg/ml ampicillin, and 0.1% glucose to pool the master library as well as rescue scFv-displaying phages. The 30ml cultures were grown at 37°C by when the OD_{600} reached ~0.6, 3mls of cultures were pooled as the master library, and the rest of cultures induced for the next steps.

5.3.3 Phage Production, Panning and Screening

To induce and rescue phage-displayed scFv, 1mM IPTG was added 30 min prior to add M13KO7 helper phage into the rest of the cultures to display scFvs on phage surface. Two hours later, 50 μg/ml of kanamycin was supplemented into the culture to selectively grow phage-infected cells. After overnight culture, phages were purified by precipitation with 1/5V of precipitation buffer (2.5M NaCl, 20% PEG-8000) and after 1-hour incubation, precipitated phage were harvested and resuspended using proper volume of PBS (pH 7.4). Phage concentration was determined as colony-forming units (cfu/ml) as described in Chapter 2.

Five rounds of panning were performed using a-cMyc and each toxin subunit or toxoid as a bait. Eight wells of ELISA plate (Costar) were coated with 1μg/ml of each bait and incubated overnight at 4°C. Input phage (80μl), which were 1 to 1 diluted into 5% nonfat milk in PBST (PBS, 0.05% Tween-20), added into the eight wells after 1 hour blocking step, and incubated for 1 hour at room temperature. Ten times washing by vigorous pipetting with PBST, bound phage were eluted with 50μl per well of 0.1N HCl

(pH 2.2) by10min incubation at room temperature and immediately neutralized with 3µl of 2M Tris base per 50µl eluted phage. To move on to the next round of panning, half of the output phage was infected into 2ml of log-phase XL1-Blue *E.coli* cells for 30min and scale up to the culture volume to 30ml of 2xYT including 5µg/ml tetracycline, 200µg/ml ampicillin, and 1% glucose. The rest of steps were described above. For each round, input and output phage titers were determined as cfu/ml as previously described and sequence diversity was monitored by identifying from the University of Texas at Austin Core Facility.

To perform high throughput screens, single colonies from output plates were inoculated into 100µl of 2xYT media with proper antibiotics (5µg/ml tetracycline, 200µg/ml ampicillin, and 1% glucose) and grown for overnight at 37°C in 96-well low binding plates. The next day, 50µl of 2xYT media with proper antibiotics including final 1mM IPTG added and shaken at room temperature for 3 hours to induce antibodies, followed by adding 50µl of 2xYT media with proper antibiotics with M13KO7 helper phage for 2 hours. After 2 hours, 50µg/ml kanamycin was added into each well and the plate was shaken for overnight at room temperature. Using ELISA, the expression and the binding affinity to Ricin or Abrin of each clone was screened.

5.4 RESULTS

5.4.1 Overview of immune library construction

The overall strategy of murine immunization and immune library construction is described in Figure 5.1.

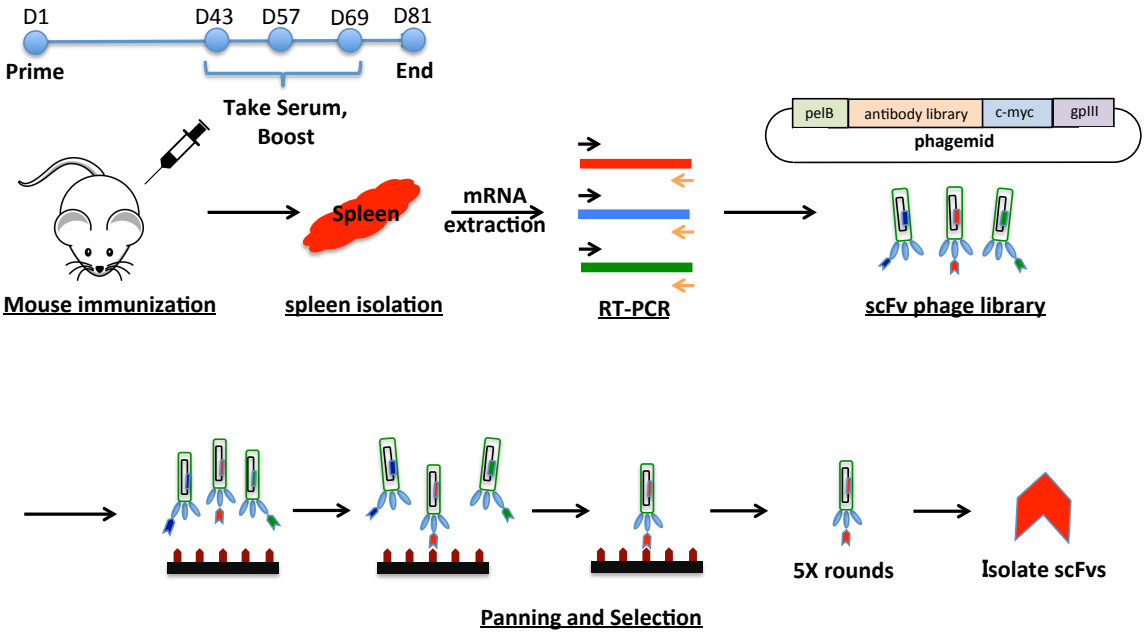


Figure 5.1: Workflow of generation of scFv libararies for isolating anti-Ricin or anti-Abrin antibodies.

Mouse was immunized by subcutaneously with A subunit or B subunit of Ricin (5µg/mouse) or Abrin (1µg/mouse) toxin. After prime and each booster, mouse blood was collected by cutting tail vein to determine serum titers, and after three-booster immunization, spleen was isolated. Total RNA was extracted from the spleen and V_H and V_L mRNA were reverse transcribed and amplified to cDNA to construct scFv libraries. In order to screen and isolate target bound scFv with high affinity, phage displayed platform was used.

We immunized three mice with purified Ricin and Abrin A or B subunit by subcutaneous route to elicit immune response and generate highly specific antibody repertoires. For all mice, three times booster immunization were performed to reach at reasonable serum titers (>1:1,000) (Figure 5.2) and after the final booster immunization, all mice were sacrificed and a spleen was isolated for generation of immune libraries. We constructed scFv libraries (V_H + V_L) based on RNAs extracted from the spleen (Figure 5.3). The size

of libraries was up to 10^5 and with diversity, even though some clones have stop codon in the middle of the sequence or unexpected missing one of the domain sequence (Table 5.1). The constructed scFv libraries were applied to each antigen (Ricin and Abrin A or B subunit) using phage surface display system, and screening was performed by to isolate high affinity binders.

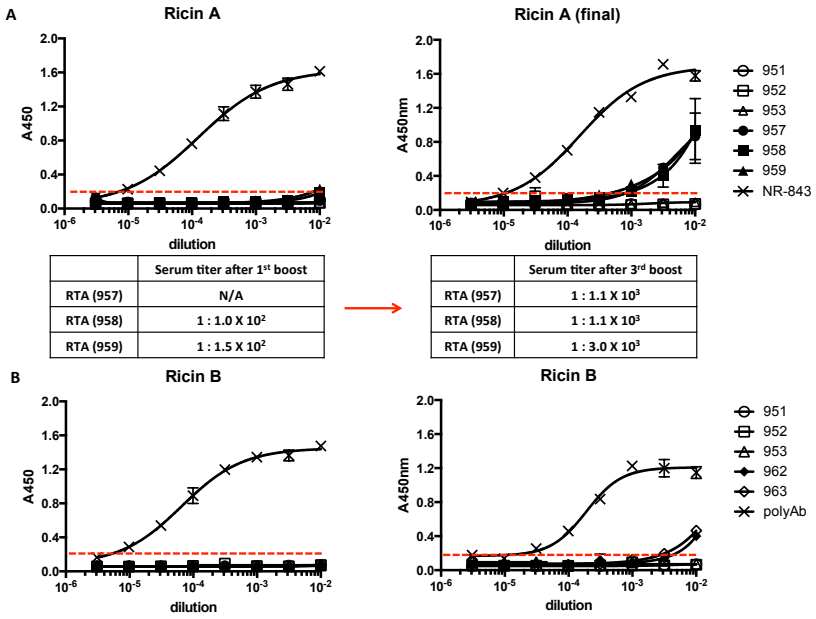


Figure 5.2 (continued)

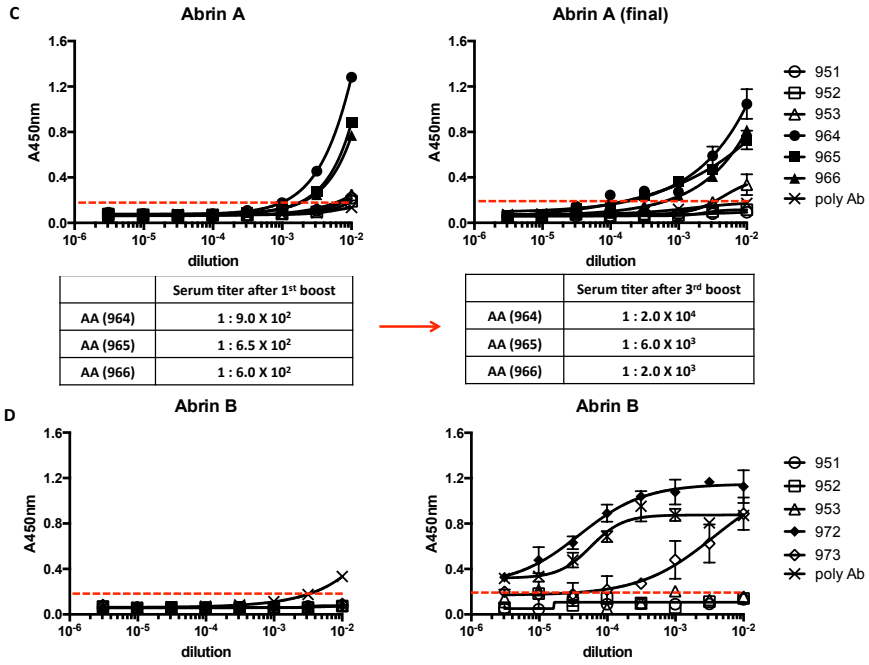


Figure 5.2: Serum titer measurement after murine immunization

After three times booster with each subunit including PBS control (#951-953), serum titers were measured by ELISA. All left panels represented the results of serum ELISA after the first time boost and all right panels indicated the results after the final booster. For both A subunits about 10 times higher titers were detected after the final boost (A and C). For Ricin B subunit, no significant serum titer was detected (B) but the result for Abrin B subunit looked promising because #972 was shown to generate more than 1:300,000 against Abrin B chain (D).

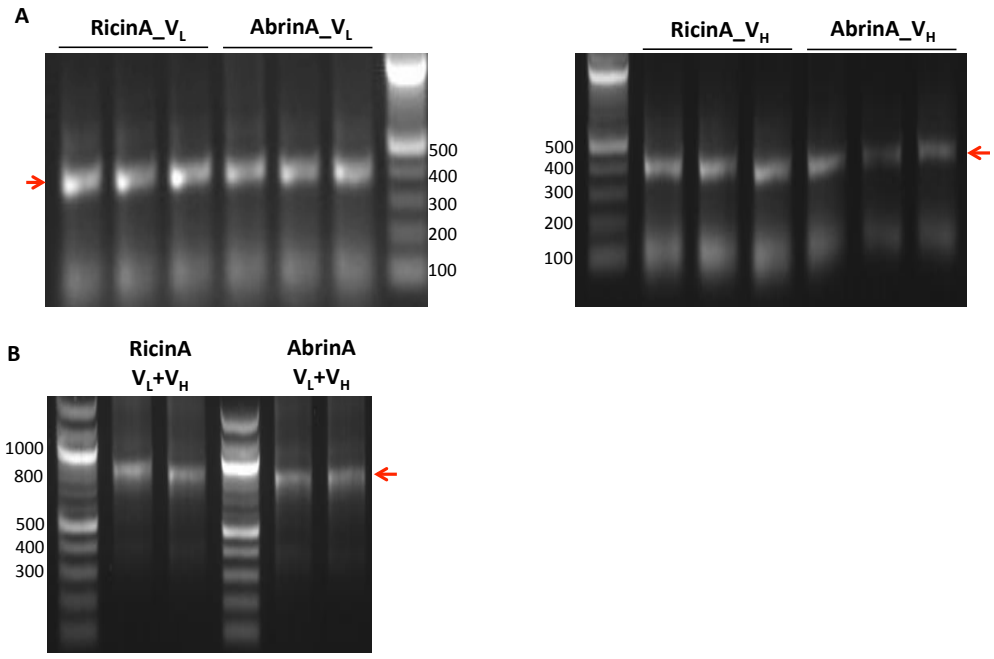


Figure 5.3: Library construction for spleen repertoires

(A) All mRNAs were extracted from mice spleen, and cDNA was generated using mRNA as template with reverse transcription. V_L and V_H were amplified using mouse V gene specific primer sets as described previously [12]. (B) For the scFv library construction, overlapping PCR was performed and after SfiI digestion, scFv libraries were subcloned into the phagemid vector, and then transformed into *E.coli* cells.

Table 5.1: Biophysical characteristics of EE peptide-binding antibody fragments.

Library	Size	Diversity	VL+VH
RicinA(1B)	1.17 x 10^5	80% (8/10)	90%
RicinA(3B)	9.7 x 10^4	100% (10/10)	60%, Stop codon:20%
RicinB	5.9 x 10^4	100% (10/10)	20% **
AbrinA(1B)	1.9 x 10^5	90% (9/10)	60% , stop codon 10%
AbrinA(3B)	2.31 x 10^5	100% (10/10)	80%
AbrinB	2.37 x 10^5	100% (10/10)	30%**

**VH only

5.4.2 Library screening via phage surface display system

As described in the previous chapters (Chapter 3 and 4), panning antibody fragment via phage display system is highly efficient to isolate high binding clones. For each library, five rounds of phage panning against each subunit were performed for selection. Before and after each round of panning, phage titers were measured to detect enrichment of phage variants after selection steps and for both toxins, the output titers increased after round three, meaning significant enrichment of phage variants was occurred (Figure 5.4).

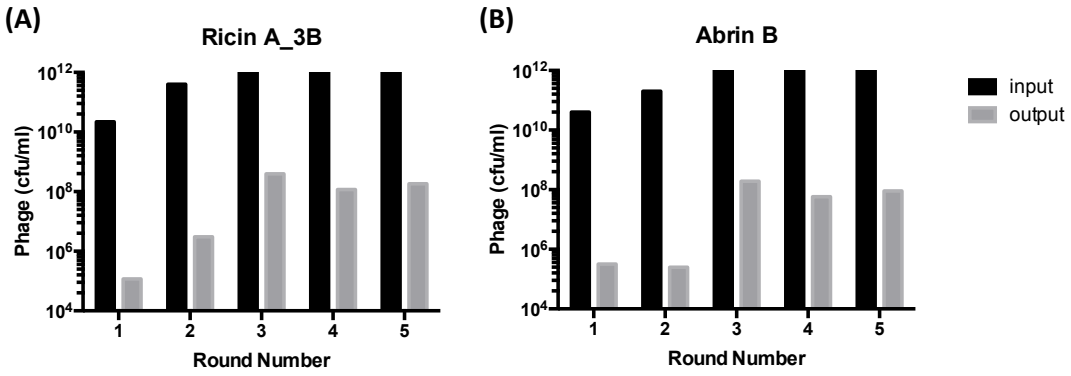


Figure 5.4: Phage panning selection of scFv libraries for anti-Ricin and anti-Abrin antibodies.

