The Dissertation Committee for Christian Wessel Cobaugh certifies that this is the approved version of the following dissertation:

SINGLE SCAFFOLD ANTIBODY LIBRARIES CREATED WITH HIGH RATES OF MUTAGENESIS OR DIVERSITY FOCUSED FOR PEPTIDE RECOGNITION

Committee:

_____

Brent L. Iverson, Supervisor

_____

George Georgiou, Co-Supervisor

_____

Charles F. Earhart

_____

George B. Kitto

_____

Jon D. Robertus

SINGLE SCAFFOLD ANTIBODY LIBRARIES CREATED WITH HIGH RATES OF MUTAGENESIS OR DIVERSITY FOCUSED FOR PEPTIDE RECOGNITION

by

Christian Wessel Cobaugh, B.A., 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

December 2007

Dedication

I dedicate this work to my wife Amy. Without her boundless love and patience, her constant support and interest, none of this would have been possible. And to my daughter, Celeste: you are the motivating force to "git 'er done".

Acknowledgments

I would like to thank Katherine Reynolds, who made my transition from civilian to student life so simple, and Dr. Klaus Linse for helping me with peptide synthesis, purification, and providing constant encouragement. I also want to give my heartfelt appreciation to the invaluable postdoctoral fellows of the BIGG lab, including Drs. Yariv Mazor and Ki Jun Jeong, who always seemed to find the time to answer my questions. I give thanks to Brent Iverson and George Georgiou for believing in my projects even when I did not.

Single Scaffold Antibody Libraries Created with High Rates of Mutagenesis or Diversity Focused for Peptide Recognition

Publication No._____

Christian Wessel Cobaugh, Ph.D.

The University of Texas at Austin, 2007

Supervisors: Brent L. Iverson

and George Georgiou

This dissertation describes several strategies used to create diversity in non-immune antibody libraries. Two of the strategies were used to create two separate peptide focused libraries. Both of these strategies used to create these antigen-class focused libraries used a single scaffold antibody gene that contained diversity only in the variable heavy region. The scaffold antibody gene one of the libraries, the M:anti-pep library, was chosen based on hypervariable loop canonical structures that are characteristic of other anti-peptide antibodies. Additionally, all of the contact residues of this antibody are commonly used contact residues in other anti-peptide antibodies. These positions and others were varied to incorporate the natural diversity of other anti-peptide antibodies. The second library, the Hu:anti-pep, is based on a widely used, unique combination of human germline

antibody segments that express well in bacterial expression. Positions were chosen for variation based on their usage as contact residues in both anti-peptide and anti-protein antibodies. The diversity was less focused than with the M:anti-pep library, incorporating all 20 amino acids at "high usage" positions and only four amino acids at "low usage" positions. Both libraries were validated by phage display selections against the peptide angiotensin (AT) and neuropeptide Y (NPY). The M:anti-pep library yielded specific antibodies to both peptides with dissociation constants as low as 14 nM against AT and 18 nM against NPY. The Hu:anti-pep library yielded specific clones with higher dissociation constants: 49 nM against NPY and 11 μM against AT.

The final strategy used to introduce diversity is widely used for affinity maturation of low affinity, previously selected antibodies. Extremely high rates of mutagenesis (2.2% of the gene to 2.7%) were used to create two libraries of the anti-digoxin antibody 26-10. The libraries had been screened by others in an attempt to examine the effects of high-rates of mutagenesis on the directed evolution of an antibody. A total of 91 isolated clones from both libraries were sequenced. Several consensus mutations were identified near the CDRH3 in the isolated clones, indicating that they had a positive, selectable effect. This study confirmed that high-error rate antibody libraries contain more active clones than expected. Combinations of the selected consensus mutations from these libraries provide moderate enhancements to the kinetics and expression of the wild-type antibody in a non-synergistic manner.

Table of Contents

LIST OF FIGURES ..... **IX**

LIST OF TABLES ..... **X**

CHAPTER 1: ANTIBODIES AND ANTIBODY ENGINEERING **1**

INTRODUCTION 1
THE ADAPTIVE IMMUNE SYSTEM: A MODEL THERAPEUTIC SYSTEM 2
ANTIBODY STRUCTURE 4
CANONICAL STRUCTURE PARADIGM 6
HYBRIDOMA TECHNOLOGY AND MONOCLONAL ANTIBODIES 9
RECOMBINANT ANTIBODIES 10
HUMANIZATION 11
ANTIBODY ENGINEERING 12
IMMUNE REPERTOIRE LIBRARIES 13
SYNTHETIC LIBRARIES 15
DIRECTED EVOLUTION 16
Diversity, sequence space, and library size *18*
Affinity maturation *18*
ANTIBODY LIBRARY DISPLAY AND SELECTION METHODS 19
Phage Display *19*
Ribosome display *24*
Microbial cell display and flow cytometry *25*
PEPTIDES 27
Peptides as alternatives to whole protein antigens *29*
Anti-peptide antibodies in detection and therapies *30*
DISSERTATION SUMMARY 31
REFERENCES 32

CHAPTER 2: A SYNTHETIC ANTIBODY LIBRARY DESIGN WITH A BIAS TOWARDS PEPTIDE BINDERS **41**

INTRODUCTION 41
RESULTS 47
Design and construction of combinatorial M:anti-pep library *47*
Design and construction of combinatorial Hu:anti-pep library *53*
Selections of scFvs with 2 different peptide targets *55*
Sequence analysis *58*

Characterization of soluble scFvs *59*
CONCLUSIONS 61
EXPERIMENTAL PROTOCOLS 68
Library Construction *68*
Phage panning *71*
Phage ELISAs *72*
Expression of soluble scFvs *73*
BIACore Analysis *74*
REFERENCES 75

CHAPTER 3: ANALYSIS OF CONSENSUS MUTATIONS IN HIGH ERROR RATE LIBRARIES OF SCFV ANTIBODIES 80

INTRODUCTION 80
RESULTS 83
Library construction and analysis *83*
Analysis of clones exhibiting increased cell fluorescence *86*
Construction and characterization of single amino acid variants of the 26-10 scFv antibody *88*
DISCUSSION 95
EXPERIMENTAL PROTOCOLS 103
Strains, Plasmids, and Reagents *103*
Creating the common mutants *104*
scAb protein expression and purification *104*
Expression levels of functional scFv *105*
Protein stability analysis *106*
Surface plasmon resonance (SPR) kinetic analysis of single chain antibodies *106*
REFERENCES 107

CHAPTER 4: CONCLUSION AND FUTURE DIRECTIONS 111

CONCLUSION 111
FUTURE DIRECTIONS 113

BIBLIOGRAPHY 115

VITA…… 126

List of Figures

FIGURE 1-1. ANTIBODY STRUCTURE: IGG TO SINGLE CHAIN FRAGMENT ANTIBODY. 6
FIGURE 1-2: CANONICAL LOOP STRUCTURES. 8
FIGURE 1-3. FILAMENTOUS PHAGE STRUCTURE. 21
FIGURE 1-4. FILAMENTOUS PHAGE DISPLAY PANNING OF SCFV LIBRARY. 23
FIGURE 1-5: WORKFLOW OF FLOW CYTOMETRY-BASED SELECTION OF ANTIBODIES USING MICROBIAL DISPLAY. 27
FIGURE 2-1: SCFV STRUCTURE WITH THE CDR LOOPS 42
FIGURE 2-2: GROSS TOPOGRAPHICAL FEATURES OF THE THREE MAJOR ANTIGEN CLASSES OF ANTIBODIES. 43
FIGURE 2-3. CANONICAL STRUCTURE PATTERNS CORRESPOND TO GROSS SPECIFICITIES45
FIGURE 2-4. SDRS OF 26-10. 49
FIGURE 2-5. STRUCTURE OF THE HU:ANTI-PEP AND M:ANTI-PEP VH LIBRARIES. 54
FIGURE 2-6. SEQUENCE OF L1 IN A27MDJκ1 55
FIGURE 2-7. POLYCLONAL PHAGE ELISAS. 56
FIGURE 2-8. MONOCLONAL PHAGE ELISA SCREENING. 57
FIGURE 3-1. PRESORT ANALYSIS OF BOTH HIGH ERROR-RATE LIBRARIES. 84
FIGURE 3-2. EXPRESSION LEVELS OF 26-10 AND MUTANTS. 93
FIGURE 3-3. FUNCTIONAL STABILITY OF 26-10 AND MUTANTS. 94
FIGURE 3-4. STRUCTURAL MODEL OF 26-10 BINDING POCKET. 98

List of Tables

TABLE 2-1. HV LOOP LENGTH OF 46 ANTI-PEPTIDE ANTIBODIES AND BOTH LIBRARY SCAFFOLDS 48
TABLE 2-2. SDRS OF 24 ANTI-PEPTIDE COMPARED TO SDRS OF 26-10 49
TABLE 2-3. M:ANTI-PEP LIBRARY TARGET DIVERSITY OF H1, H2. 51
TABLE 2-4: H3 DIVERSITY. 52
TABLE 2-5. ALIGNMENT OF CLONES ISOLATED FROM THE HU:ANTI-PEP LIBRARY: 58
TABLE 2-6. ALIGNMENT OF CLONES ISOLATED FROM THE M:ANTI-PEP LIBRARY 59
TABLE 2-7. BIACORE ANALYSIS AND EXPRESSION YIELDS OF SCFV ANTIBODY FRAGMENTS SPECIFIC TO NPY OR AT. 61
TABLE 2-8. OLIGONUCLEOTIDE PRIMERS USED FOR THE CONSTRUCTION OF THE M:ANTI-PEP LIBRARY. 70
TABLE 3-1. TYPES OF MUTATIONS. 85
TABLE 3-2. LIBRARY ACTIVITY 86
TABLE 3-3. COMMONLY OCCURRING MUTATIONS. 88
TABLE 3-4. VARIANT CLONES CREATED BY SITE-DIRECTED MUTAGENESIS. 89
TABLE 3-5. KINETICS OF 26-10 AND MUTANTS DETERMINED WITH BIACORE. 91

Chapter 1: Antibodies and Antibody Engineering

INTRODUCTION

The immune systems of vertebrates consist of a complex network of tissues, organs, specialized cells, and proteins. This network is able to evolve, or adapt, over time as the immune system fights off pathogens which themselves are adapting to overcome the immune system[1]. The adaptive mammalian immune system produces billions of different antibody proteins which are a key element of the immune system's defense against a variety of antigens presented by invading pathogens as well as cancer cells[2]. One class of antibodies, known as immunoglobulin G (IgG), is produced by B-cell lymphocytes and secreted into the serum. The role of IgG is to bind with both specificity and high-affinity to the antigen-containing foreign molecules, including proteins, peptides, carbohydrates, and others, that the immune system encounters. The immune system generates IgG repertoire diversity through V(D)J recombination, followed by iterative rounds of mutation of the antibody genes, and antigen-mediated selection for tight binders[1]. The strategies and even specific proteins of the immune system have been leveraged into therapies ever since the use of passive polyclonal serum therapy to treat infection emerged over 100 years ago[3]. Laboratory methods that mimic the highly adaptive nature of the immune system began to emerge in the late 20th century[4]. The field of antibody engineering uses combinatorial methods to recapitulate the immune system's diversity, while methods like phage display allow for the isolation of monoclonal antibodies that are simultaneously highly specific and bind with high-affinity to the antigen of interest.

THE ADAPTIVE IMMUNE SYSTEM: A MODEL THERAPEUTIC SYSTEM

Under normal circumstances, the adaptive immune system is triggered into action upon recognition of foreign antigens on the surface of invading microbes or on the surface of virally infected host cells. The clonally unique cell surface receptors of naïve B-cells that recognize antigens are essentially low affinity, membrane bound antibody molecules. Upon binding to the antigen, the B-cell engulfs the antigen, digests it, and presents peptide fragments of the antigen on its surface through the major histocompatibility complexes (MHCs). The antigen fragment-MHC binds stimulated T-cells, which secrete cytokines that activate the B-cell expressing the recognizing antibody to undergo differentiation into one of several types of specialized B-cells. While the clonal progeny of the original B-cell divide and migrate to the secondary lyphoid organs,, the genes responsible for the variable domains of the recognizing antibody undergo somatic hypermutation to the point where an even higher-affinity B-cell receptor (*i.e.* cell surface antibody) is produced through an affinity maturation process. The clones with the highest-affinity receptors are selected by out-competing the lower-affinity clones for the limited antigen[1]. Upon completion of this process, the surviving clones have differentiated into either plasma B-cells or memory B-cells. The plasma B-cells are short lived, but they secrete large quantities of soluble IgG antibodies that will bind to the antigen with high specificity and affinity.

Antibody coated pathogens are prevented from binding to, and invading, host cells and can be targeted for destruction and/or clearance. The antigen bound antibodies initiate complex effector functions to rid the host of undesired antigens by recruiting phagocytic cells such as macrophages. Once these cells are in close enough proximity to

the antibody-antigen complex, Fc receptors on the phagocytic cells' surfaces bind the Fc region of the antibody and initiate the engulfing and destruction of antibody coated antigens[1]. Alternatively, the binding of the antigen-complexed antibodies to the Fc receptors on natural killer cells (NKs) can activate a process known as antibody-dependent cell-mediated cytotoxicity (ADCC) which causes the NKs to release cytokines, pore-forming proteins, and proteases, all of which attack the antibody coated pathogen[5]. Antibodies bound to pathogens can also stimulate the complement system that opsonizes the pathogens and attracts neutrophils to phagocytose the opsonized pathogens[5]. Memory B-cells survive for much longer periods in the secondary lymphoid organs where they continue to secrete antibodies and are primed to divide into clonal populations of plasma B-cells if the pathogen is encountered in subsequent infections[2].

Polyclonal anti-sera, containing a myriad of different IgG clones of varying isotypes and to different epitopes on the same antigen, were once the only option for therapy and research. Polyclonal anti-sera, which are purified from the plasma of humans or animals whose immune systems have responded to the antigen of interest, are easier to make than monoclonal antibodies (MAbs) as long as immune donors are available[6]. However, due to the heterogeneity of the proteins in the serum, the specific activity is low because so few of the antibodies in the preparation are specific to the antigen. For example, a 0.7 mg of a mixture of two MAbs had the same neutralizing activity as 100 mg–170 mg of tetanus immune globulin[7]. Additionally, because polyclonal anti-sera is purified from donors, it can be expensive to prepare and can transmit infection from donors to recepients[8]. Other drawbacks of polyclonal anti-sera include ethical and safety issues regarding immunizing donors to create hyper-immune plasma and lot-to-lot

heterogeneity[7]. Specific activity and high levels of purity can be rigorously controlled during the manufacture of MAbs, explaining why they have become such popular therapeutic agents[6].

ANTIBODY STRUCTURE

IgG antibodies are homodimer glycoproteins made up of four polypeptide chains: two identical heavy chains, each about 250 amino acids long, and two identical light chains, each about 150 amino acids long. A disulfide bond connects a cysteine at the *C*-terminus of each light chain to a cysteine of a heavy chain[9]. In turn, the heavy chains are associated with each other *via* two disulfide bonds in the hinge region, which spans the variable and constant regions of each heavy chain (Figure 1-1). The sequences of different antibodies are highly conserved throughout most of the antibody[10]. The invariable regions are referred to as constant (C) domains. The *N*-terminal regions of each chain, known as the variable (V) domains (V_L for variable domain of the light chain, and V_H for the variable domain of the heavy chain) contains three hypervariable (HV) loops of high sequence and length variation[11]. The six HV loops each contain an antigen binding region, better known as the six *complementarity-determining regions* (CDRs)[12]. The CDRs contain the amino acids of the antibody that recognize and interact with an antigen[12]. The IgG is actually a bivalent molecule with two binding sites, each with six CDRs. The conserved sequences connecting the CDRs, known as framework regions (FRs), make up the β-sheet structural folds of antibodies[13]. Due to the conserved nature of the FR sequences, the β-sheet fold of antibodies is a conserved, easily recognizable

feature[13]. The *C*-terminal domain of the heavy chain makes up the constant fragment (Fc region), which is capable of eliciting effector functions in cells that contain Fcγ receptors.

Proteolysis of IgG molecules yields several stable antibody fragments[14]. Two examples of such fragments, the Fab and Fv are shown in Figure 1-1. Both of these fragments are monovalent because they both contain only one pair of V domains (versus a pair of V domains in whole IgG). In addition to containing the paired V domains, the Fab also contains the first C domain of each chain. The Fv fragment contains only the paired V domains with no C domains[14].

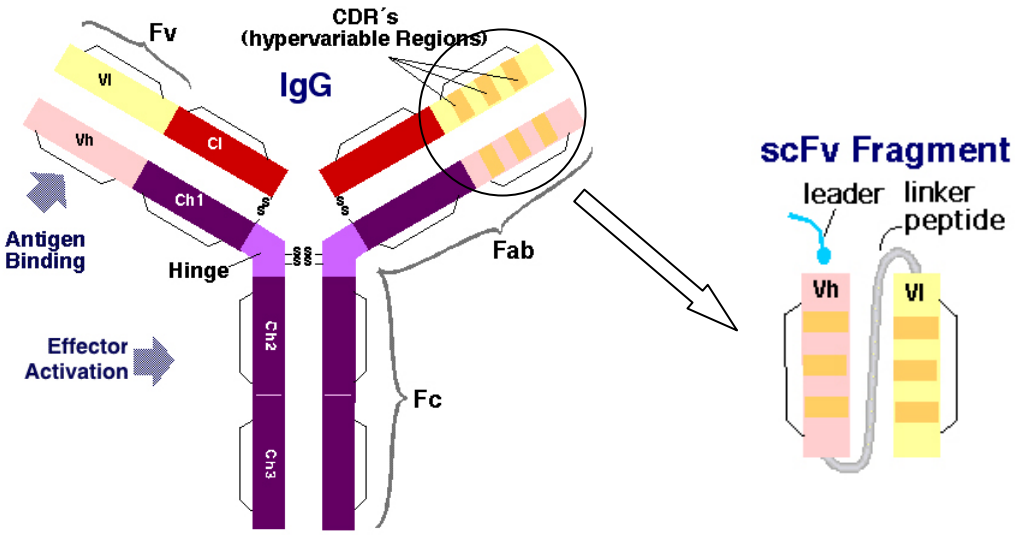


Figure 1-1. Antibody Structure: IgG to single chain fragment antibody.
IgG is a dimer of dimers (Two heavy/light chain dimers). Each chain of the IgG contains one variable domain (labeled Vl for the light chain and Vh for the heavy chain) and constant domains (one constant domain, labeled Cl, for the light chain and three constant domains, labeled Ch 1, 2 and 3 for the heavy chain). The CDRs, pictured above as brown stripes on the variable domains, contain the antigen recognition and binding sites of the antibody. A Fab fragment consists of the V_H and the V_L plus the first constant regions of both chains, while the single chain fragment (scFv) consists of the V_H and the V_L connected by peptide linker. (figure courtesy of Stephen Dubel)

CANONICAL STRUCTURE PARADIGM

The loops of the HV regions are connected by a highly structured and easily recognizable β-sheet framework that defines the immunoglobulin superfamily. The framework regions of both the V_H and V_L have a limited number of main chain conformations evidenced by the conserved nature of the structures of most antibody variable domains[13; 15; 16]. Within the framework regions, several highly conserved residues are responsible for limiting the number of possible variable domain structures[11; 15; 17; 18]. These conserved amino acids were initially observed in alignments of variable domains, providing clues about their role as structural determinants[18]. Their role as

structural determinants was further buttressed by observations that the conserved residues of the variable chains act as anchor points in solved antibody structures[15; 17; 19].

In addition to the conserved framework structure of antibody variable domains, the HV regions, also have a limited number of main chain conformations. From alignments of antibody sequences, Chothia and Lesk identified relatively invariant or *canonical* amino acids in both the FR and the HV regions[15; 19]. They established relationships between canonical residues of the FR and HV regions, lengths of the HV loops, and a limited number of HV chain conformations. The canonical residues, some of which are found outside the HV region and in the β-sheet framework, dictate the conformations of the known structures through packing, hydrogen bonding, or the ability to assume unusual values for torsion angles ϕ, ψ, or ω[15]. These limited and well described main chain conformations of HV regions in loops one and two of the V_H and in loops one, two, and three of the V_L are also known as *canonical structures*[15; 19; 20]. HV loop three of the V_H has eluded the canonical structure paradigm due high variation in its sequence composition and length. Canonical structure assignments (Figure 1-2) for an HV loop can be determined from the primary sequence of the antibody. Several algorithms have been created to predict canonical structure assignment for the HV loops based on two factors: 1) HV loop length and 2) the position and identity of the conserved (a.k.a. canonical) residues[15; 19; 20; 21].