Five rounds of phage panning against each subunit were performed to select high binding scFv from the libraries. Phage titers were measured before and after each round of panning. For both antigens, the output titers increased after round three, meaning an enrichment of phage variants after selection steps.

After the selection process, individual clones (up to 50 individual clones) were screened by phage ELISA for their binding ability against toxin subunits. Unexpectedly, less than 10% of clones showed detectable signals to antigens on ELISA and had high

background signal, meaning non-specific binding may be caused by partially misfolded antibodies. For the further assay, we selected and sequenced 6 clones that represented twice higher signal than the background, and they turned out one unique clone after affinity maturation. Even thought the background was higher than we expected, we needed to confirm the affinity to Ricin A chain using monophage ELISA (Figure 5.5).

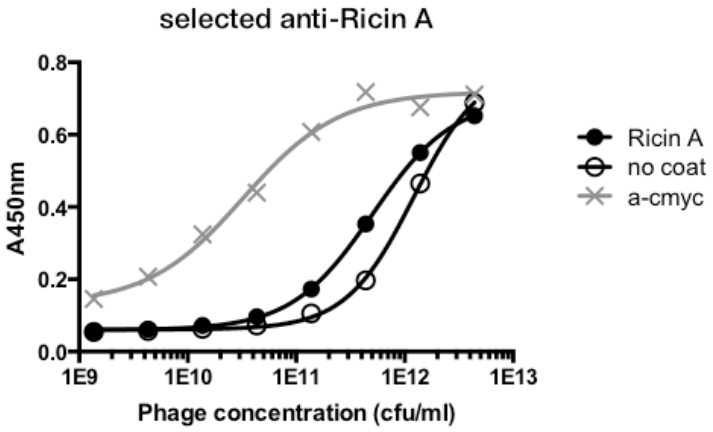


Figure 5.5: Validation of binding affinity using monophage ELISA

Mono-phages that displayed selected anti-Ricin antibody on surface were applied to the wells where were coated with Ricin A chain (for antigen specificity and affinity), a-cmyc (for assuming the level of antibody display), and no coat control.

5.4.3 Some designed anti-Ricin and anti-Abrin antibodies retain antigen specificity

Along with the immune libraries, we designed anti-Ricin and anti-Abrin antibodies using computational method as an alternative approach. Previously, crystal structure of Fab fragment from a protective anti-Ricin monoclonal antibody has been reported (4H20.pdb) [135]. In addition, highly potent single domain camelid V_HH that can neutralize ricin *in vitro* and *in vivo* has been engineered that is proved to be passively protective[136]. Based on the sequence of the anti-Ricin antibodies (4H20 for the A-chain

specific antibodies for both toxins and RTB-B7 for the B-chain specific antibodies) and using computationally designing strategies, we initially designed antibodies to get improved affinity. The results provided us total 28 different predicted antibody sequences (5 for anti-Ricin A, 8 for anti-Ricin B, 9 for anti-Abrin A and 6 for anti-Abrin B chain antibodies) and we tested top scored antibodies for the further assays. To determine whether designed antibodies still retain antigen specificity or not, we performed mono-phage ELISA against each toxin subunit (Figure 5.6). For all tested 6 antibodies, they could highly expressed and displayed on phage surface as indicated as a-cMyc signal level. Among the 6 antibodies, only 2 of them were shown to retain antigen specificity after the engineering. Designed anti-Ricin A antibody (1_27_66_34_14A) showed moderate binding affinity to purified recombinant ricin A subunit as well as ricin toxoid in a similar manner, meaning that the epitope where the antibody bind may completely overlap in the A chain. And one of anti-Ricin B antibody (1alpha_0060_A6_0036) was able to highly bind to purified ricin B chain rather than the toxoid, assuming that more than one epitope may be available for this antibody (Figure 5.6A). Interestingly, all anti-Abrin antibodies could not possess their affinity and specificity to Abrin even though they were engineered the based on an original sequence that is known to be RTA or RTB-specific neutralization antibody (Figure 5.6B). Even if the structure and function of Ricin and Abrin are similar each other, due to all computationally design were done using only crystal structures of Ricin subunits and their specific antibodies, more insights about the structures of Abrin subunits should be required.

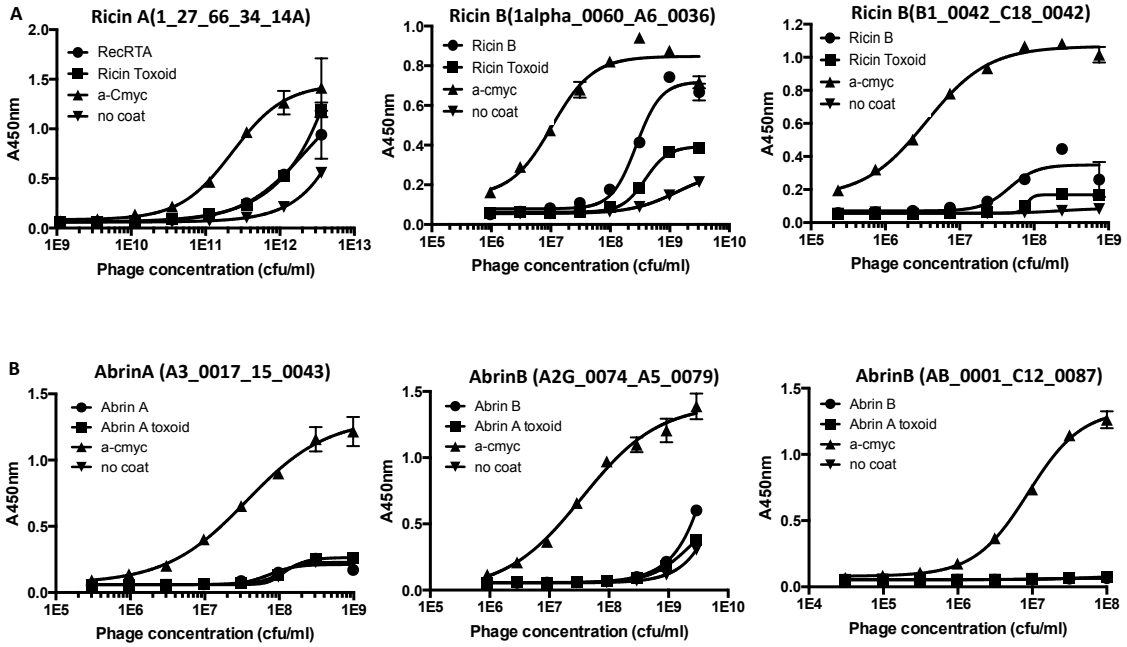


Figure 5.6: Antibody binding affinity to toxin subunit by phage ELISA

Binding affinity and specificity of computational designed anti-Ricin and anti-Abrin subunit antibodies were tested by phage ELISA as described previously. (A) Designed anti-Ricin A and B chain antibodies or (B) anti-Abrin A and B chain antibodies were applied on ELISA plate wells that coated with purified toxin subunit (circle), toxoid (square), a-cmyc (triangle), and no coat control (inverted triangle). All antibodies could display well on phage surface (level of signals against a-cmyc), but among six designed antibodies, only two of them (1_27_66_34_14A and 1alpha_0060_A6_0036) showed affinity to antigen over control.
**All computational works were done by Monica Berronado and Pragati Prasad in Macromoltek.

5.5 DISCUSSION

To date, using highly neutralizing monoclonal antibodies is the most efficient and promising ways to intoxicate post-exposed toxin. In order to develop anti-Ricin and anti-Abrin monoclonal antibodies for toxin neutralization, the toxoid[137], isolated or purified

toxin subunits [128,132,134], or the holotoxin [138] have been used as immunization reagents. Recent study demonstrated that antibody affinity and the ability to neutralize ricin *in vitro*, is highly correlated as well as with post –exposure protection [138]. Even the isolated antibodies are able to bind to the same epitope, the neutralization potency is different depending on the affinity. Therefore, identifying monoclonal antibodies with high affinity or improving binding affinity for already defined toxin antibodies are the most priority to isolate protective antibodies. Lately, multiple antibodies with high affinity (sub-nanomolar) to Ricin A chain have been isolated by combination with Next-Generation Sequencing (NGS) based antibody repertoire analysis and yeast surface display based antibody screening technology [134]. Previously several studies reported that unfavorable conformations resulting in poor biological properties often occurred due to random or non-native pairings of V_H and V_L, and non-native expression host [139–142]. Therefore, the NGS in combinatorial library screening can be one of the solutions to reduce huge amount of efforts and times to discover high affinity antibodies with better biophysical properties (expression level and stability) as validated in the study.

Here, we performed two efficient methods to identify high specific binders to A or B subunit to block biological activities and toxin intoxication. We successfully construct scFv format of immune libraries against each toxin subunit and Ricin and Abrin with diversity and based on the known anti-Ricin antibody crystal structure, *in silico* designed antibodies were produced and evaluated. Enough number of antibodies from each immune libraries were screened for binding affinity to the antigen, almost none of antibodies were identified as potent binders, assuming that the starting pool of V genes

for scFv constructs was not enough to recover promising antibodies or unexpected pairing of V_H and V_L resulting in misfolded conformation of scFv to contribute poor bindings. To improve the binding affinity of previously reported anti-Ricin A and anti-Ricin B[135,136], and mature the binding specificity and affinity to Ricin-related toxin, Abrin, antibody structures and sequences were predicted by computational modeling. Among three anti-Ricin antibodies and three anti-Abrin antibodies, only two of anti-Ricin antibodies have been proved to retain binding affinity with its own antigen, and one of anti-Ricin B antibody and all designed anti-Abrin antibodies could not possess binding specificity and affinity after the modeling, meaning optimization and fine-tune of modeled antibodies should be required. More experimental tests are currently ongoing with the aim of isolating and investigating the potential monoclonal antibodies for immunotherapeutic treatment of both toxin poisoning. Even though we are not fully successful yet, our strategies may shed light on the development of therapeutic or prophylactic vaccines aimed at inducing a similar immune response.

Chapter 6: Conclusions and Future Directions

6.1 Conclusion

In this dissertation, various platforms of epitope specific antibodies developed by engineering different size of antibody format from Fab to sdAb, and generating computationally-designed, or targeted synthetic or immune libraries. We have successfully improved antibodies (Fab/EE) binding affinity and stability by expanding hydrophilic areas of crystal contacts for using crystallization chaperone (Chapter 2), altered binding specificity by randomly mutated CDR regions (Chapter 3, anti-RTX) or computationally designed CDR regions (Chapter 4, anti-FLAG tag), and developing novel and high affinity neutralization antibodies using mouse immunization (Chapter 5, anti-Ricin and Abrin). In each study, the main technology to select and screen the highly specific, stable and binding to antigen was phage-displayed system because this display technology is fast and easy to handle in a lab, therefore, we are expecting the technology can be employed in future antibody engineering in many purposes. After we isolated promising antibodies from the library, most of antibodies should be required to optimize and improve its own biophysical characteristics such as stability or solubility, or binding affinity for better behavior. But still, the information from the engineered antibodies has been inspired us to guide the design of novel antibodies. Finally, our engineered Fab/EE as well as our computational designed anti-FLAG antibodies can be applied in co-crystallization serving as a general crystallization chaperone due to their peptide-specificity. This peptide specific platform will be more promising than the currently available protein-specific chaperone technology because by switching a binding epitope to short peptide specificity, limitless number of proteins can have potential to be crystallized. And generating synthetic or immune libraries, which are good sources of

identifying and isolating novel antigen-specific antibodies can be used for targets those are challenging to crystallize or immunize due to their toxicity.

6.2 RECOMMENDATIONS FOR FUTURE WORKS

While we have successfully identified sdAbs possessing RTX binding affinity, further biophysical improvements should be necessary using additional engineering methods. Two of variants (anti-RTX_V1 and V3), showed high binding affinity but exhibited low expression level and stability compared to WT (Chapter 3) which are key factors for hypercrystallizable trait. This can be accomplished by selecting thermo-resistant sdAbs using additional set of engineering with phage display system [143].

In Chapter 4, we proved that our most engineered anti-FLAG Fabs (EEf15.4 or EEf15.4a) possessing high binding affinity to minimal FLAG peptide, high stability and expression level. But the structural information of both antibodies has not been addressed due to their purification quality. Due to the contamination with free light chain after a size-exclusion chromatography causing crystallization disturbance, separation of the Fab and the free light chain should be required. This can be performed by applying alternative purification method such as ion exchange chromatography because the theoretical pI of Fab and free light chain is different as 8.7 and 6.9, respectively. Once we can get clean Fab, we can set up crystallization trays and have more chance to have crystals and ultimately, we have more knowledge about the structural information and can compare and contrast with predicted Fab structure.

Immune libraries of anti-Ricin and anti-Abrin antibodies were successfully generated using phage display formats with decent diversity (Chapter 5). After three times booster, we expected to have significant titers (>1:10,000) against targets. But we had only up to 1:1,000 for all subunits except Abrin B subunit and this resulted in less chance to isolate high binding antibodies. Therefore, in order to identify novel antibodies with high affinity alternative methods can be applied. First, we can try to immunize toxoid rather than only purified subunits to elicit higher immune response for both subunits (A or B subunit of toxins). Another available approach is developing computational designed Ricin or Abrin specific antibodies. This approach cannot guarantee whether we can identify the neutralization antibodies or not, but we can save enormous time and efforts to find antibodies possessing binding affinity.

Appendix A: Laboratory Protocols

A.1 WHOLE PLASMID MUTAGENESIS- PCR POINT MUTAGENESIS

1. Methylation of template DNA

Template DNA (100ug)	1µl
3.2mM SAM (S-adenosylmethionine)	1.5ul (final concentration: 160µM)
NEB buffer #2	3ul
CpG Methyltransferase (M.sssI)	1ul
Total	30ul

Incubate at 37°C for 1hour and 15min.