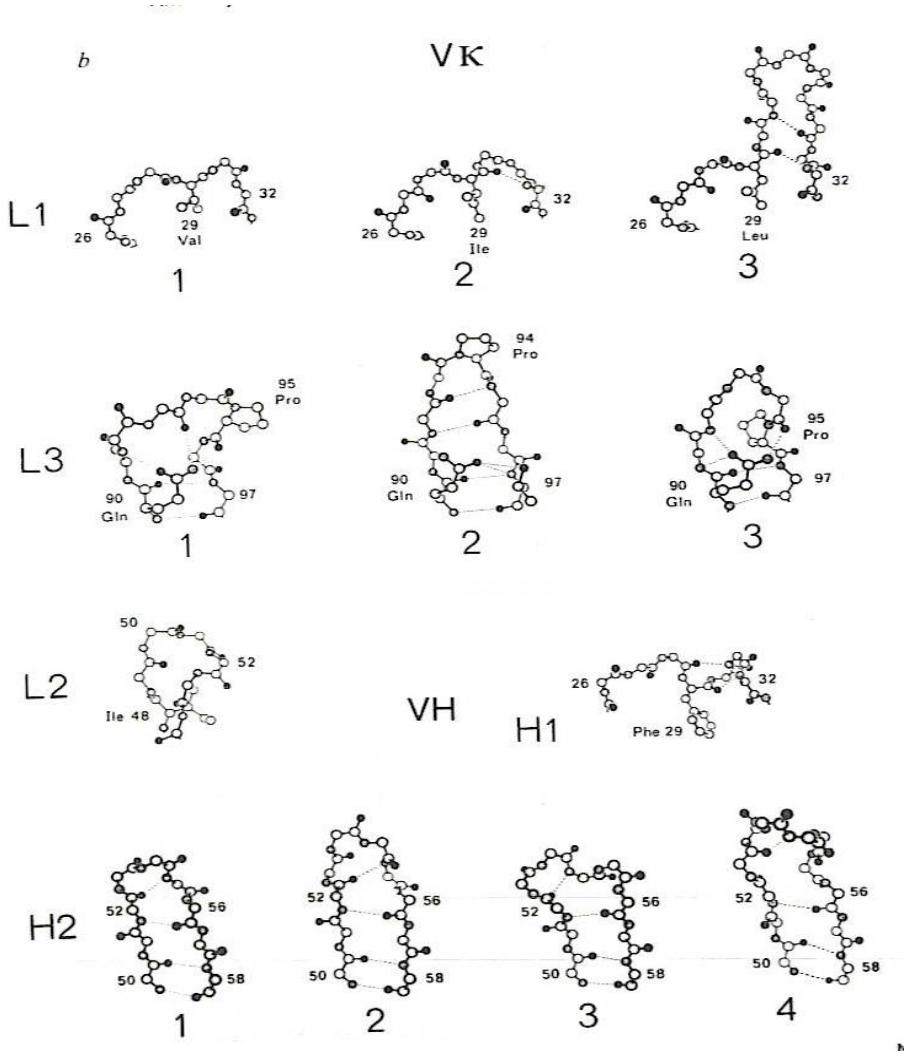


Figure 1-2: Canonical loop structures.
Canonical structures of all three HV loops of $V_L(\kappa)$ domains and HV loops one and two of the V_H domains. HV loop three of the VH domain is too variable in sequence composition and length for any canonical structures to be identified. (figure adapted from Chothia, *et al.*[19])

Each of these HV loop canonical structures is represented by an HV loop reference structure in the protein database. The canonical structure paradigm and the accuracy of the prediction algorithms has been validated by comparing the predicted HV loop structures to the solved atomic structures[19; 22].

HYBRIDOMA TECHNOLOGY AND MONOCLONAL ANTIBODIES

The main sources of MAbs in the immune system, monoclonal plasma B-cells, are very difficult to maintain in a laboratory culture. Previous observations showed that clonal populations of myeloma cells, that is cancerous B-cells, grow indefinitely in culture[23]. Myeloma cell lines that had lost their ability to produce antibodies were fused to antibody producing B-cells to create an immortal chimeric cell line that can survive in culture and provide a virtually unlimited supply of antibodies[24]. After fusion, the myeloma-spleen chimeric cells are isolated with a selection media and diluted into microwells at a concentration of one cell/well to ensure clonality[24]. The secreted antibodies are collected from this microculture and assayed for the ability to bind the target antigen using an enzyme linked immunosorbent assay (ELISA). The positive clones with high sensitivity and specificity are maintained in culture, where they produce secreted MAbs that can be collected and purified. One alternative to producing the MAbs in culture is to inject a hybridoma into the peritoneal cavity of mice. The hybridoma creates a tumor which produces an antibody-rich fluid called ascites that is collected and used as a crude source of MAb.

Although this hybridoma technology is still the backbone of monoclonal antibody isolation and production, antibodies from hybridomas have diminished value as therapeutics because of the unpredictable effects they have on the host's immune system. Since the hybridomas used to produce antibodies are usually derived from the immune system of mice, their direct use as therapies in humans results in an immune response known as the human anti-mouse antibody (HAMA) response[25] whose consequences range from inflammation to anaphylaxis[26]. Even if the HAMA response is tolerated,

subsequent doses of the monoclonal antibody will challenge the immune system and result in rapid antibody clearance thus decreasing efficacy[27]. Additionally, mouse MAbs are limited in their ability to elicit typically therapeutic effector functions such as complement activation through their Fc regions[25]. Nonetheless, MAbs from hybridomas are useful reagents for detection in assays such as Western blots, ELISAs, immunohistochemistry, and for purification by immunoprecipitation or affinity chromatography. Producing hybridomas and screening them is a laborious process.

RECOMBINANT ANTIBODIES

In the 1980s, recombinant DNA technologies enabled the cloning of virtually any gene, including antibody variable genes, into bacteria[28]. The use of recombinant DNA technologies has led to another useful antibody fragment: the single chain fragment antibody (scFv) in which the V_H and the V_L of the Fv, described above, are connected by a flexible peptide linker. The variable domains of antibody genes can be isolated from donor tissue and cloned with the inter-chain linker to create scFv genes using the polymerase chain reaction (PCR) to create scFvs[29; 30] (see Figure 1-1 for scFv structure). Similar PCR-based techniques can be used to isolate the variable domains from hybridoma cells. The scFv antibody fragments can retain the affinity of a respective whole IgG, but are easier to express in bacteria than whole IgG and can be purified to relatively high yields from bacterial cultures (yields of 0.5-5mg purified antibody/L culture are not uncommon). Due to the ease with which genes can be manipulated in bacterial expression systems, antibody fragments are amenable to mutation using molecular biological techniques to enhance the function of the protein (*i.e.* affinity,

stability, etc). It is difficult to perform these types of targeted manipulations on whole IgG in mammalian expression systems or hybridomas.

HUMANIZATION

Ethical and technical considerations diminish the use of humans as immune donors of plasma B-cells to create human hybridomas. Hybridomas from human B-cells are notoriously unstable and not scalable for large-scale production of the antibody[31]. Furthermore, the process known as "tolerance induction", whereby the human immune system is unable to recognize self-antigens, prevents the maturation and isolation of antibodies to important self protein targets such as tumors. To overcome ethical and technical concerns, the self antigen issue, and safety issues, murine MAbs are "humanized" using recombinant DNA technologies to allow for the production of monoclonal antibodies which have non-essential mouse features replaced with human features. "Complementarity determining region (CDR)-replacement" (now, better known as CDR-"grafting") was the first described humanization strategy and involved replacing only the CDRs, which constitute the antigen binding region of an antibody, of a human acceptor antibody with the CDRs of a high-affinity mouse antibody[32]. The sequence of the framework regions of the human acceptor antibody are usually selected based on high homology with the mouse antibody[33]. "Resurfacing" which involves replacement of only the solvent exposed residues of the folded antibody with human sequences has also been used to humanize antibodies derived from murine sources[34]. In nearly all cases of humanization, the resulting antibody must be further engineered to restore the high-affinity of the parent antibody[25].

Humanization of an antibody can affect the stability and affinity of the antibody[25]. One alternative to humanization involves using transgenic mice that contain repertoires of human antibody genes to create hybridomas[35] that produce human antibodies. The germline variable heavy and kappa genes of the transgenic mice are disrupted and human germline variable region segment genes were introduced along with a few human constant regions[36; 37]. The hybridomas that are derived from the B-cells of these transgenic mice produce fully human antibodies due to the human repertoires of the mice[36; 37].

ANTIBODY ENGINEERING

Since the invention of hybridoma technology, the roles of monoclonal antibodies have expanded from biological reagents to therapeutics. Antibody engineering was born out of the idea that recombinant monoclonal antibodies could be isolated, produced, and purified to fill these roles. The field of antibody engineering focuses on using molecular biology methods and recombinant approaches to provide antibodies with defined specificities and affinities for research and therapeutic purposes[38]. Recombinant DNA techniques were used to clone antibody variable genes directly from hybridomas[28]. These cloning milestones were followed by the cloning of fragment antibody libraries (*e.g.* libraries in which the antibodies are expressed as Fabs or scFvs) directly from lymphocytes into bacterial expression systems[39; 40]. In present day antibody engineering, hybridoma technology is often being bypassed as a source of monoclonal antibodies[38]. Sources of antibody diversity were expanded from animal immune repertoires[41] to human immune repertoires[42], and eventually to completely synthetic repertoires[43]. High-

throughput, affinity-based selection methods to isolate recombinant monoclonal antibodies such as phage, yeast-, bacteria-, and ribosome-display technologies have been developed to keep up with the advances in the creation of large combinatorial antibody fragment libraries. Some of the continuing efforts in the field include humanizing mouse antibodies, creating novel antibody formats, generating libraries of antibodies, and screening those libraries.