2. PCR reaction

Methylated DNA template	3ul
Forward primers	1ul
Reverse primers	1ul
10X Thermopol buffer	5ul
10mM dNTPs	1ul
Vent polymerase	1ul
ddH_2O	38ul
Total	50ul

PCR conditions

1	95°C	4min
2	95°C	30sec
3	Annealing temperature (Tm of primer-5°C)	30sec
4	72°C	4min 30sec → go to step2 and additional 39cycles
5	72°C	10min

3. Digestion of Methylated DNA by adding 0.5ul of DpnI and incubated at RT for o/n

4. Transform into a competent cells by heat shock or electroporation

A.2 2ND LIBRARY CONSTRUCTION

Construction of megaprimers

Template: pMoPac24_aRTX_A5,G4 or B6
Primer set #1 : 5' pAkpel (46.9°C) and 3'scforlong (64.7°C)
Primer set #2 : 5' scback (73.9°C) abd 3'scforlong (64.7°C)

1st PCR (error-prone PCR using mutazymeII)

Template	100ng – 200ng
10X buffer	5µl
40mM dNTPs	1µl
Primer (10µM)	1µl per primer
mutazymeII (stratagene)	1µl
ddH_2O	41.5µl
Total	50µl

Use primer set #1 and #2
Annealing temp (41.9C for set#1 and 59.7C for set#2)

PCR conditions

1	95°C	2min
2	95°C	30sec
3	Annealing temperature (Tm of primer-5°C)	30sec
4	72°C	1min → go to step2 and additional 29cycles
5	72°C	10min

Gel extraction

2nd round PCR

Template (gel extracted pcr product)	1µl
10X Thermopol buffer	5µl
10mM dNTPs	1µl
Primer (3'scforlong, 10µM)	1µl
Vent polymerase	1µl
dH_2O	41µl
Total	50µl

Annealing temperature: 59.7°C

PCR conditions

1	95°C	4min
2	95°C	30sec
3	Annealing temperature (Tm of primer-5°C)	30sec
4	72°C	30sec → go to step2 and additional 29cycles
5	72°C	2min

A.3 ETHANOL PRECIPITATION OF DNA

Reagents Needed:
- 3M sodium acetate (pH5.2)
- 100% ethanol
- 70% ethanol
- DNA

Protocol
1. Add 1/10 volume of sodium acetate (final concentration of 0.3M) and 3V of cold 100% ethanol
2. Mix well and incubate on ice (or -20degrees) for 30min
3. Spin it for 10min at the max speed
4. Carefully decant supernatant
5. Add 1ml of 70% ethanol
6. Repeat 3 and 4.
7. Air dry pellet.
8. Resuspend pellet in the appropriate volume of water.

A.4 TM EXPERIMENT

Reagent needed:
- SYPRO Orange gel stain (Sigma-Aldrich)
- Fast optical 96-well PCR plate (L0032, Biosci store) and Optical Adhesive cover (L0014, Biosci store)

Monitoring Protein unfolding using fluorescent based thermal scanning
1. Prepare each running sample and blank control in triplicate
   - 20 – 250μM protein in HBS
   - 15X SYPRO (final, 1μl of 300X SYRPO Orange stock)
   - 20μl total
2. Set up the machine (ViiA™7)
   - Fast 96-well
   - Melt curve
   - SYBR green reagents
   - Standard (Fast)
   - Define reporter as SYBR (ROX), quencher as none, and passive reference as one
3. Run method
   - Select and use only dissociation stage (starting temperature: 25°C and ending temperature: 90°C) Increment: 0.016 °C/sec
   - Select optical filter: Excitation (465nm)/ emission (590nm)
4. Run
5. Analysis
 - Go to home of ViiA7 software and click Analysis: instrument console → Select the machine that you used (A, B or C) → Manage instrument → Manage file → Download and export → Save as the data as excel format (.xlsx)
- Tab over to Dissociation curve and display the graph as a raw plot of the data. This gives your fluorescence vs. temperature graph

Or follow company protocols (Protein Thermal Shift™ Dye Kit, Applied biosystems by life technologies

A.5 COMPETENT CELL PREP

(Reference: Annalee Nguyen's note)

A.5.1 Calcium Chloride competent cells

Reagents Needed: All reagents and materials should be autoclaved and pre-chilled
- 100ml of LB
- 50ml of 0.1M $CaCl_2$
- 10ml of 0.1M $CaCl_2$/15% glycerol
- Sterilized e-tubes and bottles

Day1
1. Inoculated a single colony into 3ml of LB media (add appropriate antibiotics if it is needed) and grow cells at 37°C for o/n

Day2
1. Transfer 1ml of o/n culture into 100ml of LB (in 500ml flask) and grow cells at 37°C until OD600 = 0.4~0.6
2. Incubate the cells on ice for 10min (Keep cold in every step)
3. Harvest cells by centrifugation in a sterilized bottle for 5min at 5krpm (4°C)
4. Decant supernatant and gently resuspend on ice using 10ml of 0.1M $CaCl_2$ (ice cold) (Treat cells gently)
5. Incubate cells on ice for 20min
6. Cell down by centrifugation for 5min at 5krpm
7. Discard supernatant and gently resuspend on ice using 4ml of 0.1M $CaCl_2$/15% glycerol
8. Aliquot in sterilized e-tubes (100µl/tube) and store at -80°C

A.5.2 Electrocompetent Cells

Reagents Needed: All reagents and materials should be autoclaved and pre-chilled
- 500ml of media (5g Tryptone, 2.5g Yeast extract, 2g NaCl) in a 2L flask
- 2 X 500ml bottles (six times washed with MilliQ dH_2O, detergent and bleach free)
- 400ml of 10% glycerol
- 400ml of dH_2O
- Sterilized e-tubes

Day1
1. Inoculated a single colony into 5ml of media (add appropriate antibiotics if it is needed) and grow cells at 37°C for o/n

Day2
1. Add glucose to the 500ml of media and transfer 5ml of o/n culture into it. Grow cells at 37°C until OD600 = 0.6 (3h for a normal strain)
2. Incubate the cells on ice for 15min (Keep cold in every step)
3. Harvest cells by centrifugation in a sterilized bottle for 20min at 3krpm (4°C)
4. Discard supernatant, add 400ml of cold water, and resuspend the pellet by swirling (treat gently)
5. Pellet cells by centrifugation for 20min at 3krpm (4°C)
6. Decant supernatant, add 5-10ml of cold 10% glycerol, and resuspend cells. Add cold 10% glycerol to halfway and centrifuge again for 10min at 3krpm (4°C)
7. Pour supernatant immediately, and resuspend pellets with remaining supernatant (~1ml liquid)
8. Aliquot in sterilized e-tubes (100µl/tube) and store at -80°C

A.6 PROTEIN EXPRESSION AND PURIFICATION

(Adapted from Maynard Lab protocol and compiled Kevin's dissertation)

Reagents Needed:

- 2 X 250ml of TB media in 1L flask (autoclaved)
- 1M IPTG, filter sterile (store at -20°C)
- Sucrose buffer (0.75M sucrose, 0.1M Tris pH8.0) filter sterile, pre-chilled
- 1mM EDTA, pH8.0, autoclaved, pre-chilled
- 10mg/ml lysozyme, dissolved in sucrose buffer or EDTA solution
- 0.5M $MgCl_2$, pre-chilled
- Dialysis buffer: 10mM Tris (pH8.0), 0.5M NaCl (2 volume of culture volume), pre-chilled
- 1M imidazole in water (autoclaved)
- Wash buffer: 20mM Tris (pH8.0), 0.5mM NaCl, 20mM imidazole (autoclaved)
- Elution buffer: 0.1M EDTA, 20mM Tris (pH8.0), 0.5M NaCl (autoclaved)

Day1

1. Inoculated a single colony into 2ml of media (add appropriate antibiotics if it is needed) and grow cells at 37°C during the day. Scale-up to 250ml and start overnight culture at 37°C (225rpm).

Day2

Expression

1. Harvest cells by centrifugation for 10min at 5krpm.
2. Decant supernatant, resuspend pellets using ~20ml of fresh media, and return to a flask.
3. Shake at 25°C for 1hr (225rpm)
4. Induce antibody by adding 250µl of IPTG (final concentration: 1mM) and let shake at 25°C for additional 5hrs
5. Harvest cells by centrifugation for 10min at 5krpm

Fractionation (Keep cold in every step)

6. Resuspend pellets in 10ml of sucrose buffer (per 250ml culture volume), and transfer equally to 3-40ml oak ridge tubes
7. Add additional 10ml of sucrose buffer (per each tube)
8. Add 250 ~ 300µl of lysozyme (10mg/ml) per tubes
9. Add 10ml of 1mM EDTA and rotate at 4°C for 40-60min
10. Add 0.5ml of $MgCl_2$ and rotate at 4°C for 40-60min
11. Pellet cells by centrifugation for 20min at max rpm (20krpm)
12. Transfer supernatant to dialysis tubing (check cut-off of dialysis tube) and start to dialyze at 4°C for o/n

Day3

1. Transfer dialysate into 50ml conical tubes and add 1M imidazole (final concentration: 10mM)
2. Add ~0.5ml of charged IMAC resin for every 50ml of solution and rotate at 4°C for 3-18hrs
3. Collect resin by centrifugation (use table-top centrifuge) at 1krpm for 2min
4. Decant all supernatant but 2ml, wash with 10ml of wash buffer and centrifuge again
5. Decant all supernatant but 2ml and collect resin by pouring trough column
6. Wash with 10ml of wash buffer for several times
7. Elute with 1ml of elution buffer (*you can see a color of resin turns white)
8. Size exclusion purification (running buffer: 1X HBS for crystallization proteins or 1X PBS as general)

***Fab expression (skp co-expression)**

Reagents Needed:
- 250ml of TB media in 1L flask (autoclaved)
- 10% glucose (filter-sterile)
- 1M IPTG, filter sterile (store at -20°C)
- 20% L-arabinose (filter-sterile)
- Sucrose buffer (0.75M sucrose, 0.1M Tris, pH8.0) filter sterile, pre-chilled
- 1mM EDTA, pH8.0, autoclaved, pre-chilled
- 10mg/ml lysozyme, dissolved in sucrose buffer or EDTA solution
- 0.5M $MgCl_2$, pre-chilled
- Dialysis buffer: 10mM Tris (pH8.0), 0.5M NaCl (2 volume of culture volume), pre-chilled
- 1M imidazole in water (autoclaved)
- Fab Wash buffer: 25mM Tris, 10% glycerol, 0.1M NaCl, 20mM imidazole, pH9.0 (autoclaved)
- Fab Elution buffer: 25mM Tris, 10% glycerol, 0.1M NaCl, 0.5M imidazole, pH9.0 (autoclaved)

Day1
1. Inoculate 12ml 2xYT/2% glucose (appropriate antibiotics) from a single colony and grow cells at 37°C for o/n

Day2
1. Transfer overnight cultures (adjust OD_{600} = 0.1) into 250ml of TB/0.1% glucose/antibiotics and grow cells at 25°C (225rpm) until OD_{600} = 0.5 (3 ~3.5hrs)
2. Induce skp by adding L-arabinose (final concentration: 0.2%) and grow cells at 25°C for 30min
3. Induce Fab by adding IPTG (final concentration: 1mM) and express cells at 25°C for additional 4hrs
4. Harvest cells by centrifugation for 10min at 5krpm

Fractionation (Keep cold in every step)

5. Resuspend pellets in 10ml of sucrose buffer (per 250ml culture volume), and transfer equally to 3-40ml oak ridge tubes
6. Add additional 10ml of sucrose buffer (per each tube), 250 ~ 300µl of lysozyme (10mg/ml), and 10ml of 1mM EDTA per tubes. Incubate at 4°C for 45min
7. Add 0.5ml of $MgCl_2$ and rotate at 4°C for 10min
8. Pellet cells by centrifugation for 20min at max rpm (20krpm)
9. Transfer supernatant to dialysis tubing (check cut-off of dialysis tube) and start to dialyze at 4°C for o/n

Day3
1. Transfer dialysate into 50ml conical tubes and add 1M imidazole (final concentration: 10mM)
2. Add ~0.5ml of charged IMAC resin for every 50ml of solution and rotate at 4°C for 3-18hrs
   Before adding to samples, IMAC resin allow to pass through with 10ml of dH_2O followed by Fab wash buffer and resuspend in 1ml Fab wash buffer to transfer to samples.
3. Collect resin by centrifugation (use table-top centrifuge) at 800rpm for 2min
4. Pour supernatant but 2ml, wash with 10ml of wash buffer and centrifuge again
5. Decant all supernatant but 2ml and collect resin by pouring trough column
6. Wash with 10ml of wash buffer for several times
7. Elute with 1ml of Fab elution buffer (*you can see slight change of color of resin)
8. Size exclusion purification (S200) (running buffer: 1X HBS pH7.4)
   Should elute in B2 – C1 fractions.

A.7 ELISA (ENZYME-LINKED IMMUNOSORBENT ASSAY)

Reagents needed
- 96 well ELISA plates and sealers
- 96 well cell culture plate
- PBS (pH7.4), PBST (0.05% Tween20), PBST + 5% milk

Day1
1. Coat plate with antigen
   - Typically, coat 50µl of 4µg/ml solution (dilute antigen in PBS)
   - Incubate at 4°C for o/n

Day2
1. Discard coated proteins and block the wells by adding 180µl of PBST + 5% milk
   Incubate at RT for 1hour
2. Wash 3 times with PBST (0.05% Tween20)
3. Add samples
   - 50µl/well, dilute in blocking solution (Serially dilute 1:2 or 1:√10)
   - Incubate at RT for 1hour
4. Wash 3 or 4 times with PBST (0.05% Tween20)
5. Add detection antibody
   - Typically antibody conjugated with HRP
   - 50µl/well, dilute in blocking solution (1:1000 ~ 1:5000)
   - Incubate at RT for 1hour
6. Wash 3 or 4 times with PBST (0.05% Tween20)
7. Develop by adding TMB solution (50µl/well)
8. Quench reaction by adding 1N HCl (50µl/well)
9. Read plate at 450nm

A.8 ISOLATION OF DU-SSDNA

(Adapted from Sidhu and Weiss (2004) Constructing Phage display libraries by oligonucleotide-directed mutagenesis. Phage Display. Eds. Clackson and Lowman)

Day1
1. Inoculate single colony of CJ236 with phagemid in 2ml TB (appropriate antibiotics) and grow cells at 37°C until OD_{600} is about 0.5
2. Add helper phage (final ~ 10^{10} virions/ml) and incubate additional 30min at 37°C
3. Transfer the culture to 30ml of 2xYT/amp/tet/uridine media and shake at 37°C for o/n

Day2
1. Pellet cells for 10min at 15,000rpm and transfer supernatant to a new bottle and precipitate the phage by adding 1/5V of 20% precipitation buffer (20% (w/v) PEG-8000, 2.5M NaCl). Incubate on rotor at 4°C for 20min
2. Centrifuge for 10min at 15,000rpm to harvest phage and decant the supernatant
3. Resuspend the phage pellet in 0.5mL of PBS (pH7.4) and centrifuge for 5minutes at max rpm to remove all debris
4. Isolate ss-dUDNA using QIAprep Spin M13 kit

A.9 PHAGE PRODUCTION AND PANNING

A.9.1 Helper phage preparation

(Adapted from Sidhu and Weiss and Kevin Entzminger's notes)

Reagents Needed:

- PBS: 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4, and pH7.2
- Phage precipitation buffer (20% PEG/NaCl): 20% PEG-8000, 2.5M NaCl
- M13K07 Helper Phage (NEB)
- 2xYT media and 2xYT agar plate freshly streaked with XL1-Blue E.coli cell
- 100% Glycerol (autoclaved)

Day1

1. Inoculate a single colony of XL1-Blue into 2ml of 2xYT/tet media and grow cells until OD_{600} = ~0.5
2. Infect M13K07 helper phage (~10^{10} pfu/ml) and incubate for 30min at 37°C
3. Scale up to 1L of 2xYT/tet/kan and grow cells at 37°C o/n with shaking

Day2

1. Pellet cells for 15min at 8000rpm
2. Transfer supernatant to a new bottle and precipitate the phage by adding 1/5V of 20% precipitation buffer
3. Vortex vigorously and incubate at R.T for 5min and at 4°C for additional 45min
4. Spin down for 10min at 8000rpm
5. Discard supernatant as much as you can
6. Resuspend precipitants using 6ml of PBS (pH7.2) and transfer them into the sterile e-tubes
7. Remove all contaminants after 5min centrifugation
   *Transfer supernatant to the new e-tubes and repeat this process until clear
8. Collect all 6ml of phages, pass through them using 0.2μm filter, and add same volume of 100% glycerol
9. Aliquot the phages and store at -20°C

A.9.2 Phage titration

1. Pick a colony of XL1-Blue from a freshly streaked plate and inoculate in 10ml of TB media with tetracyclin (5μl/ml). Grow cells at 37°C until OD_{600} is ~0.5
2. Dilute phages in TB media (dilution range: 10^{-6} to 10^{-10})
3. Split the grown XL1-Blue culture in sterile test tubes (0.5ml/tube) and infect phages by adding 10μl of diluted phages. Incubate at 37°C for 20min.
4. Plate 10μl and 100μl of each aliquot on 2xYT/Amp plate and incubate at 37°C o/n
5. The next day, count the number of colonies to determine the concentration of phage (colony forming unit per ml= cfu/ml) * Ideal number of colonies on a plate will be b/n 30-300.