IMMUNE REPERTOIRE LIBRARIES

The natural immune repertoires of living organisms, including humans, were the original sources of diversity for recombinant antibody libraries[40; 41]. The first V_H and V_L domain repertoires of immunized mice, described in separate papers[40; 41], were cloned out of immune tissues and into vectors that resulted in the libraries being expressed as antibody fragments. Later peripheral blood lymphocytes (PBLs) of human donors were used to create the variable domain diversity of fully human immune libraries[44]. The pools of variable genes are simple to clone because the regions flanking the variable regions gene sequences are conserved and readily available[45]. There are two categories of immune libraries:

1. Naïve immune libraries: the immune system has not been challenged by the target antigen[46]. As a result, the variable genes recovered have not undergone somatic hypermutation and *in vivo* affinity maturation for the target antigen. These libraries are well suited for a broad range of targets. Since the functional diversity relative to each target antigen is low in an un-stimulated immune repertoire, in

order to enhance the chances of isolating specific high-affinity clones the libraries need to be very large ($>10^9$)[38]. A larger library size increases the likelihood that a higher fraction of the immune repertoire's diversity is captured in the library[46]. Nonetheless, isolated antibodies from such libraries are usually low affinity and require additional *in vitro* affinity enhancements.

2. Biased immune libraries: the immune system has been repeatedly challenged by the target antigen[44], developing primary and secondary responses to the antigen. Many of the variable genes are primed for antigen binding after undergoing somatic hypermutation and affinity maturation, often resulting in high affinity binders[40; 41]. These libraries are most often applicable to only the target antigen used to immunize the donor, meaning different target antigens will require the creation of separate libraries. Human versions of biased immune libraries are rare[44] due to ethical considerations of injecting human donors with target antigens for the sake of creating antibodies. Biased human immune libraries have been isolated from transgenic mice that have had their native antibody germline genes replaced with human versions[35] and from the PBLs and spleen cells of diseased human donors[44]. Because the immune system has already selected for high affinity binders to the antigen of interest, the number of relevant sequences in a biased library will be higher than in a naïve immune library[38].

Cloning the variable genes from immune repertoires can introduce unusual V_H/V_L pairs. As a result, many of the displayed library members may not be stable or express well[45]. Antibodies targeting self-antigens, which are important in recognizing some

cancer markers, have been isolated from biased immune libraries using donors afflicted with autoimmune diseases[44]. Such antibodies are generally not accessible from naïve libraries. Naïve immune libraries are the only option of the two types of immune libraries for isolating antibodies to toxic or unstable antigens. Despite the successes of immune repertoire libraries, their shortcomings in the areas of stability, expression, and diversity have led to alternative sources of library diversity.

SYNTHETIC LIBRARIES

Rather than using the entire spectrum of frameworks and diversity contained within the immune system, antibody libraries can be created by introducing diversity using synthetic DNA into the CDRs of highly stable, well-expressed variable domain frameworks[45]. These framework "scaffold" genes are selected to increase the number of functional library clones[47]. Synthetic libraries can be used to isolate antibodies for a broad range of antigens and can be based on human frameworks to avoid HAMA responses[48]. The synthetic DNA used to introduce diversity contains codon variation, or degeneracy, at specific positions resulting in libraries of antibodies that encode an ensemble of amino acids at those positions. However, introducing diversity can result in unusual combinations of amino acids causing misfolded proteins and aggregation[45]. To limit such inappropriate genes in the library, limited variation can be used at the randomized sites. Knowledge of the library scaffold(s)' structure and function can provide details that allow the diversity to be tailored to specific sites in the CDRs and to limit the amino acid possibilities at those sites[47; 48; 49]. Information regarding improvements to antibodies has been determined from structural studies of known

antibody/antigen complexes, computer models, and previous mutagenesis studies[50]. This information can be used to introduce mutations in synthetic libraries to reflect potential improvements[50]. For example, if the parent scaffold's structure is known, CDR positions with side-chains oriented towards the binding pocket are chosen as candidate sites for randomization. The ability to control precisely diversity and choose frameworks with optimal characteristics for particular applications makes synthetic antibody libraries a useful option for antibody engineering.

DIRECTED EVOLUTION

Directed evolution is a protein engineering strategy that attempts to reconstitute Darwinian selection to evolve proteins with novel functions or physical characteristics. Since it is an iterative process, when used in antibody engineering, directed evolution is not unlike rounds of somatic hypermutation/affinity maturation that B-cell receptors undergo *in vivo*. The process can be used to isolate high affinity, well expressed, stable, and specific antibodies from antibody libraries. Diversifying library DNA of the sequences from a parent gene is the first step in directed evolution. Error-prone PCR[51], *in vitro* recombination[52], and site-directed mutagenesis are commonly used to introduce diversity. The diversification step can be applied to the CDRs of scaffold antibodies in the case of synthetic libraries or to a previously isolated low affinity antibody in the case of affinity maturation (more on affinity maturation, below). Selecting improved antibody variants from the entire pool of genes and/or eliminating non-functional variants from the library is the second step (selection techniques described, below). After selection, the isolated variant genes are amplified to provide a large amount of genetic material for

sequence analysis and/or for subsequent rounds of selection. Multiple rounds of selection are usually required for significant enhancements to a protein's function. In some cases, variation is re-introduced to the library genes between rounds of selection. Once enrichment, a term that refers to the overall improvement of the library, has reached the desired level, the remaining variant clones are screened on a smaller scale that permits the desired characteristics, *i.e.* improved affinity in the case of antibodies, to be carefully analyzed and compared. The screening stage usually provides a set of variants small enough to be analyzed in more painstaking detail using biochemical assays that require purified protein.

During the selection steps, coupling the phenotypes of improved clones with their genotypes is required to carry-over and amplify the "survivors" in subsequent rounds of selection, screening, and clonal isolation. By isolating the genes of improved clones at any point in the selection process, the protein sequences of the clones can be determined by DNA sequencing. In between rounds of selection, the genes of improved proteins can be further diversified using mutagenesis for subsequent rounds of selection and screening. The genes recovered from each round of selection are used to express proteins for subsequent rounds of selection or screening. During the directed evolution of antibodies, proteins which are selected with enhancements such as improved affinity, stability, and expression are coupled to their respective genes. The display technologies used in the selection process of antibodies physically link phenotype and genotype[45].

Diversity, sequence space, and library size

The maximum library size of ~10^{13} available to cell-free antibody libraries created by *in vitro* protein synthesis [53] (*e.g.* ribosome display) is several orders of magnitude larger than the maximum library sizes of ~10^{9}-10^{10} available to cell-based antibody libraries[46] (*e.g.* phage display or microbial cell surface)[54]. Even the highest throughput *in vivo* selection method, namely phage display, is limited to screening ~10^{10} total variants[46]. Selections from large antibody libraries usually results in the isolation of higher-affinity antibodies than from smaller libraries[38; 55]. There are several bottlenecks that limit the size of libraries and the throughput of screening libraries[38] which in turn affect the amount of sequence space that can be screened. In order to deal with these bottlenecks, restricting the library's diversity must be balanced with maximizing the number of functional clones relevant to the target antigen. This challenging balancing act can be crucial to the success of antibody engineering selections.

Affinity maturation

As described above, the natural immune system uses a process known as affinity maturation to create high-affinity antibodies from relatively low-affinity B-cell receptor progenitors. The process is mimicked *in vitro* to enhance MAbs through antibody engineering[56]. During the initial steps of the immune response, individual clones of B-cells expressing rearranged germline variable domains as cell surface receptors that recognize the antigen with low-affinity are selected and stimulated to divide and establish a clonal population. In the immune system, somatic mutations are introduced into the

variable domains to provide diversity to the genes encoding the low-affinity receptors. The entire population of variant B-cell clones is continuously exposed to the antigen. Variant clones whose receptors have enhanced affinity and higher expression levels are provided with selective advantages because higher receptor occupation by the antigen leads to increased cell division and survival. The process of somatic hypermutation and selection continues until only one or a few high-affinity clones dominate the immune repertoire. The immune system's ability to introduce diversity to the variable domains of selected, low-affinity antibodies and select higher-affinity variants has similarities to the *in vitro* methods of affinity maturation of engineered antibodies. *In vitro* DNA sequence diversity is introduced using mutagenic PCR[56], DNA shuffling[57], or *via* gene propagation in mutator strain[58] bacteria. Libraries of mutated genes are subjected to selections for clones of enhanced affinity mimicking the clonal selection and amplification of the natural immune system.

ANTIBODY LIBRARY DISPLAY AND SELECTION METHODS

Phage Display

Display on the surface of M13, a filamentous bacteriophage, is perhaps the most widespread and successful method used to isolate antibodies from large combinatorial antibody libraries[42]. Phage display has been used to isolate high-affinity human antibodies from many types of libraries, including biased and naïve immune antibody libraries and synthetic antibody libraries [48; 59; 60]. With phage display, the genes of proteins or peptides can be fused to the gene of one of the coat proteins of the phage virion. The minor coat protein pIII, which mediates the phage attachment to the cell

during infection, is usually used as the fusion partner in phage display of antibody fragments. This leads to the production of phage particles with the protein or peptide displayed on the surface of the phage as a fusion to the outer coat protein. To display antibody fragment libraries on phage, the library is first cloned into a special type of plasmid called phagemid, between the carboxyl terminus of a secretion leader sequence and the amino terminus of pIII. Phagemids contains a phage origin of replication to enable packing of the phagemid DNA into phage particles and also a selectable marker to ensure selection of the cells containing the phagemid. The phagemid DNA is then transformed into the host *E. coli*. Therefore, the size of resulting phage library is limited by the transformation inefficiencies of *E. coli*.

A culture of the cells transformed with phagemid is infected with M13 "helper" phage that provides all of the genes required for replication and packaging of the phagemid into the phage virion. The helper phage contain a defective origin of replication or packaging signal resulting in the preferential packing of the phagemid genome over the helper phage genome. After infection, the expression of the antibody-pIII fusions from the phagemid is induced. The wild-type pIII, encoded by the helper phage genome, and the antibody-pIII fusion, encoded by the phagemid, compete for incorporation into the virion. As a result, the antibody is not displayed on every one of the five pIII of each particle. Between zero and five copies of the antibody will be displayed (Figure 1-3). Some of the phage particles that emerge from a cell will be decorated with the antibody variant encoded by the phagemid contained within the particle (connecting the phenotype to the genotype). These particles can be panned for binding to a target antigen and amplified.

It is desirable to maintain a low valency of antibody-pIII fusions on the surface of phage particles in antibody engineering directed evolution efforts so that avidity is minimized and high-affinity clones can be enriched. Assuming that an immobilized antigen results in a densely coated binding surface, avidity effects will occur if multiple copies of a displayed antibody on one phage particle bind to the immobilized antigen. Multiple binding events per phage particle will enable even low-affinity clones to be selected. Because wild-type pIII is necessary for the phage to infect, low valency helps ensure that there are enough copies of the wild type pIII on the selected phage particles to enable infection after each round of panning.

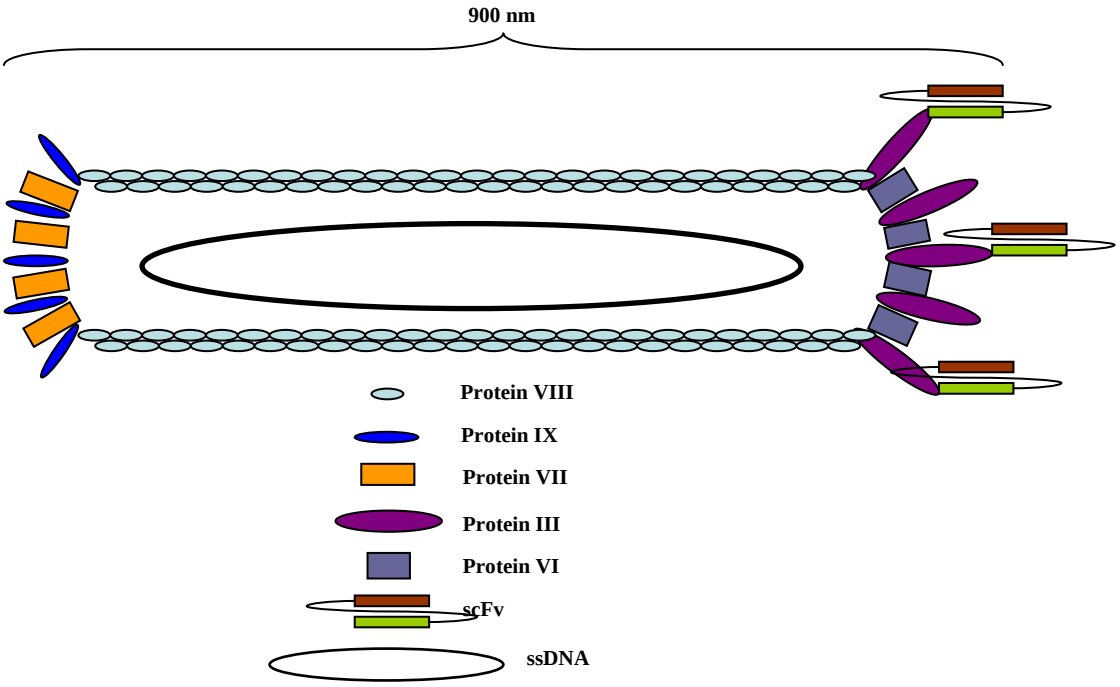


Figure 1-3. Filamentous phage structure.
Filamentous phage particles consist of single-stranded DNA encapsulated in the bacteriophage capsid made up of the major coat protein, pVIII. The particle is ejected from infected bacteria with the end containing pIX and pVII leaving first. pIII and pVI are located at the other end. pIII is important as it is the protein that recognizes the pilus of the bacterium to initiate infection. Fusions of antibody libraries to pIII are the most common type of phage display.

Panning selections consist of iterative rounds of adsorption of the phage-library and desorption of bound and washed phage particles (Figure 1-4). The library population is enriched from round-to-round by increasing the stringency of binding to select for clones exhibiting enhanced affinity. Stringency can be increased by limiting the concentration of the antigen or by using extensive washings, both of which will select for antibodies with slower dissociation rates. Faster association rates can be selected by limiting the exposure time of the library to the antigen[38]. Solution-phase capture is another strategy to attenuate the avidity effects leading to the selection of low-affinity antibodies[56]. In solution-phase capture, the phage antibodies and a biotin-labeled antigen are allowed to interact in solution until equilibrium has been reached. The antibody-antigen complex is then pulled down, typically using the biotin on the antigen and a streptavidin on a solid support such as a bead. By allowing the binding interactions to reach equilibrium in solution, avidity effects are minimized and the highest affinity interactions emerge from the selections.

Between rounds of selection, the eluted phage particles are used to infect *E. coli* in order to propagate enriched clones and amplify them. This output culture of *E. coli* is used to produce phage particle for subsequent rounds of selection or for screening.

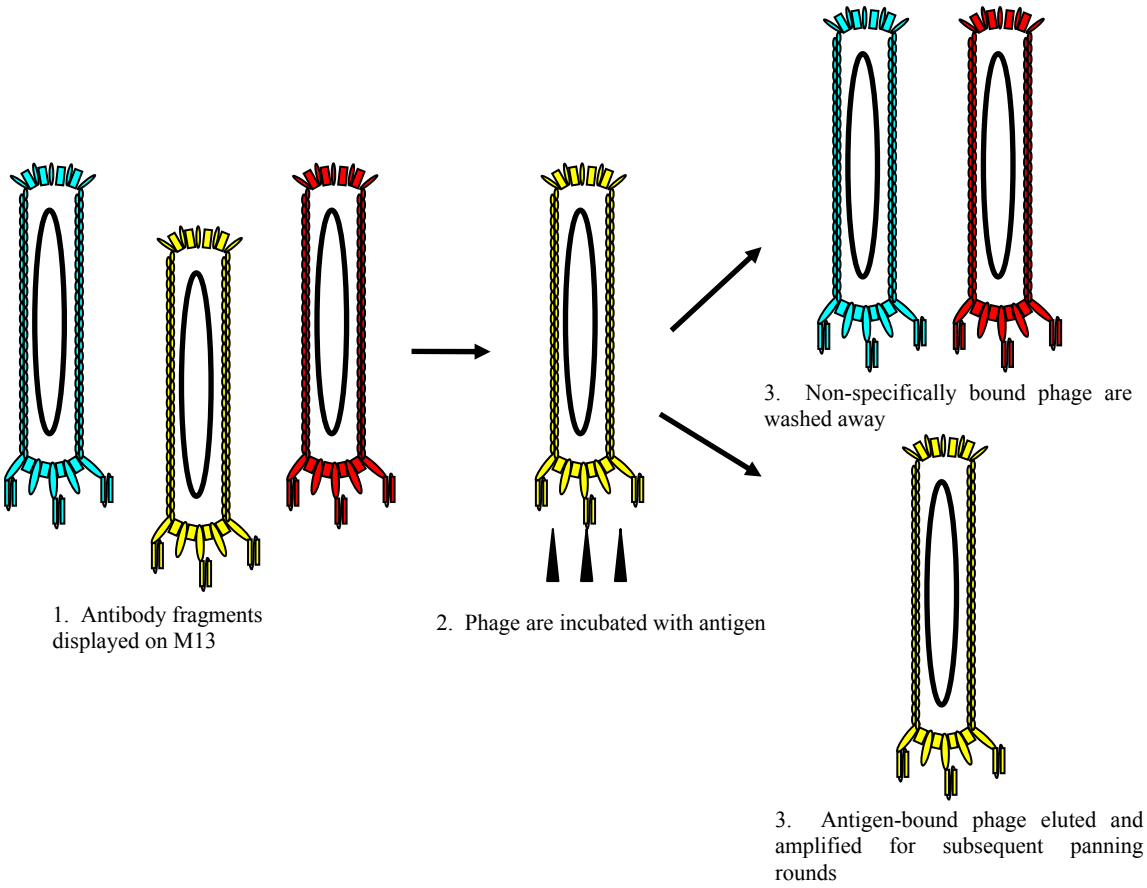


Figure 1-4. Filamentous phage display panning of scFv library.
In step 1, the phage particles displaying the antibody fragment library on an outer coat protein have been secreted from host *E. coli*. In step 2, the phage library is incubated in solution with the target antigen, represented by the triangles. The non-specific or low-affinity phage are washed away during the washing stage of step 3, while the specifically, tightly bound phage are eluted, collected, and used to infect a culture of the host *E. coli*.

Enrichment can be monitored using polyclonal phage ELISA. Briefly, dilutions of normalized solutions of phage from each round of selection are incubated with the target antigen that has been immobilized on an ELISA plate. The binding of the polyclonal phage is detected using an antibody specific to the phage coat. Enrichment between rounds is evident when equal numbers of phage from one round elicits a stronger signal than phage from a previous round. This stronger signal indicates that more of the phage particles are binding than from the sample with the weaker signal. Once enrichment

reaches an acceptable level, individual clones are screened for enhancements. Sequence diversity following different rounds of selection is monitored by DNA sequencing of randomly picked clones. The diversity of the polyclonal population is high in the early rounds of panning, generally becoming much lower in later rounds.

Ribosome display

Ribosome display is a method used for selecting proteins, including antibody fragments, from large combinatorial libraries expressed by cell-free, *in vitro* transcription and translation (IV-T&T) [61; 62]. Due to the cell-free nature of this method, the library size is not limited by transformation efficiency. As a result, libraries with diversities on the order of 10^{13} or greater can be screened[63]. Using IV T&T, the ribosome translates transcribed genes from an scFv library. Each library variant is translated along with a spacer arm to tether the folded scFv to the polypeptide portion of the ribosome[61]. A rabbit reticulocyte IV T&T system provides monovalent display of the antibodies which lowers avidity effects during the ribosome display selections[64]. After the scFv synthesis has been completed, and before peptide release, the scFv-ribosome-mRNA complex is stabilized using a combination of high salt and low temperatures, preventing the release of the mRNA and the polypeptide[61]. One of the drawbacks of using ribosome display to engineer antibodies is that the conditions required to stabilize the scFv-ribosome-mRNA complex, are not always ideal for antibody-antigen interactions. These particles are then panned against an antigen of interest and eluted. The recovered particles contain the genes of the isolated clones which are amplified and reverse transcribed back into cDNA for subsequent rounds of selection or screening. During the amplification step, non-

proofreading polymerases provide additional diversity between rounds of selection[61]. This method has been used to isolate anti-GCN4 (a yeast transcription factor) scFvs from a mouse immune library[53], to affinity and stability mature DNA-shuffled scFvs[57], and to select high-affinity binders from naïve immune libraries[65].