A.9.3 Phage panning

Reagents Needed:
- PBS, PBST (0.5% Tween-20), and PBST + milk (5%)
- M13K07 Helper Phage (NEB)
- 2xYT media and 2xYT agar plate freshly streaked with XL1-Blue E.coli cell
- Acidic Elution Buffer (0.1M HCl in water, pH 2.2: pH adjust with glycine)
- Neutralization Buffer (2M Tris Base in water)
- Storage Buffer (PBS + 25% (w/v) glycrol)

Day 1
1. Transform 40µl of CCC-dsDNA to 400µl of XL1-Blue electrocompetent cells by electroporation
2. Add 3ml of TB (total) and transfer cells to 25ml of TB/tet/0.1% glucose. Incubate at 37°C for 20min
3. In order to measure library diversity, plate 10µl and 100µl on 2xYT/amp.
4. For the rest, add Amp and grow cells until OD_{600} ~0.5, and add IPTG (1mM). Shake at 25°C for 2h
   * Save 1ml of culture by mixing with 0.5ml of 50% glycerol at -80°C before adding IPTG
5. Add Helper phage (final ~ 10^{10} virions/ml) and incubate at 25°C for 2h
6. Add Kan (25µl/ml) and incubate at 25°C o/n
7. Coat target proteins (typically 4µg/ml, 50µl/well)

Day2
1. Precipitate phage (see above)
   * Label isolated phage as 'Input' and save 100µl phage at -20°C
2. Phage panning (*Save all phages at 4°C)
   i. Shake out coating solution and block wells with 300µl PBST + 5% milk at 25°C for at least 1h
   ii. Shake out blocking buffer and wash wells with PBST (3X)
   iii. Add 160µl Phage/milk solution and incubate at RT for 1h30min
   iv. Extract non-binders (*Save and label as 'Recovery')
   v. Add 200µl of PBST per well and pipette vigorously to wash and repeat 5 – 10X.
   vi. Add 50µl of Acidic Elution buffer per well and incubate at RT for 10min. Pipette vigorously and transfer to sterile eppendorf tube
   vii. Add 3µl of Neutralization buffer per 50µl of acidic buffer. (*Label this as 'Output')

A.9.4 Colony screening for phage library using ELISA

Reagents needed
- 96 well ELISA plates and sealers
- 96 well cell culture plate
- 2xYT Agar/amp plate
- TB media
- PBS, PBST (0.05% Tween20), PBST + 5% milk
- M13K07 Helper phage (10^{13} cfu/ml)

Day1
1. Pick a colony, streak onto an agar plate, and then pipette into each well containing 100µl of TB/tet/amp media
2. Grow cells at 37°C o/n

Day2
1. Induce antibodies with IPTG (add 0.25µl of IPTG in 50µl of TB/tet/amp per well) and allow cells to grow at 25°C for 3hrs
2. Add the M13K07 helper phages (add 0.25µl of helper phage(~10^{14} cfu/ml) in 50µl of TB/tet/amp per well) and shake another 2hrs
3. Add the Kanamycin (add 0.25µl of 25mg/ml of Kan in 50µl of TB/tet/amp per well) and shake at 25°C for o/n
4. Coat wells for ELISA (a-cmyc; 1µg/ml, and target proteins, 50µl/well)

Day3
1. Shake out the coating solutions and block the wells with PBST + 5% milk for 1hr at RT
2. Spin down cells for 30min at max speed
3. Shake out the blocking solution and wash with PBST for 3X
4. Take 50µl of supernatant and add them into each well, incubate for 1hr at RT
5. Shake out phages and wash with PBST for 4X
6. Add anti-M13-HRP (1:3000) and incubate at RT for 1hr
7. Shake out phages and wash with PBST for 4X
8. Add 50µl with TMB for develop
9. Add 50µl with 1N HCl to quench the reaction

A.10 REGENERATION OF IMAC RESIN

Used IMAC resin is stored in 20% EtOH.

Reagents Needed:

- Resin wash buffer: 50mM Tris 8.0, 50mM EDTA, 2% SDS, 0.5% b-Me
- Regeneration buffer: 6M GuHCl, 0.2M acetic acid)
- 25%, 50%, 75% and 100% EtOH
- 100mM EDTA (pH8.0)
- Recharging buffer: 500mM $NiSO_4$, $NiCl_2$, $CuCl_2$, or $CuSO_4$

Wash resin (in 50ml conical tube):

1. Spin down 1000g for 2min, decant sup
2. Wash with resin wash buffer by incubating at 37°C for 3hours
3. Repeat 3-4 times
4. (If you want to confirm all proteins removed from resin, run 10ul of washed resin on SDS-PAGE gel. If clean, proceed to the next step.)

Regeneration of resin

1. 5 volume of water (ddH_2O)
2. 2 volume of regeneration buffer
3. 5 volume of water
4. 3 volume of 2%SDS
5. 1 volume of 25% EtOH
6. 1 volume of 50% EtOH
7. 1 volume of 75% EtOH
8. 1 volume of 100% EtOH
9. 1 volume of 75% EtOH
10. 1 volume of 50% EtOH
11. 1 volume of 25% EtOH
12. 1 volume of water
13. 5 volume of 100mM EDTA
14. Wash with water until clear
15. Recharge with 2 volume of recharging buffer, rotate at room temperature at least 1hour
16. Wash with water until completely clear
17. Store in 2 volume of 20% EtOH

A.11 RNA EXTRACTION USING TRIZOL

*Use RNase-free filter tip
*All steps should be done in a hood
1. Weigh pre-tared mouse spleen in round bottom, 2ml eppendorf tube and add ~1ml Trizol/ 50mg tissue.
2. Homogenize spleen using plunger and needle method. Briefly, with the back of a sterile 5ml syringe plunger, mash the spleen into small enough debris. Use a new 1ml syringe and different gauge of needles (from 18G to 26G), repeat draw and expel the suspension several times until the viscosity drops.
3. Incubate for 5min at room temperature to allow complete dissociation of nucleoprotein complexes.
4. Centrifuge at max speed for 3min and transfer supernatant to a new eppendorf tube to remove remaining cell mass.
5. Add 0.1ml of 1-bromo-3-chloropropane in a tube and shake vigorously by hand for 15sec (DO NOT vortex).
6. Allow standing for 1min, shake additional 15sec, and centrifuge at max speed for 15min.
7. Carefully transfer the colorless upper phase (containing RNA) to a new tube.
8. Add 100% ethanol to final concentration of 35% and vortex to disperse any visible white precipitant.
9. Transfer the sample to the RNA spin cartridge (Invitrogen PureLink RNA kit) and centrifuge at max speed for 15sec. Discard the flow-through.
10. Add 700ul of Wash buffer I to the spin column and centrifuge at max speed for 15sec. Discard the flow-through.
11. Place the spin column into a new collection tube and add 500ul of Wash buffer II. Centrifuge at max speed for 15sec. Discard the flow-through.
12. Repeat steps 11 once.
13. Centrifuge the spin cartridge at max speed for 1min to dry the membrane with attached RNA
14. Transfer the spin column into an RNA recovery tube (supplied by the kit).
15. Add 100ul of RNase-free water to the center of membrane, and incubate for 1min at room temperature to elute the RNA.
16. Centrifuge for 2min at max speed and store the eluted RNA at -80°C.

Appendix B

B.1 Sequences of Engineered Proteins

B.1.1 pAK400/4D5.B1ag

```
E  V  Q  L  V  E  S  G  G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A  A  S  G  F  N
GAGGTTCAGCTGGTGGAGTCTGGCGGTGGCCTGGTGCAGCCAGGGGGCTCACTCCGTTTGTCCTGTGCAGCTTCTGGCTTCAAC
I  K  D  T  Y  I  G  W  V  R  R  A  P  G  K  G  E  E  L  V  A  R  I  Y  P  T  N  G
ATTAAAGACACCTATATAGGTTGGGTGCGTCGCGCCCCGGGTAAGGGCGAGGAATTGGTTGCAAGAATTTATCCTACGAATGGT
Y  T  R  Y  A  D  S  V  K  G  R  F  T  I  S  A  D  T  S  K  N  T  A  Y  L  Q  M  N
TATACTAGATATGCCGATAGCGTCAAGGGCCGTTTCACTATAAGCGCAGACACATCCAAAAACACAGCCTACCTACAAATGAAC
S  L  R  A  E  D  T  A  V  Y  Y  C  A  R  W  G  G  D  G  F  Y  A  M  D  Y  S  G  Q
AGCTTAAGAGCTGAGGACACTGCCGTCTATTATTGTGCTCGCTGGGGAGGGGACGGCTTCTATGCTATGGACTACTCGGGTCAA
G  T  L  V  T  V  S  S  A  S  G  A  D  H  H  H  H  H  H  *
GGAACACTAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

B.1.2 pAK400/4D5.B1ag_A5

```
E  V  Q  L  V  E  S  G  G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A  A  S  G  F  N
GAGGTTCAGCTGGTGGAGTCTGGCGGTGGCCTGGTGCAGCCAGGGGGCTCACTCCGTTTGTCCTGTGCAGCTTCTGGCTTCAAC
I  K  T  V  M  A  G  W  V  R  R  A  P  G  K  G  E  E  L  V  A  R  I  Y  P  Y  N  A
ATTAAAACGGTTATGGCTGGTTGGGTGCGTCGCGCCCCGGGTAAGGGCGAGGAATTGGTTGCAAGAATTTATCCTTATAATGCT
Y  T  D  Y  A  D  S  V  K  G  R  F  T  I  S  A  D  T  S  K  N  T  A  Y  L  Q  M  N
TATACTGATTATGCCGATAGCGTCAAGGGCCGTTTCACTATAAGCGCAGACACATCCAAAAACACAGCCTACCTACAAATGAAC
S  L  R  A  E  D  T  A  V  Y  Y  C  A  R  L  Y  S  H  S  P  V  Y  A  M  D  Y  S  G
AGCTTAAGAGCTGAGGACACTGCCGTCTATTATTGTGCTCGCTTGTACAGCCACAGCCCGGTGTATGCTATGGACTACTCGGGT
Q  G  T  L  V  T  V  S  S  A  S  G  A  D  H  H  H  H  H  H  *
CAAGGAACACTAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

B.1.3 pAK400/4D5.B1ag_G4

```
E  V  Q  L  V  E  S  G  G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A  A  S  G  F  N
GAGGTTCAGCTGGTGGAGTCTGGCGGTGGCCTGGTGCAGCCAGGGGGCTCACTCCGTTTGTCCTGTGCAGCTTCTGGCTTCAAC
I  K  S  V  M  Y  G  W  V  R  R  A  P  G  K  G  E  E  L  V  A  R  I  Y  P  Y  G  S
ATTAAAAGTGTTATGTATGGTTGGGTGCGTCGCGCCCCGGGTAAGGGCGAGGAATTGGTTGCAAGAATTTATCCTTATGGTTCT
A  T  Y  Y  A  D  S  V  K  G  R  F  T  I  S  A  D  T  S  K  N  T  A  Y  L  Q  M  N
GCTACTTATTATGCCGATAGCGTCAAGGGCCGTTTCACTATAAGCGCAGACACATCCAAAAACACAGCCTACCTACAAATGAAC
S  L  R  A  E  D  T  A  V  Y  Y  C  A  R  R  V  K  H  S  G  H  Y  A  M  D  Y  S  G
AGCTTAAGAGCTGAGGACACTGCCGTCTATTATTGTGCTCGCCGCGTCAAGCACAGCGGCCACTATGCTATGGACTACTCGGGT
Q  G  T  L  V  T  V  S  S  A  S  G  A  D  H  H  H  H  H  H  *
CAAGGAACACTAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

B.1.4 pAK400/4D5.B1ag_B6

```
E  V  Q  L  V  E  S  G  G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A  A  S  G  F  N
GAGGTTCAGCTGGTGGAGTCTGGCGGTGGCCTGGTGCAGCCAGGGGGCTCACTCCGTTTGTCCTGTGCAGCTTCTGGCTTCAAC
I  K  G  P  V  S  G  W  V  R  R  A  P  G  K  G  E  E  L  V  A  R  I  Y  P  T  N  G
ATTAAAGGTCCTGTGTCTGGTTGGGTGCGTCGCGCCCCGGGTAAGGGCGAGGAATTGGTTGCAAGAATTTATCCTACGAATGGT
Y  T  R  Y  A  D  S  V  K  G  R  F  T  I  S  A  D  T  S  K  N  T  A  Y  L  Q  M  N
TATACTAGATATGCCGATAGCGTCAAGGGCCGTTTCACTATAAGCGCAGACACATCCAAAAACACAGCCTACCTACAAATGAAC
S  L  R  A  E  D  T  A  V  Y  Y  C  A  R  S  R  V  L  S  T  D  Y  A  M  D  Y  S  G
AGCTTAAGAGCTGAGGACACTGCCGTCTATTATTGTGCTCGCTCGCGGGTCTTGAGCACCGACTATGCTATGGACTACTCGGGT
Q  G  T  L  V  T  V  S  S  A  S  G  A  D  H  H  H  H  H  H  *
CAAGGAACACTAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

B.1.5 pBAD33/Skp

```
A  D  K  I  A  I  V  N  M  G  S  L  F  Q  Q  V  A  Q  K  T  G  V  S  N  T  L  E  N
GCTGACAAAATTGCAATCGTCAACATGGGCAGCCTGTTCCAGCAGGTAGCGCAGAAAACCGGTGTTTCTAACACGCTGGAAAAT
E  F  K  G  R  A  S  E  L  Q  R  M  E  T  D  L  Q  A  K  M  K  K  L  Q  S  M  K  A
GAGTTCAAAGGCCGTGCCAGCGAACTGCAGCGTATGGAAACCGATCTGCAGGCTAAAATGAAAAAGCTGCAGTCCATGAAAGCG
G  S  D  R  T  K  L  E  K  D  V  M  A  Q  R  Q  T  F  A  Q  K  A  Q  A  F  E  Q  D
GGCAGCGATCGCACTAAGCTGGAAAAAGACGTGATGGCTCAGCGCCAGACTTTTGCTCAGAAAGCGCAGGCTTTTGAGCAGGAT
R  A  R  R  S  N  E  E  R  G  K  L  V  T  R  I  Q  T  A  V  K  S  V  A  N  S  Q  D
CGCGCACGTCGTTCCAACGAAGAACGCGGCAAACTGGTTACTCGTATCCAGACTGCTGTGAAATCCGTTGCCAACAGCCAGGAT
I  D  L  V  V  D  A  N  A  V  A  Y  N  S  S  D  V  K  D  I  T  A  D  V  L  K  Q  V
ATCGATCTGGTTGTTGATGCAAACGCCGTTGCTTACAACAGCAGCGATGTAAAAGACATCACTGCCGACGTACTGAAACAGGTT
K
AAA
```

B.1.6 pFabF/EE

```
D  Y  K  D  I  V  M  T  Q  T  P  S  S  L  P  V  S  L  G  D  Q  A  S  I  S  C  R  S
GACTACAAAGATATTGTGATGACCCAGACTCCATCCTCCCTGCCTGTCAGTCTTGGAGATCAAGCCTCCATCTCTTGCAGATCT
S  Q  S  I  V  H  S  N  G  N  T  Y  L  E  W  Y  L  Q  K  P  G  Q  S  P  K  L  L  I
AGTCAGAGCATTGTACATAGTAATGGAAACACCTATTTAGAATGGTACCTGCAGAAACCAGGCCAGTCTCCAAAGCTCCTGATC
Y  K  V  S  N  R  F  S  G  V  P  D  R  F  S  G  S  G  S  G  T  D  F  T  L  K  I  S
TACAAAGTTTCCAACCGATTTTCTGGGGTCCCAGACAGGTTCAGTGGCAGTGGATCAGGGACAGATTTCACACTCAAGATCAGC
R  V  E  A  E  D  L  G  I  Y  Y  C  F  Q  G  S  L  V  P  P  T  F  G  A  G  T  K  L
AGAGTGGAGGCTGAGGATCTGGGAATTTATTACTGCTTTCAAGGTTCACTTGTTCCTCCCACGTTCGGTGCTGGGACCAAGCTG
E  L  K  R  G  A  A  A  P  S  V  F  I  F  P  P  S  D  E  Q  L  K  S  G  T  A  S  V
GAGCTGAAACGTGGTGCGGCCGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTT
V  C  L  L  N  N  F  Y  P  R  E  A  K  V  Q  W  K  V  D  N  A  L  Q  S  G  N  S  Q
GTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAG
E  S  V  T  E  Q  D  S  K  D  S  T  Y  S  L  S  S  T  L  T  L  S  K  A  D  Y  E  K
GAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAA
H  K  V  Y  A  C  E  V  T  H  Q  G  L  S  S  P  V  T  K  S  F  N  R  G  E  C  G  G
CACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGTTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGCGGCGGT
G  G  H  H  H  H  H  H  H  H  H  H  .  .
GGCGGTCACCATCACCACCATCATCATCACCACCATTAATAA
```

```
Q  V  Q  L  Q  Q  S  G  P  E  D  V  K  P  G  A  S  V  K  I  S  C  K  A  S  G  Y  S
CAGGTTCAGCTGCAGCAGTCTGGGCCTGAGGATGTGAAGCCCGGCGCGAGCGTGAAAATCAGTTGTAAAGCCTCTGGATATTCA
L  S  T  S  G  M  G  V  N  W  V  K  Q  S  P  G  K  G  L  E  W  L  A  H  I  Y  W  D
CTGAGCACTTCTGGTATGGGTGTGAACTGGGTTAAACAGAGCCAGGAAAGGGTCTGGAGTGGCTGGCACACATTTACTGGGATG
D  D  K  R  Y  N  P  S  L  K  S  R  A  T  L  T  V  D  K  S  S  S  T  V  Y  L  E  L
ACGATCAAGCGCTATAACCCATCCCTGAAGAGCAGGGCCACACTCACCGTGGATAAGTCCAGCAGCACGGTATACCTCGAGCTC
R  S  L  T  S  E  D  S  S  V  Y  Y  C  A  R  R  G  G  S  S  H  Y  Y  A  M  D  Y  W
AGGAGTCTGACCAGCGAAGATAGTTCCGTATACTACTGTGCTCGAAGAGGGGGTAGTTCCCATTACTATGCTATGGACTACTGG
G  Q  G  T  T  V  T  V  S  S  A  S  F  K  G  P  S  V  F  P  L  A  P  S  S  K  S  T
GGTCAAGGAACCACAGTCACCGTCTCCTCGGCAAGCTTCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACC
S  G  G  T  A  A  L  G  C  L  V  K  D  Y  F  P  E  P  V  T  V  S  W  N  S  G  A  L
TCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTG
T  S  G  V  H  T  F  P  A  V  L  Q  S  S  G  L  Y  S  L  S  S  V  V  T  V  P  S  S
ACCAGCGGCGTGCACACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGC
S  L  G  T  Q  T  Y  I  C  N  V  N  H  K  P  S  N  T  K  V  D  K  K  V  E  P  K  S
AGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCT
C  D  K  T  H  T  G  G  G  G  D  Y  K  D  D  D  D  K  .  .
TGTGACAAAACTCACACAGGCGGTGGCGGTGATTACAAAGATGACGATGACAAATAATGA
```

B.1.7 pFabF/EEh14.3

```
D  I  V  M  T  Q  T  P  S  S  L  P  V  S  L  G  D  Q  A  S  I  S  C  R  S  S  Q  S
GACATAGTAATGACACAAACACCAAGCAGCCTACCAGTAAGCCTAGGAGACCAAGCAAGCATAAGCTGCAGAAGCAGCCAAAGC
I  V  H  S  N  G  N  T  Y  L  E  W  Y  L  Q  K  P  G  Q  S  P  K  L  L  I  Y  K  V
ATAGTACACAGCAACGGAAACACATACCTAGAATGGTACCTACAAAAACCAGGACAAAGCCCAAAACTACTAATATACAAAGTA
S  N  R  F  S  G  V  P  D  R  F  S  G  S  G  S  G  T  D  F  T  L  K  I  S  R  V  E
AGCAACAGATTCAGCGGAGTACCAGACAGATTCAGCGGAAGCGGAAGCGGAACAGACTTCACACTAAAAATAAGCAGAGTAGAA
A  E  D  L  G  I  Y  Y  C  F  Q  G  S  L  V  P  P  T  F  G  A  G  T  K  L  E  L  K
GCAGAAGACCTAGGAATATACTACTGCTTCCAAGGAAGCCTAGTACCACCAACATTCGGAGCAGGAACAAAACTAGAACTAAAA
R  G  A  A  A  P  S  V  F  I  F  P  P  S  D  E  Q  L  K  S  G  T  A  S  V  V  C  L
AGAGGAGCGGCCGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTG
L  N  N  F  Y  P  R  E  A  K  V  Q  W  K  V  D  N  A  L  Q  S  G  N  S  Q  E  S  V
CTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTC
T  E  Q  D  S  K  D  S  T  Y  S  L  S  S  T  L  T  L  S  K  A  D  Y  E  K  H  K  V
ACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTC
Y  A  C  E  V  T  H  Q  G  L  S  S  P  V  T  K  S  F  N  R  G  E  C  G  G  G  G  H
TACGCCTGCGAAGTCACCCATCAGGGCCTGAGTTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGCGGCGGTGGCGGTCAC
H  H  H  H  H  H  H  H  H  H
CATCACCACCATCATCATCACCACCAT

Q  V  Q  L  Q  Q  S  G  P  E  D  V  K  P  G  A  S  V  K  I  S  C  K  A  S  G  D  S
CAAGTACAACTACAACAAAGCGGACCAGAAGACGTAAAACCAGGAGCAAGCGTAAAAATAAGCTGCAAAGCAAGCGGAGACAGC
L  S  S  Y  N  A  G  V  N  W  V  K  Q  S  P  G  K  G  L  E  W  L  A  H  M  A  G  V
CTGAGCAGCTACAACGCAGGAGTAAACTGGGTAAAACAAAGCCCAGGAAAAGGACTAGAATGGCTAGCACACATGGCAGGCGTG
S  T  R  Y  N  P  S  L  K  S  R  A  T  L  T  V  D  K  T  S  S  T  V  Y  L  E  L  R
AGCACAAGATACAACCCAAGCCTAAAAAGCAGAGCAACACTAACAGTAGACAAAACAAGCAGCACAGTATACCTAGAACTAAGA
S  L  T  S  E  D  S  S  V  Y  Y  C  V  R  N  E  W  S  G  A  F  W  G  Q  G  T  T  V
AGCCTAACAAGCGAAGACAGCAGCGTATACTACTGCGTGCGCAACGAATGGAGCGGCGCGTTCTGGGGACAAGGAACAACAGTA
T  V  S  S  G  A  S  F  K  G  P  S  V  F  P  L  A  P  S  S  K  S  T  S  G  G  T  A
ACAGTAAGCAGCGGGGCAAGCTTCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCG
A  L  G  C  L  V  K  D  Y  F  P  E  P  V  T  V  S  W  N  S  G  A  L  T  S  G  V  H
GCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
T  F  P  A  V  L  Q  S  S  G  L  Y  S  L  S  S  V  V  T  V  P  S  S  S  L  G  T  Q
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAG
T  Y  I  C  N  V  N  H  K  P  S  N  T  K  V  D  K  K  V  E  P  K  S  C  D  K  T  H
ACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCAC
T  G  G  G  G
ACAGGCGGTGGCGGT
```

B.1.8 pFabF/EEh13.6

D I V M T Q T P S S L P V S L G D Q A S I S C R S S Q S
GACATAGTAATGACACAAACACCAAGCAGCCTACCAGTAAGCCTAGGAGACCAAGCAAGCATAAGCTGCAGAAGCAGCCAAAGC
I V H S N G N T Y L E W Y L Q K P G Q S P K L L I Y K V
ATAGTACACAGCAACGGAAACACATACCTAGAATGGTACCTACAAAAACCAGGACAAAGCCCAAAACTACTAATATACAAAGTA
S N R F S G V P D R F S G S G S G T D F T L K I S R V E
AGCAACAGATTCAGCGGAGTACCAGACAGATTCAGCGGAAGCGGAAGCGGAACAGACTTCACACTAAAAATAAGCAGAGTAGAA
A E D L G I Y Y C F Q G S L V P P T F G A G T K L E L K
GCAGAAGACCTAGGAATATACTACTGCTTCCAAGGAAGCCTAGTACCACCAACATTCGGAGCAGGAACAAAACTAGAACTAAAA
R G A A A P S V F I F P P S D E Q L K S G T A S V V C L
AGAGGAGCGGCCGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTG
L N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S V
CTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTC
T E Q D S K D S T Y S L S S T L T L S K A D Y E K H K V
ACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTC
Y A C E V T H Q G L S S P V T K S F N R G E C G G G G H
TACGCCTGCGAAGTCACCCATCAGGGCCTGAGTTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGCGGCGGTGGCGGTCAC
H H H H H H H H H H
CATCACCACCATCATCATCACCACCAT

Q V Q L Q Q S G P E D V K P G A S V K I S C K A S G D S
CAAGTACAACTACAACAAAGCGGACCAGAAGACGTAAAACCAGGAGCAAGCGTAAAAATAAGCTGCAAAGCAAGCGGAGACAGC
L S S F N A G V N W V K Q S P G K G L E W L A H G A V M
CTGAGCAGCTTCAACGCAGGAGTAAACTGGGTAAAACAAAGCCCAGGAAAAGGACTAGAATGGCTAGCACACGGAGCAGTAATG
S T R Y N P S L K S R A T L T V D K T S S T V Y L E L R
AGCACAAGATACAACCCAAGCCTAAAAAGCAGAGCAACACTAACAGTAGACAAAACAAGCAGCACAGTATACCTAGAACTAAGA
S L T S E D S S V Y Y C A K S T G R Y D F W G Q G T T V
AGCCTAACAAGCGAAGACAGCAGCGTATACTACTGCGCAAAAAGCACAGGAAGATACGACTTCTGGGGACAAGGAACAACAGTA
T V S S G A S F K G P S V F P L A P S S K S T S G G T A
ACAGTAAGCAGCGGGGCAAGCTTCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCG
A L G C L V K D Y F P E P V T V S W N S G A L T S G V H
GCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
T F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAG
T Y I C N V N H K P S N T K V D K K V E P K S C D K T H
ACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCAC
T G G G G
ACAGGCGGTGGCGGT

B.1.9 pFabF/EEf15.4

D I V M T Q T P S S L P V S L G D Q A S I S C R S S N A
GACATAGTAATGACACAAACACCAAGCAGCCTACCAGTAAGCCTAGGAGACCAAGCAAGCATAAGCTGCCGTTCTTCTAACGCT
R S G S L E W Y L Q K P G Q S P K L L I Y D G N N R F S
CGTTCTGGTTCTCTGGAATGGTACCTACAAAAACCAGGACAAAGCCCAAAACTACTAATATACGACGGTAACAACCGTTTCTCT
G V P D R F S G S G S G T D F T L K I S R V E A E D L G
GGAGTACCAGACAGATTCAGCGGAAGCGGAAGCGGAACAGACTTCACACTAAAAATAAGCAGAGTAGAAGCAGAAGACCTAGGA
I Y Y C S A F D Q T N K Y V G F G A G T K L E L K R G A
ATATACTACTGCTCTGCGTTCGATCAGACCAACAAATACGTGGGCTTCGGAGCAGGAACAAAACTAGAACTAAAAAGAGGAGCG
A A P S V F I F P P S D E Q L K S G T A S V V C L L N N
GCCGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAAC
F Y P R E A K V Q W K V D N A L Q S G N S Q E S V T E Q
TTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAG
D S K D S T Y S L S S T L T L S K A D Y E K H K V Y A C
GACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGC
E V T H Q G L S S P V T K S F N R G E C G G G G H H H H
GAAGTCACCCATCAGGGCCTGAGTTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGCGGCGGTGGCGGTCACCATCACCAC
H H H H H H
CATCATCATCACCACCAT