Microbial cell display and flow cytometry

Another method used to link genotype to phenotype during selections is the use of microbial cells that display the proteins encoded by the library at sites that can be exposed to added antigen. Such sites may include the cell surface, inner membrane, or the periplasmic space of Gram-negative bacteria[66; 67; 68; 69]. The most commonly used microbes for protein display are the bacterium *E. coli* and the yeast *Saccharomyces cerevisiae*. Because microbial cells are relatively large, flow cytometry (FC) can be used to screen these libraries[70] (Figure 1-5). To use this method for affinity selections, the microbes displaying an antibody library are incubated with a fluorescently-labeled antigen[67]. FC is able to interrogate individual cells quantitatively at rates of up to 10^9 cells/hour. The light scattering (side and forward scatter) of each cell triggers the flow cytometer to measure the fluorescence intensity of the cell. The fluorescence intensity is a direct measure of the amount of bound fluorescent antigen. FC sorting parameters are set to collect the most fluorescent cells, a pool which corresponds to the cells that have the most bound antigen[67; 71].

The number of bound antigen molecules per cell depends on several factors: the affinity of the expressed antibody for the antigen, the expression level of the antibody, *i.e.* the number of protein molecules expressed on the bacteria, and the proper folding of the

protein. While the primary parameter of antibody selection is typically affinity, the latter two parameters are important to the efficient scale-up production of the protein, if necessary. Microbial display provides for selection based on all three parameters in contrast to ribosome and phage display which have very little selective pressure for expression. As a result, antibodies isolated by ribosome or phage display are often expressed poorly in bacteria or higher cells. The real-time quantitative analysis of parameters such as the kinetics and equilibrium coupled with high-throughput screening[71] makes FC and microbial display a powerful combination when it comes to screening antibody libraries. Enrichment between rounds of selection can be assayed without any additional experiments by comparing the fluorescence of each round's total population of cells. The fluorescence corresponds to the antigen binding potential, or enrichment, of the cell population from each round of selection.

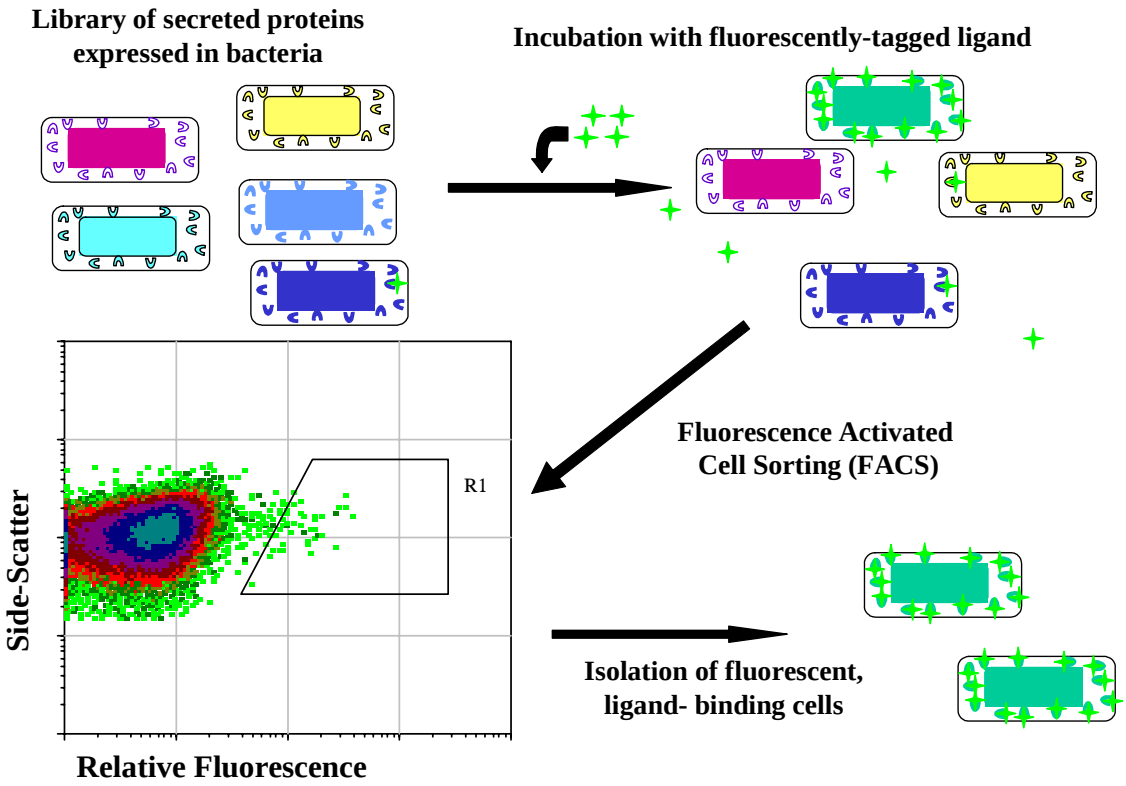


Figure 1-5: Workflow of flow cytometry-based selection of antibodies using microbial display.
The library of antibodies is expressed and anchored to either the inner membrane or outer membrane of the host microbe. The cells are incubated with a fluorescent version of the antigen and sorted on a flow cytometer. The most fluorescent cells are collected for subsequent rounds of selection or characterization. (adapted from Barrett Harvey)

PEPTIDES

Peptides are short chains of amino acids linked by amide bonds. Peptides are usually distinguished from proteins in that they typically lack extensive secondary structure. Peptides have physiological relevance as hormones, neurotransmitters, or other types of signaling molecules in humans. They are most often linear and flexible, but at times, they may have a constrained lariat structure as a result of covalent linking *e.g.* by

disulfide bonds, at the termini. Due to the flexible nature of peptides in the unbound state and the loss of entropy in the antibody-bound state[72], they have been an elusive target for high-affinity antibodies. In most of the cases involving large, commercial, synthetic libraries based on human germline sequences, the published K_Ds of the isolated anti-peptide antibodies are no better than > 1μM[48; 73]. Although there are a few exceptions of anti-peptide antibodies with K_Ds in the high-nanomolar range from synthetic libraries[74], most of the libraries are validated against protein or hapten targets[44; 59; 60]. Selections from even the largest of these libraries can result in few or no reported hits against peptide antigens, suggesting a lack of peptide specificity[59; 60; 73]. Immune repertoire libraries have resulted in antibodies with far greater affinity (picomolar K_Ds) to peptides[53]. Affinity maturation using ribosome display of an antibody isolated from an immune repertoire library has resulted in an anti-peptide antibody with a K_D of 1 pM[75].

In vertebrates, the endocrine system secretes dozens of peptide hormones. Other peptides modulate the digestive system and secretions, help regulate the heart, stimulate smooth muscle, and act as neurotransmitters[76]. A single peptide hormone or neuropeptide can have several different functions depending on the cell type targeted and the receptor activated[76]. For example, the tachykinin neuropeptide called *substance P*, an 11 amino acid peptide, plays a role in vomiting, sensing pain, cell growth, mood, and even vasodilation[77]. After being cleaved from the precursor angiotensinogen, the liver deca-peptide hormone, angiotensin, affects blood pressure through the feedback of the renin-angiotensin system, stimulates the release of aldosterone, and even causes the thirst sensation[78]. Neuropeptide Y, a 36 amino acid neurotransmitter, causes vasoconstriction, is involved in regulating eating behaviors in response to energy needs, and plays a role in

memory and epilepsy[79]. Peptides are also used by lower vertebrates and prokaryotes. Some peptides secreted by host-invading prokaryotes can act as toxins in infected hosts, while some animals use peptides as venom to subdue prey and discourage predators[80].

Peptides as alternatives to whole protein antigens

Most antibody epitopes are conformational and specified by interactions with topographic determinants on the surface of the antigenic molecule. To define epitopes on proteins for a particular antibody, peptide libraries have been developed to help identify critical amino acids, and structural and functional features that are important for binding of the antibody-of-interest to the epitope [81]. Isolated peptides can then be mapped back to the protein using sequence, structural, or functional homology to identify the epitope. This concept of cross-reactivity between peptide mimics and protein epitopes has been adapted to the panning of phage-displayed antibody libraries. In these cases, panning a library against peptides that represent fragments of a larger protein has enabled isolation of antibodies that are cross-reactive against the larger protein. This method can be advantageous when the protein target is rare, expensive, or unstable in its native form (*e.g.* transmembrane proteins). For example, the principal antibody-mediated neutralizing determinant of the human immunodeficiency virus type 1 (HIV-1) is contained in the third hypervariable region (V3) of envelope glycoprotein, gp120[82]. This entire determinant can be represented by smaller, easier to make, more stable peptides which can serve as the target for antibody libraries. Antibody clones isolated from these libraries are typically cross-reactive with the complete gp120 protein and can neutralize the virus *in vivo*[83]. Through immunization with the gp120 peptide mimics, broadly

neutralizing antibodies to V3 have been isolated from mice laying the groundwork for the peptides to be used as a vaccine against HIV-1[84]. Peptide mimics of carbohydrate, protein, and even hapten targets are now routinely used to safely immunize mice in order to isolate antibodies against mimicked target protein, virus, or other types of pathogens[85; 86; 87; 88]. Due to the cross-reactivity of some antibodies to both peptide mimics and full-sized proteins, libraries that are biased for peptide binders ought to contain clones that can recognize peptides as well as related proteins. If the peptide mimic maps to a sequence on the related protein, it is simple to identify the epitope of any antibodies that are cross-reactive between the two.

Anti-peptide antibodies in detection and therapies

High-affinity, peptide-specific antibodies have been used to redefine expression patterns determined using lower-affinity *protein* specific antibodies. For example, antibodies generated against a 20-amino acid peptide representing an epitope of cytochrome 450-B26 (a liver enzyme) were able to detect femtomolar amounts of the cytochrome protein, while an antibody isolated from a screen using the entire cytochrome protein was unable to detect any of the protein[89]. A monoclonal antibody generated to a peptide mimic of fibrinogen is capable of monitoring fibrinogenolysis in biochemical assays[90]. Anti-peptide antibodies isolated from rabbits immunized with a 12 amino acid mimotope of the 636 amino acid endotoxin A from *Pseudomonas aeruginosa* are able to neutralize the toxin *in vivo*[91; 92].

DISSERTATION SUMMARY

Chapter two examines the construction and experimental validation of synthetic libraries biased towards anti-peptide antibodies. Canonical structure patterns and the natural diversity of anti-peptide antibodies were analyzed. Common features of anti-peptide antibodies gleaned from this analysis were incorporated into two different anti-peptide synthetic libraries. For each library, diversity was limited by randomizing only positions that are commonly used to make contact with the antigen (peptides) in known peptide binders. These positions were varied to incorporate only the diversity observed in natural anti-peptide antibodies. To limit sequence space, both libraries randomize positions in only one domain (V_H only) of a single scaffold antibody. The libraries were selected against two peptides, angiotensin I (10 amino acids) and neuropeptide Y (36 amino acids) encompassing a broad size range, using phage display with the isolation of several specific, high-affinity clones. This is the first report of canonical structures patterns being leveraged to create antigen-class specific libraries.