Q V Q L Q Q S G P E D V K P G A S V K I S C K A S G D S
CAAGTACAACTACAACAAAGCGGACCAGAAGACGTAAAACCAGGAGCAAGCGTAAAAATAAGCTGCAAAGCAAGCGGAGACAGC
L S S F N A G V N W V K Q S P G K G L E W L A H G A V M
CTGAGCAGCTTCAACGCAGGAGTAAACTGGGTAAAACAAAGCCCAGGAAAAGGACTAGAATGGCTAGCACACGGAGCAGTAATG
S T R Y N P S L K S R A T L T V D K T S S T V Y L E L R
AGCACAAGATACAACCCAAGCCTAAAAAGCAGAGCAACACTAACAGTAGACAAAACAAGCAGCACAGTATACCTAGAACTAAGA
S L T S E D S S V Y Y C A K S T G R Y D F W G Q G T T V
AGCCTAACAAGCGAAGACAGCAGCGTATACTACTGCGCAAAAAGCACAGGAAGATACGACTTCTGGGGACAAGGAACAACAGTA
T V S S G A S F K G P S V F P L A P S S K S T S G G T A
ACAGTAAGCAGCGGGGCAAGCTTCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCG
A L G C L V K D Y F P E P V T V S W N S G A L T S G V H
GCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
T F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAG
T Y I C N V N H K P S N T K V D K K V E P K S C D K T H
ACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCAC
T G G G G
ACAGGCGGTGGCGGT

B.1.10 pAk400/14B7_Flag (N-term)

```
D  I  V  M  T  Q  T  P  S  S  L  P  V  S  L  G  D  Q  A  S  I  S  C  R  S  S  N  A
GACATAGTAATGACACAAACACCAAGCAGCCTACCAGTAAGCCTAGGAGACCAAGCAAGCATAAGCTGCCGTTCTTCTAACGCT

D  Y  K  D  I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I  S  C  R  A
GACTACAAAGACATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCA
S  Q  D  I  R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y  Y  T  S  R
AGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGA
L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N  L  E  Q  E
TTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAA
D  I  G  T  Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E  I  R  R  G
GATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGT
G  G  G  S  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q  Q  S  G  P
GGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCT
E  L  V  K  P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W  M  N  W  V
GAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTG
K  Q  R  P  G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N  G  K  F  K
AAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAG
G  K  A  T  L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V  D  S  A  V
GGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTC
Y  F  C  A  R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V  S  S  A  S
TATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTC**GGCC**TCG
G  A  D  H  H  H  H  H  H  *
GG**GGCC**GATCACCATCATCACCATCATTAG
```

B.1.11 pAk400/14B7_Flag_mut1

```
D  Y  K  D  D  D  D  K  I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I
GACTACAAAGACGATGACGACAAGATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATC
S  C  R  A  S  Q  D  I  R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y
AGTTGCAGGGCAAGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTAC
Y  T  S  R  L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N
TACACATCAAGATTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAAC
L  E  Q  E  D  I  G  T  Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E
CTGGAGCAAGAAGATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAA
I  R  R  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q
ATAAGACGTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAG
Q  S  G  P  E  L  V  K  P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W
CAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGG
M  N  W  V  K  Q  R  P  G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N
ATGAACTGGGTGAAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAAT
G  K  F  K  G  K  A  T  L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V
GGGAAGTTCAAGGGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTG
D  S  A  V  Y  F  C  A  R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V
GACTCTGCGGTCTATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTC
S  S  A  S  G  A  D  H  H  H  H  H  H  *
TCCTC**GGCC**TCGGG**GGCC**GATCACCATCATCACCAT CATTAG
```

B.1.12 pAk400/14B7_Flag_mut2

A D Y K D I V L I Q S T S S L S A S L G D R V T I S C R
GCAGACTACAAAGACATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGG
A S Q D I R N Y L N W Y Q Q K P D G T V K L L I Y Y T S
GCAAGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCA
R L Q S G V P S R F S G S G S G T D Y S L T I S N L E Q
AGATTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAA
E D I G T Y F C Q Q G N T L P W T F G G G T K L E I R R
GAAGATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGT
G G G G S G G G G S G G G G S G G G G S E V Q L Q Q S G
GGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGA
P E L V K P G A S V K I S C K D S G Y A F S S S W M N W
CCTGAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGG
V K Q R P G Q G P E W I G R I Y P G D G D T N Y N G K F
GTGAAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTC
K G K A T L T A D K S S S T A Y M Q L S S L T S V D S A
AAGGGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCG
V Y F C A R S G L L R Y A M D Y W G Q G T S V T V S S A
GTCTATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTC**GGCC**
S G A D H H H H H H *
TCGGG**GGCC**GATCACCATCATCACCATCATTAG

B.1.13 pAk400/14B7_Flag_mut3

D F K D I V L I Q S T S S L S A S L G D R V T I S C R A
GACTTCAAAGACATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCA
S Q D I R N Y L N W Y Q Q K P D G T V K L L I Y Y T S R
AGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGA
L Q S G V P S R F S G S G S G T D Y S L T I S N L E Q E
TTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAA
D I G T Y F C Q Q G N T L P W T F G G G T K L E I R R G
GATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGT
G G G S G G G G S G G G G S G G G G S E V Q L Q Q S G P
GGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCT
E L V K P G A S V K I S C K D S G Y A F S S S W M N W V
GAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTG
K Q R P G Q G P E W I G R I Y P G D G D T N Y N G K F K
AAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAG
G K A T L T A D K S S S T A Y M Q L S S L T S V D S A V
GGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTC
Y F C A R S G L L R Y A M D Y W G Q G T S V T V S S A S
TATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTC**GGCC**TCG
G A D H H H H H H *
GG**GGCC**GATCACCATCATCACCATCATTAG

B.1.14 pAk400/14B7_Flag_mut4

```
D  Y  R  D  I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I  S  C  R  A
GACTACAGGGACATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCA
S  Q  D  I  R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y  Y  T  S  R
AGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGA
L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N  L  E  Q  E
TTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAA
D  I  G  T  Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E  I  R  R  G
GATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGT
G  G  G  S  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q  Q  S  G  P
GGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCT
E  L  V  K  P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W  M  N  W  V
GAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTG
K  Q  R  P  G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N  G  K  F  K
AAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAG
G  K  A  T  L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V  D  S  A  V
GGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTC
Y  F  C  A  R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V  S  S  A  S
TATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCGGCCTCG
G  A  D  H  H  H  H  H  H  *
GGGGCCGATCACCATCATCACCATCATTAG
```

B.1.15 pAk400/14B7_Flag_mut5

```
E  Y  K  E  I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I  S  C  R  A
GAATACAAAGAAATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCA
S  Q  D  I  R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y  Y  T  S  R
AGTCAGGACATTAGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGA
L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N  L  E  Q  E
TTACAGTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAA
D  I  G  T  Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E  I  R  R  G
GATATTGGCACTTACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGT
G  G  G  S  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q  Q  S  G  P
GGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCT
E  L  V  K  P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W  M  N  W  V
GAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTG
K  Q  R  P  G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N  G  K  F  K
AAGCAGAGGCCTGGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAG
G  K  A  T  L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V  D  S  A  V
GGCAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTC
Y  F  C  A  R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V  S  S  A  S
TATTTCTGTGCAAGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCGGCCTCG
G  A  D  H  H  H  H  H  H  *
GGGGCCGATCACCATCATCACCATCATTAG
```

B.1.16 pAk400/14B7_Flag_mut6

```
I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I  S  C  R  A  S  Q  D  I
ATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCAAGTCAGGACATT
R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y  Y  T  S  R  L  Q  S  G
AGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGATTACAGTCAGGA
V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N  L  E  Q  E  D  I  G  T
GTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAAGATATTGGCACT
Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E  I  R  R  G  G  G  G  S
TACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGTGGTGGTGGTTCT
G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q  Q  S  G  P  E  L  V  K
GGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCTGAGCTGGTGAAG
P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W  M  N  W  V  K  Q  R  P
CCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTGAAGCAGAGGCCT
G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N  G  K  F  K  G  K  A  T
GGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAGGGCAAGGCCACA
L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V  D  S  A  V  Y  F  C  A
CTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTCTATTTCTGTGCA
R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V  S  S  A  S  G  A  D  H
AGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCAC
H  H  H  H  H  *
CATCATCACCATCATTAG
```

B.1.17 pAk400/14B7_Flag (C-term)

```
I  V  L  I  Q  S  T  S  S  L  S  A  S  L  G  D  R  V  T  I  S  C  R  A  S  Q  D  I
ATTGTTCTCATCCAGTCTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACCATCAGTTGCAGGGCAAGTCAGGACATT
R  N  Y  L  N  W  Y  Q  Q  K  P  D  G  T  V  K  L  L  I  Y  Y  T  S  R  L  Q  S  G
AGGAATTATTTAAACTGGTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACTACACATCAAGATTACAGTCAGGA
V  P  S  R  F  S  G  S  G  S  G  T  D  Y  S  L  T  I  S  N  L  E  Q  E  D  I  G  T
GTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCTCACCATTAGCAACCTGGAGCAAGAAGATATTGGCACT
Y  F  C  Q  Q  G  N  T  L  P  W  T  F  G  G  G  T  K  L  E  I  R  R  G  G  G  G  S
TACTTTTGCCAACAGGGTAATACGCTTCCGTGGACGTTCGGTGGAGGCACCAAGCTGGAAATAAGACGTGGTGGTGGTGGTTCT
G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  V  Q  L  Q  Q  S  G  P  E  L  V  K
GGTGGTGGTGGTTCTGGCGGCGGCGGCTCCGGTGGTGGTGGATCCGAGGTTCAGCTTCAGCAGTCTGGACCTGAGCTGGTGAAG
P  G  A  S  V  K  I  S  C  K  D  S  G  Y  A  F  S  S  S  W  M  N  W  V  K  Q  R  P
CCTGGGGCCTCAGTGAAGATTTCCTGCAAAGATTCTGGCTACGCATTCAGTAGCTCTTGGATGAACTGGGTGAAGCAGAGGCCT
G  Q  G  P  E  W  I  G  R  I  Y  P  G  D  G  D  T  N  Y  N  G  K  F  K  G  K  A  T
GGACAGGGTCCTGAGTGGATTGGACGGATTTATCCTGGAGATGGAGATACTAACTACAATGGGAAGTTCAAGGGCAAGGCCACA
L  T  A  D  K  S  S  S  T  A  Y  M  Q  L  S  S  L  T  S  V  D  S  A  V  Y  F  C  A
CTGACTGCAGACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGCGGTCTATTTCTGTGCA
R  S  G  L  L  R  Y  A  M  D  Y  W  G  Q  G  T  S  V  T  V  S  S  D  Y  K  D  D  D
AGGTCGGGGTTACTACGTTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCGGATTACAAGGATGATGAC
D  K  A  S  G  A  D  H  H  H  H  H  H  *
GACAAGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

B.1.18 pAK400/DO11.10

```
E  A  A  V  T  Q  S  P  R  N  K  V  A  V  T  G  E  K  V  T  L  S  C  N  Q  T  N  N
GAGGCTGCAGTCACCCAAAGCCCAAGAAACAAGGTGGCAGTAACAGGAGAGAAGGTGACATTGAGCTGTAATCAGACTAATAAC
H  N  N  M  Y  W  Y  R  Q  D  T  G  H  G  L  R  L  I  Y  Y  S  Y  G  A  G  S  T  E
CACAACAACATGTACTGGTATCGGCAGGACACGGGGCATGGGCTGAGGCTGATCTACTATTCATATGGTGCTGGCAGCACTGAG
K  G  D  I  P  D  G  Y  K  A  S  R  P  S  Q  E  N  F  S  L  T  L  E  S  A  T  P  S
AAAGGAGATATCCCTGATGGATACAAGGCCTCTAGACCAAGCCAAGAGAACTTCTCCCTCACTCTGGAGTCGGCTACCCCCTCT
Q  T  S  V  Y  F  C  A  S  G  S  G  T  T  N  T  E  V  F  F  G  K  G  T  R  L  T  V
CAGACATCAGTGTACTTCTGTGCCAGCGGCTCCGGGACAACAAACACAGAAGTCTTCTTTGGTAAAGGAACCAGACTCACAGTT
V  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  E  Q  V  E  Q  L  P
GTAGGTGGAGGCGGTTCAGGCGGAGGTGGCTCCGGAGGCGGTGGTTCTGGAGGAGGAGGATCCGAGCAGGTGGAGCAGCTTCCT
S  I  L  R  V  Q  E  G  S  S  A  S  I  N  C  T  Y  E  N  S  A  S  N  Y  F  P  W  Y
TCCATCCTGAGAGTCCAGGAGGGATCCAGTGCCAGCATCAACTGCACTTATGAGAACAGTGCCTCCAACTACTTCCCTTGGTAT
K  Q  E  P  G  E  N  P  K  L  I  I  D  I  R  S  N  M  E  R  K  Q  T  Q  G  L  I  V
AAGCAAGAACCTGGAGAGAATCCTAAGCTCATCATTGACATTCGTTCAAATATGGAAAGAAAGCAGACCCAAGGACTCATCGTT
L  L  D  K  K  A  K  R  F  S  L  H  I  T  D  T  Q  P  G  D  S  A  M  Y  F  C  A  A
TTACTGGATAAGAAAGCCAAACGCTTCTCCCTGCACATCACAGACACCCAGCCTGGAGACTCAGCCATGTACTTCTGTGCTGCA
S  P  N  Y  N  V  L  Y  F  G  S  G  T  K  L  T  V  E  P  N  A  A  S  G  A  D  H  H
AGTCCTAATTACAACGTGCTTTACTTCGGATCTGGCACCAAACTCACTGTAGAGCCAAACGCGGCCTCGGGGGCCGATCACCAT
H  H  H  H  *
CATCACCATCATTAG
```

B.1.19 pAK400/3D5g48

```
D  Y  K  D  I  V  M  T  Q  T  P  S  S  L  P  V  S  L  G  D  Q  A  S  I  S  C  R  S
GACTACAAAGATATTGTGATGACCCAGACTCCATCCTCCCTGCCTGTCAGTCTTGGAGATCAAGCCTCCATCTCTTGCAGATCT
S  Q  S  I  V  H  S  N  G  N  T  Y  L  E  W  Y  L  Q  K  P  G  Q  S  P  K  L  L  I
AGTCAGAGCATTGTACATAGTAATGGAAACACCTATTTAGAATGGTACCTGCAGAAACCAGGCCAGTCTCCAAAGCTCCTGATC
Y  K  V  S  N  R  F  S  G  V  P  D  R  F  S  G  S  G  S  G  T  D  F  T  L  K  I  S
TACAAAGTTTCCAACCGATTTTCTGGGGTCCCAGACAGGTTCAGTGGCAGTGGATCAGGGACAGATTTCACACTCAAGATCAGC
R  V  E  A  E  D  L  G  I  Y  Y  C  F  Q  G  S  L  V  P  P  T  F  G  A  G  T  K  L
AGAGTGGAGGCTGAGGATCTGGGAATTTATTACTGCTTTCAAGGTTCACTTGTTCCTCCCACGTTCGGTGCTGGGACCAAGCTG
E  L  K  R  G  G  G  G  S  G  G  G  G  S  G  G  G  G  S  S  G  G  G  S  Q  V  Q  L
GAGCTGAAACGTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGCGGCGGCGGCTCCAGTGGTGGTGGATCCCAGGTTCAGCTG
Q  Q  S  G  P  E  D  V  K  P  G  A  S  V  K  I  S  C  K  A  S  G  Y  S  L  S  T  S
CAGCAGTCTGGGCCTGAGGATGTGAAGCCCGGCGCGAGCGTGAAAATCAGTTGTAAAGCCTCTGGATATTCACTGAGCACTTCT
G  M  G  V  N  W  V  K  Q  S  P  G  K  G  L  E  W  L  A  H  I  Y  W  D  D  D  K  R
GGTATGGGTGTGAACTGGGTTAAACAGAGCCAGGAAAGGGTCTGGAGTGGCTGGCACACATTTACTGGGATGACGATCAAGCGC
Y  N  P  S  L  K  S  R  A  T  L  T  V  D  K  S  S  S  T  V  Y  L  E  L  R  S  L  T
TATAACCCATCCCTGAAGAGCAGGGCCACACTCACCGTGGATAAGTCCAGCAGCACGGTATACCTCGAGCTCAGGAGTCTGACC
S  E  D  S  S  V  Y  Y  C  A  R  R  G  G  S  S  H  Y  Y  A  M  D  Y  W  G  Q  G  T
AGCGAAGATAGTTCCGTATACTACTGTGCTCGAAGAGGGGGTAGTTCCCATTACTATGCTATGGACTACTGGGGTCAAGGAACC
T  V  T  V  S  S  A  S  G  A  D  H  H  H  H  H  H  *
ACAGTCACCGTCTCCTCGGCCTCGGGGGCCGATCACCATCATCACCATCATTAG
```

Glossary/Abbreviation

ccc-dsDNA: covalently closed circular heteroduplex double-stranded DNA

CDR: complementary determining regions

cfu: colony forming units

dU-ssDNA: uracil containing single-stranded DNA

EE: EYMPME peptide

ELISA: enzyme-linked immunosorbent assay

Fab: fragment antigen-binding

Flag: DYKDDDDK peptide

IMAC: immobilized metal affinity chromatography

nAb: nanobody

scFv: single-chain variable fragment

sdAb: single domain antibody

SEC: size exclusion chromatography

TMB: 3,3',5,5'-tetramethylbenzidine substrate

V_H: variable heavy chain

V_L: variable light chain

References

1. Griffin, L. & Lawson, a. Antibody fragments as tools in crystallography. *Clin. Exp. Immunol.* **165,** 285–91 (2011).
2. Koide, S. Engineering of recombinant crystallization chaperones. *Curr. Opin. Struct. Biol.* **19,** 449–57 (2009).
3. Tereshko, V. *et al.* Toward chaperone-assisted crystallography : Protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (V H H) scaffold. 1175–1187 (2008). doi:10.1110/ps.034892.108.1
4. Löw, C. *et al.* Nanobody mediated crystallization of an archeal mechanosensitive channel. *PLoS One* **8,** e77984 (2013).
5. Ermolenko, D. N., Zherdev, A. V, Dzantiev, B. B. & Popov, V. O. Antiperoxidase Antibodies Enhance Refolding of Horseradish Peroxidase. *Biochem. Biophys. Res. Commun.* **965,** 959–965 (2002).
6. Pai, J. C. *et al.* Conversion of scFv peptide-binding specificity for crystal chaperone development. *Protein Eng. Des. Sel.* **24,** 419–28 (2011).
7. Johnson, J. L. *et al.* Structural and biophysical characterization of an epitope-specific engineered Fab fragment and complexation with membrane proteins : implications for co-crystallization research papers. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 896–906 (2015).
8. Roosild, T. P. & Choe, S. Structure of anti-FLAG M2 Fab domain and its use in the stabilization of engineered membrane proteins. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62,** 835–839 (2006).
9. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. Making antibodies fragments using phage display libraries. *Nature* **352,** 624–628 (1991).
10. Barbas, C. F., Kang, A. S., Lerner, R. A. & Benkovict, S. J. Assembly of combinatorial antibody libraries on phage surfaces : The gene III site. *Proc. Natl. Acad. Sci.* **88,** 7978–7982 (1991).
11. Hoet, M. *et al.* Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nat. Biotechnol.* **23,** 344–348 (2005).
12. Krebber, A. *et al.* Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods* **201,** 35–55 (1997).
13. Presta, L. G. Engineering of therapeutic antibodies to minimize immunogenicity and optimize function B. *Adv. Drug Deliv. Rev.* **58,** 640–656 (2006).

14. Knappik, A. *et al*. Fully Synthetic Human Combinatorial Antibody Libraries (HuCAL) Based on Modular Consensus Frameworks and CDRs Randomized with Trinucleotides. *J. Mol. Biol.* **296,** 57–86 (2000).

15. Sidhu, S. S. *et al*. Phage-displayed Antibody Libraries of Synthetic Heavy Chain Complementarity Determining Regions. *J. Mol. Biol.* **338,** 299–310 (2004).

16. Morea, V., Lesk, A. M. & Tramontano, A. Antibody Modeling : Implications for Engineering and Design. *Methods* **279,** 267–279 (2000).

17. Clark, L. A. *et al*. An antibody loop replacement design feasibility study and a loop-swapped dimer structure. *Protein Eng. Des. Sel.* **22,** 93–101 (2009).

18. Clark, L. A. *et al*. Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. *Protein Sci.* **15,** 949–960 (2006).

19. Sivasubramanian, A., Sircar, A., Chaudhury, S. & Gray, J. J. Toward high-resolution homology modeling of antibody F. *Proteins* **74,** 497–514 (2009).

20. Pantazes, R. J. & Maranas, C. D. OptCDR : a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. *Protein Eng. Des. Sel.* **23,** 849–858 (2010).

21. Haard, H. J. De *et al*. A Large Non-immunized Human Fab Fragment Phage Library That Permits Rapid Isolation and Kinetic Analysis of High Affinity Antibodies *. *J. Biol. Chem.* **274,** 18218–18230 (1999).

22. Wiiger, M. T., Gehrken, H. B., Fodstad, O., Maelandsmo, G. M. & Andersson, Y. A novel human recombinant single-chain antibody targeting CD166 / ALCAM inhibits cancer cell invasion in vitro and in vivo tumour growth. *Cancer immunol Immunother* **59,** 1665–1674 (2010).

23. Smith, G. P. Filamentous Fusion Phage : Novel Expression Vectors that Display Cloned Antigens on the Virion Surface. *Science (80-.).* **228,** 1315–1317 (1985).

24. Smith, G. P. & Petrenko, V. A. Phage Display. **2665,** (1997).

25. Wern, H. & St, S. Biotechnological applications for surface-engineered bacteria. *biotechno. Allp. Biochem.* **40,** 209–228 (2004).

26. Farajnia, S. *et al*. Development trends for generation of single-chain antibody fragments. *Immunopharmacol. Immunotoxicol.* **36,** 297–308 (2014).

27. Hanes, J. & Pluckthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci.* **94,** 4937–4942 (1997).

28. Vaughan, T. J. *et al*. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* **14,** 309–314 (1996).

29. Griffiths, A. D. *et al*. Isolation of high affinity human antibodies directly from

large synthetic repertoires. *Embo J.* **13,** 3245–3260 (1994).

30. Ponsel, D., Neugebauer, J., Ladetzki-Baehs, K. & Tissot, K. High Affinity, Developability and Functional Size: The Holy Grail of Combinatorial Antibody Library Generation. *molecules* **16,** 3675–3700 (2011).
31. Lieberman, R. L., Culver, J. a, Entzminger, K. C., Pai, J. C. & Maynard, J. a. Crystallization chaperone strategies for membrane proteins. *Methods* **55,** 293–302 (2011).
32. Xu, L. *et al*. Directed Evolution of High-Affinity Antibody Mimics Using mRNA Display. *Chem. Biol.* **9,** 933–942 (2002).
33. Mattheakis, L. C., Bhattt, R. R. & Dower, W. J. An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl. Acad. Sci.* **91,** 9022–9026 (1994).
34. Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display : selecting and evolving proteins in vitro that specifically bind to a target. **4,** 269–279 (2007).
35. Nagumo, Y., Fujiwara, K., Horisawa, K., Yanagawa, H. & Doi, N. PURE mRNA display for in vitro selection of single-chain antibodies. *J. Biochem.* **159,** 519–526 (2016).
36. Harvey, B. R. *et al*. Anchored periplasmic expression , a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. *Proc. Natl. Acad. Sci.* **101,** 9193–9198 (2004).
37. Georgiou, G. *et al*. Display of hetetologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotechnol.* **15,** 29–34 (1997).
38. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypepride libraries. *Nat. Biotechnol.* **15,** 553–557 (1997).
39. Daugherty, P. S., Chen, G., Olsen, M. J., Iverson, B. L. & Georgiou, G. Antibody affinity maturation using bacterial surface display. *Protein Eng.* **11,** 825–832 (1998).
40. Huang, R., Fang, P. & Kay, B. K. Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries. *Methods* **58,** 10–7 (2012).
41. Pantazes, R. J. & Maranas, C. D. OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. *Protein Eng. Des. Sel.* **23,** 849–858 (2010).
42. Bertheleme, N. *et al*. Unlocking the secrets of the gatekeeper: methods for stabilizing and crystallizing GPCRs. *Biochim. Biophys. Acta* **1828,** 2583–91 (2013).
43. Rosenbaum, D. M. *et al*. GPCR engineering yields high-resolution structural

insights into beta2-adrenergic receptor function. *Science (80-.).* **318,** 1266–1237 (2007).

44. Uysal, S. *et al*. Crystal structure of full-length KcsA in its. *Proc. Natl. Acad. Sci.* **106,** 6644–6649 (2009).

45. Mujić-Delić, A., de Wit, R. H., Verkaar, F. & Smit, M. J. GPCR-targeting nanobodies: attractive research tools, diagnostics, and therapeutics. *Trends Pharmacol. Sci.* **35,** 247–255 (2014).

46. Jaakola, V.-P. & Ijzerman, A. P. The crystallographic structure of the human adenosine A2A receptor in a high-affinity antagonist-bound state: implications for GPCR drug screening and design. *Curr. Opin. Struct. Biol.* **20,** 401–14 (2010).

47. Koldobskaya, Y. *et al*. A portable RNA sequence whose recognition by a synthetic antibody facilitates structural determination. *Nat. Struct. Mol. Biol.* **18,** 100–106 (2011).

48. Piccirilli, J. a & Koldobskaya, Y. Crystal structure of an RNA polymerase ribozyme in complex with an antibody fragment. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366,** 2918–28 (2011).

49. Kalyoncu, S. *et al*. Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments. *Proteins Struct. Funct. Bioinforma.* **82,** 1884–1895 (2014).

50. Hino, T., Iwata, S. & Murata, T. Generation of functional antibodies for mammalian membrane protein crystallography. *Curr. Opin. Struct. Biol.* **23,** 563–568 (2013).

51. Johnson, J. L. *et al*. Structural and biophysical characterization of an epitope-specific engineered Fab fragment and complexation with membrane proteins: implications for co-crystallization. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 896–906 (2015).

52. O'Malley, M. a *et al*. Progress toward heterologous expression of active G-protein-coupled receptors in Saccharomyces cerevisiae: Linking cellular stress response with translocation and trafficking. *Protein Sci.* **18,** 2356–70 (2009).

53. Enteropathogenic, Y. *et al*. Crystal Structures of the Outer Membrane Domain of Intimin and Invasin from Enterohemorrhagic. *Struct. Des.* **20,** 1233–1243 (2012).

54. Pai, J. C., Entzminger, K. C. & Maynard, J. A. Restriction enzyme-free construction of random gene mutagenesis libraries in Escherichia coli. *Anal. Biochem.* **421,** 640–648 (2012).

55. Röthlisberger, D., Honegger, A. & Plückthun, A. Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J. Mol. Biol.* **347,** 773–89 (2005).

56. Levy, R., Weiss, R., Chen, G., Iverson, B. L. & Georgiou, G. Production of Correctly Folded Fab Antibody Fragment in the Cytoplasm of Escherichia coli trxB gor Mutants via the Coexpression of Molecular Chaperones. *Protein Expr. Purif.* **23,** 338–347 (2001).

57. Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357,** 289–298 (2006).

58. Krebs, B. *et al.* High-throughput generation and engineering of recombinant human antibodies. *J. Immunol. Methods* **254,** 67–84 (2001).

59. Muyldermans, S., Cambillau, C. & Wyns, L. Recognition of antigens by single-domain antibody fragments: The superfluous luxury of paired domains. *Trends Biochem. Sci.* **26,** 230–235 (2001).

60. Holt, L. J., Herring, C., Jespers, L. S., Woolven, B. P. & Tomlinson, I. M. Domain antibodies : proteins for therapy. *Trends Biotechnol.* **21,** 484–490 (2003).

61. Conrath, K. E., Lauwereys, M., Wyns, L. & Muyldermans, S. Camel Single-domain Antibodies as Modular Building Units in Bispecific and Bivalent Antibody Constructs *. *J. Biol. Chem.* **276,** 7346–7350 (2001).

62. Conrath, K. *et al.* Efficient Targeting of Conserved Cryptic Epitopes of Infectious. *J. Biol. Chem.* **279,** 1256–1261 (2004).

63. Aetselier, P. D. E. B. & Evets, H. R. EFFICIENT TUMOR TARGETING BY SINGLE-DOMAIN ANTIBODY FRAGMENTS OF CAMELS. *Int. J. cancer* **462,** 456–462 (2002).

64. Ewert, S., Cambillau, C., Conrath, K. & Plu, A. Biophysical Properties of Camelid V HH Domains Compared to Those of Human. *Biochemistry* 3628–3636 (2002).

65. Barthelemy, P. a. *et al.* Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. *J. Biol. Chem.* **283,** 3639–3654 (2008).

66. Ma, X., Barthelemy, P. a, Rouge, L., Wiesmann, C. & Sidhu, S. S. Design of synthetic autonomous VH domain libraries and structural analysis of a VH domain bound to vascular endothelial growth factor. *J. Mol. Biol.* **425,** 2247–59 (2013).

67. Barthelemy, P. a *et al.* Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. *J. Biol. Chem.* **283,** 3639–54 (2008).

68. Tanha, J. *et al.* Improving solubility and refolding efficiency of human V H s by a novel mutational approach. **19,** 503–509 (2006).

69. Miller-Gallacher, J. L. *et al.* The 2.1 Å Resolution Structure of Cyanopindolol-Bound β1-Adrenoceptor Identifies an Intramembrane Na+ Ion that Stabilises the Ligand-Free Receptor. *PLoS One* **9,** e92727 (2014).

70. Rasmussen, S. G. F. *et al*. Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature* **469,** 175–80 (2011).

71. Abskharon, R. N. N. *et al*. Crystallization and preliminary X-ray diffraction analysis of a specific VHH domain against mouse prion protein. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66,** 1644–1646 (2010).

72. Loris, R. *et al*. Crystal Structure of the Intrinsically Flexible Addiction Antidote MazE *. **278,** 28252–28257 (2003).

73. Abskharon, R. N. N. *et al*. Probing the N - Terminal β - Sheet Conversion in the Crystal Structure of the Human Prion Protein Bound to a Nanobody. (2014).

74. Lam, A. Y., Pardon, E., Korotkov, K. V, Hol, W. G. J. & Steyaert, J. Nanobody-aided structure determination of the EpsI:EpsJ pseudopilin heterodimer from Vibrio vulnificus. *J. Struct. Biol.* **166,** 8–15 (2009).

75. Domanska, K., Vanderhaegen, S., Srinivasan, V., Pardon, E. & Dupeux, F. Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic β 2-microglobulin variant. *Proc. Natl. Acad. Sci.* **108,** 1314–1319 (2011).

76. Ward, A. B. *et al*. Structures of P-glycoprotein reveal its conformational fl exibility and an epitope on the nucleotide-binding domain. *Proc. Natl. Acad. Sci.* **110,** 13386–13391 (2013).

77. Korotkov, K. V, Pardon, E., Steyaert, J. & Hol, W. G. J. Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure* **17,** 255–265 (2009).

78. Wang, X. & Maynard, J. a. The Bordetella adenylate cyclase repeat-in-toxin (RTX) domain is immunodominant and elicits neutralizing antibodies. *J. Biol. Chem.* **290,** 3576–91 (2015).

79. Bond, C. J., Marsters, J. C. & Sidhu, S. S. Contributions of CDR3 to VHH Domain Stability and the Design of Monobody Scaffolds for Naive Antibody Libraries. *J. Mol. Biol.* **332,** 643–655 (2003).

80. Sidhu, S. S., Lowman, H. B., Cunnningham, B. C. & Wells, J. A. Phage Display for Selection of Novel Binding Peptides. **328,** 333–363 (2000).

81. Fellouse, F. a, Barthelemy, P. a, Kelley, R. F. & Sidhu, S. S. Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. *J. Mol. Biol.* **357,** 100–14 (2006).

82. Yan, J., Li, G., Hu, Y., Ou, W. & Wan, Y. Construction of a synthetic phage-displayed Nanobody library with CDR3 regions randomized by trinucleotide cassettes for diagnostic applications. *J. Transl. Med.* **12,** 343 (2014).

83. Pardon, E. *et al*. A general protocol for the generation of Nanobodies for structural biology. *Nat. Protoc.* **9,** 674–93 (2014).

84. Perchiacca, J. M., Lee, C. C. & Tessier, P. M. Optimal charged mutations in the

complementarity-determining regions that prevent domain antibody aggregation are dependent on the antibody scaffold. *Protein Eng. Des. Sel.* **27,** 29–39 (2014).

85. Lee, C. C. *et al.* Toward aggregation-resistant antibodies by design. *Trends Biotechnol.* **31,** 612–620 (2013).

86. Lee, C. C. *et al.* Design and Optimization of Anti-amyloid Domain Antibodies Specific for NL -Amyloid and Islet Amyloid Polypeptide *. **291,** 2858–2873 (2016).

87. Pandit, R. A. *et al.* Isolated CyaA-RTX subdomain from Bordetella pertussis: Structural and functional implications for its interaction with target erythrocyte membranes. *Biochem. Biophys. Res. Commun.* 6–11 (2015). doi:10.1016/j.bbrc.2015.08.110

88. Shukla, A. K., Gupta, C., Srivastava, A. & Jaiman, D. *Antibody Fragments for Stabilization and Crystallization of G Protein-Coupled Receptors and Their Signaling Complexes. Membrane proteins Engineering, Purification and Crystallization* **557,** (Elsevier Inc., 2015).

89. Ecker, D. M. *et al.* The therapeutic monoclonal antibody market The therapeutic monoclonal antibody market. *MAbs* **7,** 9–14 (2015).

90. Vanhee, P. *et al.* Computational design of peptide ligands. *Trends Biotechnol.* **29,** 231–239 (2011).

91. Fjell, C. D., Hiss, J. A., Hancock, R. E. W. & Schneider, G. Designing antimicrobial peptides: form follows function. *Nat. Rev. Drug Discov.* **11,** 37–51 (2012).

92. Azoitei, M. L. *et al.* Computational design of protein antigens that interact with the CDR H3 loop of HIV broadly neutralizing antibody 2F5. *Proteins* **82,** 2770–2782 (2014).

93. Fleishman, S. J. *et al.* of Influenza Hemagglutinin. *Science* **332,** 816–821 (2011).

94. Kwong, P. D., Mascola, J. R. & Nabel, G. J. Broadly neutralizaing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat. Rev. Immunol.* **13,** 693–701 (2013).

95. Chothia, C. & Tt, A. M. L. Canonical Structures for the Hypervariable Regions of Immunoglobulins. *J. Mol. Biol.* **196,** 901–917 (1987).

96. Whitelegg, N. R. J. & Rees, A. R. WAM: an improved algorithm for modelling antibodies on the WEB. *Protein Eng.* **13,** 819–824 (2000).

97. Miklos, A. E. *et al.* Brief Communication Structure-Based Design of Supercharged , Highly Thermoresistant Antibodies. *Chem. Biol.* **19,** 449–455 (2012).

98. Kuroda, D., Shirai, H., Jacobson, M. P. & Nakamura, H. Computer-aided antibody design. *Protein Eng. Des. Sel.* **25,** 507–521 (2012).

99. Marvin, J. S. & Lowman, H. B. Redesigning an antibody fragment for faster association with its antigen. *Biochemistry* **42,** 7077–7083 (2003).

100. Chennamsetty, N., Voynov, V., Kayser, V., Helk, B. & Trout, B. L. Design of therapeutic proteins with enhanced stability. *Proc. Natl. Acad. Sci.* **106,** 11937–11942 (2009).

101. Sormanni, P., Aprile, F. A. & Vendruscolo, M. Rational design of antibodies targeting specific epitopes within intrinsically disordered proteins. *Proc. Natl. Acad. Sci.* **112,** 9902–9907 (2015).

102. Hopp, T. P. *et al.* A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* **6,** 1204–1210 (1988).

103. Roosild, T. P., Castronovo, S. & Choe, S. Structure of anti-FLAG M2 Fab domain and its use in the stabilization of engineered membrane proteins. *Acta Crystallogr.* **F62,** 835–839 (2006).

104. Leysath, C. E. *et al.* Crystal Structure of the Engineered Neutralizing Antibody M18 Complexed to Domain 4 of the Anthrax Protective Antigen. *J. Mol. Biol.* **387,** 680–693 (2009).

105. Brenke, R. *et al.* Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics* **28,** 2608–2614 (2012).

106. Pantazes, R. J., Grisewood, M. J., Li, T., Gifford, N. P. & Maranas, C. D. The Iterative Protein Redesign and Optimization (IPRO) suite of programs. *J. Comput. Chem.* **36,** 251–263 (2015).

107. Dunbrack Jr, R. L. Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.* **12,** 431–440 (2002).

108. Fazelinia, H., Cirino, P. C. & Maranas, C. D. Extending Iterative Protein Redesign and Optimization (IPRO) in protein library design for ligand specificity. *Biophys. J.* **92,** 2120–2130 (2007).

109. Reddy, S. T. *et al.* Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat. Biotechnol.* **28,** 965–969 (2010).

110. Krebber, A. *et al.* Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods* **201,** 35–55 (1997).

111. Hayhurst, A. *et al.* Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. *J. Immunol. Methods* **276,** 185–196 (2003).

112. Maynard, J. A. *et al.* Protection against anthrax toxin by recombinant antibody fragments correlates with antigen affinity. *Nature* **20,** 597–601 (2002).

113. Cobaugh, C. W., Almagro, J. C., Pogson, M., Iverson, B. & Georgiou, G. Synthetic antibody libraries focused towards peptide ligands. *J. Mol. Biol.* **378,**

622–33 (2008).

114. Honegger, A. Engineering antibodies for stability and efficient folding. *Handb. Exp. Pharmacol.* **181,** 47–68 (2008).

115. Ladiwala, A. R. a *et al.* Rational design of potent domain antibody inhibitors of amyloid fibril assembly. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 19965–70 (2012).

116. Pantazes, R. J., Grisewood, M. J. & Maranas, C. D. Recent advances in computational protein design. *Curr. Opin. Struct. Biol.* **21,** 467–472 (2011).

117. Fleishman, S. J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science (80-.).* **332,** 816–821 (2011).

118. Tinberg, C. E. *et al.* Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501,** 212–6 (2013).

119. Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* **321,** 522–525 (1986).

120. Marcatili, P., Rosi, A. & Tramontano, A. PIGS: automatic prediction of antibody structures. *Bioinformatics* **24,** 1953–1954 (2008).

121. Sircar, A., Kim, E. T. & Gray, J. J. RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Res.* **37,** W474–W479 (2009).

122. Teplyakov, A. *et al.* Antibody modeling assessment II. Structures and models. *Proteins* **82,** 1563–1582 (2014).

123. Pantazes, R. J. & Maranas, C. D. MAPs: a database of modular antibody parts for predicting tertiary structures and designing affinity matured antibodies. *BMC Bioinformatics* **14,** 168 (2013).

124. Li, T., Pantazes, R. J. & Maranas, C. D. OptMAVEn – A New Framework for the de novo Design of Antibody Variable Region Models Targeting Specific Antigen Epitopes. *PLoS One* **9,** e105954 (2014).

125. Olsnes, S. The history of ricin , abrin and related toxins. *Toxicon* **44,** 361–370 (2004).

126. Gill, D. M. Bacterial Toxins : a Table of Lethal Amounts. *Microbiol. Rev.* **46,** 86–94 (1982).

127. Kimura, M., Sumizawa, T. & Funatsu, G. The Complete Amino Acid Sequences of the B- Chains of Abrin-a and Abrin-b , Toxic Proteins from the Seeds of Abrus precatorius The Complete Amino Acid Sequences of the B-Chains of Abrin-a and Abrin-b , Toxic Proteins from the Seeds of Abrus precatorius. *Biosci. Biotechnol. Biochem.* **57,** 166–169 (1993).

128. Prigent, J. *et al.* Neutralising Antibodies against Ricin Toxin. *PLoS One* **6,** e20166 (2011).

129. Pelat, T. *et al*. Isolation of a human-like antibody fragment (scFv) that neutralizes ricin biological activity. *BMC Biotechnol*. **9,** 60–73 (2009).

130. Maddaloni, M. *et al*. Immunological Characteristics Associated with the Protective Efficacy of Antibodies to Ricin. *J. Immunol*. **172,** 6211–6228 (2004).

131. Smallshaw, J. E., Richardson, J. A. & Vitetta, E. S. RiVax , a recombinant ricin subunit vaccine , protects mice against ricin delivered by gavage or aerosol. *Vaccine* **25,** 7459–7469 (2007).

132. Zhang, T. *et al*. Truncated abrin A chain expressed in Escherichia coli : A promising vaccine candidate Truncated abrin A chain expressed in Escherichia coli : A promising vaccine candidate. *Hum. Vaccine Immunother*. **10,** 2648–2655 (2014).

133. Goldman, E. R. *et al*. Llama-Derived Single Domain Antibodies Specific for Abrus Agglutinin. *Toxins (Basel)*. 1405–1419 (2011). doi:10.3390/toxins3111405

134. Wang, B. *et al*. Discovery of high affinity anti-ricin antibodies by B cell receptor sequencing and by yeast display of combinatorial V H : V L libraries from immunized animals. *MAbs* **0,** 1–10 (2016).

135. Zhao, Z., Worthylake, D., Lecour, L., Maresh, G. A. & Pincus, S. H. Crystal Structure and Computational Modeling of the Fab Fragment from a Protective Anti-Ricin Monoclonal Antibody. *PLoS One* **7,** 1–10 (2012).

136. Vance, D. J., Tremblay, J. M., Mantis, N. J. & Shoemaker, C. B. Stepwise Engineering of Heterodimeric Single Domain Camelid V H H Antibodies That Passively Protect Mice from Ricin Toxin *. *J. Biol. Chem*. **288,** 36538–36547 (2013).

137. Cohen, O. *et al*. Characterization and Epitope Mapping of the Polyclonal Antibody Repertoire Elicited by Ricin Holotoxin-Based Vaccination. *Clin. Vaccine Immunol*. **21,** 1534–1540 (2014).

138. Primates, N., Kronman, C., Ordentlich, A. & Mazor, O. Isolation of Anti-Ricin Protective Antibodies Exhibiting High Affinity from Immunized. *Toxins (Basel)*. **8,** 64–79 (2016).

139. Wo, A., Plu, A. & Institut, B. Stability Engineering of Antibody Single-chain Fv Fragments. *J. Mol. Biol*. **305,** 989–1010 (2001).

140. Chang, H. & Sun, Z. Increasing stability of antibody via antibody engineering: Stability engineering on an anti-hVEGF. *Proteins* **82,** 2620–2630 (2014).

141. Tiller, T. *et al*. A fully synthetic human Fab antibody library based on fixed VH / VL framework pairings with favorable biophysical properties A fully synthetic human Fab antibody library based on fixed VH / VL framework pairings with favorable biophysical properties. *MAbs* **5,** 445–470 (2013).

142. Jayaram, N. & Bhowmick, P. Germline V H / V L pairing in antibodies. *Protein*

Eng. Des. Sel. **25,** 523–529 (2012).

143. Entzminger, K. C., Johnson, J. L., Hyun, J., Lieberman, R. L. & Maynard, J. A. Increased Fab thermoresistance via V H -targeted directed evolution. *Protein Eng. Des. Sel.* **28,** 365–377 (2015).

Vita

Jeongmin Hyun was born in Seoul, Korea. After graduating Myungduk Foreign Language high school, she attended Yonsei University where she received her Bachelor of Science in Biotechnology in 2007. Her interest in biology led to her to continue a Masters degree in the Life Science and Biotechnology Department with the guidance of Dr. Baik-Lin Seong. Upon graduation, she started her doctoral study in 2010 in the Cell and Molecular Biology Department at the University of Texas at Austin.

Email address: jmhyun1111@gmail.com

This dissertation was typed by Jeongmin Hyun