When using affinity maturation in antibody engineering, error-rates of less than 1% are commonly used to introduce diversity[56; 58]. These low error rates are used to simulate the gradual rate of genetic change thought to occur in natural evolution, to avoid the accumulation of deleterious mutations that would limit the library's overall fitness, and to limit the amount of sequence space to match what can reasonably be surveyed by current selection methods[93]. However, there are several reports that show benefits to using high

error-rates of mutagenesis in combinatorial libraries to survey unusual, but beneficial combinations of mutations in protein libraries[93; 94; 95].

In chapter three, high error-rate mutagenesis is extended in the search for rational, enhancing mutations to an antibody. Mutational hotspots were identified from sequences of selected affinity-enhanced clones from high error-rate libraries. Combinations of these consensus mutations were analyzed for enhancements to the wild-type antibody. It was discovered that several of the mutations provide minor, non-synergistic enhancements to the kinetics and expression of the wild-type antibody.

REFERENCES

1. Medzhitov, R. & Janeway, C., Jr. (2000). Innate immunity. *N Engl J Med* **343**, 338-44.
2. Medzhitov, R. & Janeway, C., Jr. (2000). Innate immune recognition: mechanisms and pathways. *Immunol Rev* **173**, 89-97.
3. Casadevall, A. & Scharff, M. D. (1995). Return to the past: the case for antibody-based therapies in infectious diseases. *Clin Infect Dis* **21**, 150-61.
4. Keller, M. A. & Stiehm, E. R. (2000). Passive immunity in prevention and treatment of infectious diseases. *Clin Microbiol Rev* **13**, 602-14.
5. Swanson, J. A. & Hoppe, A. D. (2004). The coordination of signaling during Fc receptor-mediated phagocytosis. *J Leukoc Biol* **76**, 1093-103.
6. Casadevall, A. (2002). Passive antibody administration (immediate immunity) as a specific defense against biological weapons. *Emerg Infect Dis* **8**, 833-41.
7. Lang, A. B., Cryz, S. J., Jr., Schurch, U., Ganss, M. T. & Bruderer, U. (1993). Immunotherapy with human monoclonal antibodies. Fragment A specificity of polyclonal and monoclonal antibodies is crucial for full protection against tetanus toxin. *J Immunol* **151**, 466-72.
8. Slade, H. B. (1994). Human Immunoglobulins for intravenous use and hepatitis C viral transmission. *Clin Diagn Lab Immunol* **1**, 613-9.
9. Amzel, L. & Poljak, R. (1979). Three-Dimensional Structure of Immunoglobulins. *Annual Review of Biochemistry* **48**, 961-997.
10. Padlan, E. A. (1977). Structural implications of sequence variability in immunoglobulins. *Proc Natl Acad Sci U S A* **74**, 2551-5.
11. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J Biol Chem* **252**, 6609-16.

12. Kabat, E. A. & Wu, T. T. (1971). Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann N Y Acad Sci* **190**, 382-93.
13. Padlan, E. A. & Davies, D. R. (1975). Variability of three-dimensional structure in immunoglobulins. *Proc Natl Acad Sci U S A* **72**, 819-23.
14. Yamaguchi, Y., Kim, H., Kato, K., Masuda, K., Shimada, I. & Arata, Y. (1995). Proteolytic fragmentation with high specificity of mouse immunoglobulin G. Mapping of proteolytic cleavage sites in the hinge region. *J Immunol Methods* **181**, 259-67.
15. Chothia, C. & Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* **196**, 901-17.
16. Alzari, P. M., Lascombe, M. B. & Poljak, R. J. (1988). Three-dimensional structure of antibodies. *Annu Rev Immunol* **6**, 555-80.
17. Padlan, E. A. (1979). Evaluation of the structural variation among light chain variable domains. *Mol Immunol* **16**, 287-96.
18. Kabat, E. A., Padlan, E. A. & Davies, D. R. (1975). Evolutionary and structural influences on light chain constant (CL) region of human and mouse immunoglobulins. *Proc Natl Acad Sci U S A* **72**, 2785-8.
19. Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R. & et al. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* **342**, 877-83.
20. Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). Structural repertoire of the human VH segments. *J Mol Biol* **227**, 799-817.
21. Martin, A. C. & Thornton, J. M. (1996). Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J Mol Biol* **263**, 800-15.
22. Chothia, C. & Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *Embo J* **5**, 823-6.
23. Milstein, C. (1999). The hybridoma revolution: an offshoot of basic research. *Bioessays* **21**, 966-73.
24. Kohler, G. & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-7.
25. Presta, L. G. (2006). Engineering of therapeutic antibodies to minimize immunogenicity and optimize function. *Adv Drug Deliv Rev* **58**, 640-56.
26. Khazaeli, M. B., Conry, R. M. & LoBuglio, A. F. (1994). Human immune response to monoclonal antibodies. *J Immunother Emphasis Tumor Immunol* **15**, 42-52.
27. Mirick, G. R., Bradt, B. M., Denardo, S. J. & Denardo, G. L. (2004). A review of human anti-globulin antibody (HAGA, HAMA, HACA, HAHA) responses to monoclonal antibodies. Not four letter words. *Q J Nucl Med Mol Imaging* **48**, 251-7.
28. Orlandi, R., Gussow, D. H., Jones, P. T. & Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc Natl Acad Sci U S A* **86**, 3833-7.

29. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S. & Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science* **242**, 423-6.
30. Raag, R. & Whitlow, M. (1995). Single-chain Fvs. *Faseb J* **9**, 73-80.
31. van Dijk, M. A. & van de Winkel, J. G. (2001). Human antibodies as next generation therapeutics. *Curr Opin Chem Biol* **5**, 368-74.
32. Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* **321**, 522-5.
33. Queen, C., Schneider, W. P., Selick, H. E., Payne, P. W., Landolfi, N. F., Duncan, J. F., Avdalovic, N. M., Levitt, M., Junghans, R. P. & Waldmann, T. A. (1989). A humanized antibody that binds to the interleukin 2 receptor. *Proc Natl Acad Sci U S A* **86**, 10029-33.
34. Zhang, W., Feng, J., Li, Y., Guo, N. & Shen, B. (2005). Humanization of an anti-human TNF-alpha antibody by variable region resurfacing with the aid of molecular modeling. *Mol Immunol* **42**, 1445-51.
35. Lonberg, N. (2005). Human antibodies from transgenic animals. *Nat Biotechnol* **23**, 1117-25.
36. Green, L. L., Hardy, M. C., Maynard-Currie, C. E., Tsuda, H., Louie, D. M., Mendez, M. J., Abderrahim, H., Noguchi, M., Smith, D. H., Zeng, Y. & et al. (1994). Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. *Nat Genet* **7**, 13-21.
37. Lonberg, N., Taylor, L. D., Harding, F. A., Trounstine, M., Higgins, K. M., Schramm, S. R., Kuo, C. C., Mashayekh, R., Wymore, K., McCabe, J. G. & et al. (1994). Antigen-specific human antibodies from mice comprising four distinct genetic modifications. *Nature* **368**, 856-9.
38. Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. *Nat Biotechnol* **23**, 1105-16.
39. Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T. & Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. *Nature* **341**, 544-6.
40. Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting-Mees, M., Burton, D. R., Benkovic, S. J. & Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246**, 1275-81.
41. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature* **352**, 624-8.
42. Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. & Winter, G. (1991). By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol* **222**, 581-97.
43. Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. (1994). Making antibodies by phage display technology. *Annu Rev Immunol* **12**, 433-55.
44. Hoet, R. M., Cohen, E. H., Kent, R. B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H., van Hegelsom, R., Neer, N. C., Nastri, H. G., Rondon, I. J., Leeds, J. A., Hufton, S. E., Huang, L., Kashin, I., Devlin, M., Kuang, G., Steukers, M., Viswanathan, M., Nixon, A. E., Sexton, D. J., Hoogenboom, H. R. & Ladner, R. C. (2005). Generation of high-affinity

human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nat Biotechnol* **23**, 344-8.
45. Sidhu, S. S. & Fellouse, F. A. (2006). Synthetic therapeutic antibodies. *Nat Chem Biol* **2**, 682-8.
46. de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P., de Bruine, A. P., Arends, J. W. & Hoogenboom, H. R. (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J Biol Chem* **274**, 18218-30.
47. Sidhu, S. S., Li, B., Chen, Y., Fellouse, F. A., Eigenbrot, C. & Fuh, G. (2004). Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. *J Mol Biol* **338**, 299-310.
48. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. & Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* **296**, 57-86.
49. Fellouse, F. A., Wiesmann, C. & Sidhu, S. S. (2004). Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc Natl Acad Sci U S A* **101**, 12467-72.
50. Morea, V., Lesk, A. M. & Tramontano, A. (2000). Antibody modeling: implications for engineering and design. *Methods* **20**, 267-79.
51. Cadwell, R. C. & Joyce, G. F. (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**, 28-33.
52. Stemmer, W. P. (1994). DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc Natl Acad Sci U S A* **91**, 10747-51.
53. Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H. R. & Pluckthun, A. (1998). Ribosome display efficiently selects and evolves high-affinity antibodies *in vitro* from immune libraries. *Proc Natl Acad Sci U S A* **95**, 14130-5.
54. Hayhurst, A. & Georgiou, G. (2001). High-throughput antibody isolation. *Curr Opin Chem Biol* **5**, 683-9.
55. Marks, J. D. & Bradbury, A. (2004). Selection of human antibodies from phage display libraries. *Methods Mol Biol* **248**, 161-76.
56. Hawkins, R. E., Russell, S. J. & Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J Mol Biol* **226**, 889-96.
57. Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Pluckthun, A. (2001). Tailoring *in vitro* evolution for protein affinity or stability. *Proc Natl Acad Sci U S A* **98**, 75-80.
58. Low, N. M., Holliger, P. H. & Winter, G. (1996). Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *J Mol Biol* **260**, 359-68.
59. Griffiths, A., Williams, S., Hartley, O., Tomlinson, I., Waterhouse, P., Crosby, W., Kontermann, R., Jones, P., Low, N., Allison, T., Prospero, T., Hoogenboom, H., Nissim, A., Cox, J., Harrison, J., Zaccolo, M., Gherardi, E. & Winter, G. (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J* **13**, 3245–3260.

60. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. & Johnson, K. S. (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat Biotechnol* **14**, 309-14.
61. Hanes, J. & Pluckthun, A. (1997). *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **94**, 4937-42.
62. Xu, L., Aha, P., Gu, K., Kuimelis, R. G., Kurz, M., Lam, T., Lim, A. C., Liu, H., Lohse, P. A., Sun, L., Weng, S., Wagner, R. W. & Lipovsek, D. (2002). Directed evolution of high-affinity antibody mimics using mRNA display. *Chem Biol* **9**, 933-42.
63. Mattheakis, L. C., Bhatt, R. R. & Dower, W. J. (1994). An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc Natl Acad Sci U S A* **91**, 9022-6.
64. Ramachandiran, V., Kramer, G. & Hardesty, B. (2000). Expression of different coding sequences in cell-free bacterial and eukaryotic systems indicates translational pausing on Escherichia coli ribosomes. *FEBS Lett* **482**, 185-8.
65. Hanes, J., Schaffitzel, C., Knappik, A. & Pluckthun, A. (2000). Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat Biotechnol* **18**, 1287-92.
66. Harvey, B. R., Georgiou, G., Hayhurst, A., Jeong, K. J., Iverson, B. L. & Rogers, G. K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. *Proc Natl Acad Sci U S A* **101**, 9193-8.
67. Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. & Curtiss, R., 3rd. (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* **15**, 29-34.
68. Boder, E. T. & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* **15**, 553-7.
69. Chen, G., Hayhurst, A., Thomas, J. G., Harvey, B. R., Iverson, B. L. & Georgiou, G. (2001). Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol* **19**, 537-42.
70. Daugherty, P. S., Chen, G., Olsen, M. J., Iverson, B. L. & Georgiou, G. (1998). Antibody affinity maturation using bacterial surface display. *Protein Eng* **11**, 825-32.
71. Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., Cochran, J. R., Heinzelman, P., Colby, D., Swers, J., Graff, C., Wiley, H. S. & Wittrup, K. D. (2003). Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. *Nat Biotechnol* **21**, 163-70.
72. Zahnd, C., Spinelli, S., Luginbühl, B., Amstutz, P., Cambillau, C. & Plückthun, A. (2004). Directed *in Vitro* Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. *J Biol Chem* **279**, 18870-18877.
73. Krebs, B., Rauchenberger, R., Reiffert, S., Rothe, C., Tesar, M., Thomassen, E., Cao, M., Dreier, T., Fischer, D., Hoss, A., Inge, L., Knappik, A., Marget, M.,

Pack, P., Meng, X. Q., Schier, R., Sohlemann, P., Winter, J., Wolle, J. & Kretzschmar, T. (2001). High-throughput generation and engineering of recombinant human antibodies. *J Immunol Methods* **254**, 67-84.

74. Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C., Danielsson, L., Carlsson, R. & Borrebaeck, C. A. (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol* **18**, 852-6.
75. Luginbuhl, B., Kanyo, Z., Jones, R. M., Fletterick, R. J., Prusiner, S. B., Cohen, F. E., Williamson, R. A., Burton, D. R. & Pluckthun, A. (2006). Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation. *J Mol Biol* **363**, 75-97.
76. Roth, J. (1973). Peptide hormone binding to receptors: a review of direct studies *in vitro*. *Metabolism* **22**, 1059-73.
77. Regoli, D., Boudon, A. & Fauchere, J. L. (1994). Receptors and antagonists for substance P and related peptides. *Pharmacol Rev* **46**, 551-99.
78. Phillips, M. I., Speakman, E. A. & Kimura, B. (1993). Levels of angiotensin and molecular biology of the tissue renin angiotensin systems. *Regul Pept* **43**, 1-20.
79. Dumont, Y., Martel, J. C., Fournier, A., St-Pierre, S. & Quirion, R. (1992). Neuropeptide Y and neuropeptide Y receptor subtypes in brain and peripheral tissues. *Prog Neurobiol* **38**, 125-67.
80. Moczydlowski, E., Lucchesi, K. & Ravindran, A. (1988). An emerging pharmacology of peptide toxins targeted against potassium channels. *J Membr Biol* **105**, 95-111.
81. Rowley, M. J., O'Connor, K. & Wijeyewickrema, L. (2004). Phage display for epitope determination: a paradigm for identifying receptor-ligand interactions. *Biotechnol Annu Rev* **10**, 151-88.
82. Rusche, J. R., Javaherian, K., McDanal, C., Petro, J., Lynn, D. L., Grimaila, R., Langlois, A., Gallo, R. C., Arthur, L. O., Fischinger, P. J. & et al. (1988). Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. *Proc Natl Acad Sci U S A* **85**, 3198-202.
83. Barbas, C. F., 3rd, Amberg, W., Simoncsits, A., Jones, T. M. & Lerner, R. A. (1993). Selection of human anti-hapten antibodies from semisynthetic libraries. *Gene* **137**, 57-62.
84. White-Scharf, M. E., Potts, B. J., Smith, L. M., Sokolowski, K. A., Rusche, J. R. & Silver, S. (1993). Broadly neutralizing monoclonal antibodies to the V3 region of HIV-1 can be elicited by peptide immunization. *Virology* **192**, 197-206.
85. Brett, P. J., Tiwana, H., Feavers, I. M. & Charalambous, B. M. (2002). Characterization of oligopeptides that cross-react with carbohydrate-specific antibodies by real time kinetics, in-solution competition enzyme-linked immunosorbent assay, and immunological analyses. *J Biol Chem* **277**, 20468-76.
86. Charalambous, B. M. & Feavers, I. M. (2000). Peptide mimics elicit antibody responses against the outer-membrane lipooligosaccharide of group B neisseria meningitidis. *FEMS Microbiol Lett* **191**, 45-50.
87. Charalambous, B. M. & Feavers, I. M. (2001). Mimotope vaccines. *J Med Microbiol* **50**, 937-9.

88. Tiwana, H., Clow, K. J., Hall, C., Feavers, I. M. & Charalambous, B. M. (2005). The immunogenicity of a conformationally restricted peptide mimetic of meningococcal lipooligosaccharide. *Scand J Immunol* **62**, 385-92.
89. Stresser, D. M. & Kupfer, D. (1999). Monospecific antipeptide antibody to cytochrome P-450 2B6. *Drug Metab Dispos* **27**, 517-25.
90. Sobel, J. H., Trakht, I., Pileggi, N. & Qi Wu, H. (1998). Antipeptide monoclonal antibodies to defined fibrinogen Aalpha chain regions: anti-Aalpha 487-498, a structural probe for fibrinogenolysis. *Blood* **91**, 1590-8.
91. Elzaim, H. S., Chopra, A. K., Peterson, J. W., Goodheart, R. & Heggers, J. P. (1998). Generation of neutralizing antipeptide antibodies to the enzymatic domain of Pseudomonas aeruginosa exotoxin A. *Infect Immun* **66**, 2170-9.
92. El-Zaim, H. S., Chopra, A. K., Peterson, J. W., Vasil, M. L. & Heggers, J. P. (1998). Protection against exotoxin A (ETA) and Pseudomonas aeruginosa infection in mice with ETA-specific antipeptide antibodies. *Infect Immun* **66**, 5551-4.
93. Daugherty, P. S., Chen, G., Iverson, B. L. & Georgiou, G. (2000). Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proc Natl Acad Sci U S A* **97**, 2029-34.
94. Shafikhani, S., Siegel, R. A., Ferrari, E. & Schellenberger, V. (1997). Generation of large libraries of random mutants in Bacillus subtilis by PCR-based plasmid multimerization. *Biotechniques* **23**, 304-10.
95. Zaccolo, M. & Gherardi, E. (1999). The effect of high-frequency random mutagenesis on *in vitro* protein evolution: a study on TEM-1 beta-lactamase. *J Mol Biol* **285**, 775-83.
96. Beckman, R. A., Mildvan, A. S. & Loeb, L. A. (1985). On the fidelity of DNA replication: manganese mutagenesis *in vitro*. *Biochemistry* **24**, 5810-7.
97. Eckert, K. A. & Kunkel, T. A. (1991). DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Appl* **1**, 17-24.
98. Arnold, F. H. (1998). Enzyme engineering reaches the boiling point. *Proc Natl Acad Sci U S A* **95**, 2035-6.
99. Suzuki, M., Christians, F. C., Kim, B., Skandalis, A., Black, M. E. & Loeb, L. A. (1996). Tolerance of different proteins for amino acid diversity. *Mol Divers* **2**, 111-8.
100. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science* **247**, 1306-10.
101. Drummond, D. A., Iverson, B. L., Georgiou, G. & Arnold, F. H. (2005). Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol* **350**, 806-16.
102. Burks, E. A., Chen, G., Georgiou, G. & Iverson, B. L. (1997). *In vitro* scanning saturation mutagenesis of an antibody binding pocket. *Proc Natl Acad Sci U S A* **94**, 412-7.
103. Chen, G., Dubrawsky, I., Mendez, P., Georgiou, G. & Iverson, B. L. (1999). *In vitro* scanning saturation mutagenesis of all the specificity determining residues in an antibody binding site. *Protein Eng* **12**, 349-56.

104. Brown, M., Rittenburg, M. B., Chen, C. & Roberts, V. A. (1996). Tolerance of single, but not multiple, amino acid replacements in antibody VH CDR 2: a means of minimizing B cell wastage from somatic hypermutation? *J Immunol* **156**, 3285-91.
105. Ito, W., Iba, Y. & Kurosawa, Y. (1993). Effects of substitutions of closely related amino acids at the contact surface in an antigen-antibody complex on thermodynamic parameters. *J Biol Chem* **268**, 16639-47.
106. Wong, Y. W., Kussie, P. H., Parhami-Seren, B. & Margolies, M. N. (1995). Modulation of antibody affinity by an engineered amino acid substitution. *J Immunol* **154**, 3351-8.
107. Short, M. K., Jeffrey, P. D., Kwong, R. F. & Margolies, M. N. (1995). Contribution of antibody heavy chain CDR1 to digoxin binding analyzed by random mutagenesis of phage-displayed Fab 26-10. *J Biol Chem* **270**, 28541-50.
108. Chen, C., Roberts, V. A., Stevens, S., Brown, M., Stenzel-Poore, M. P. & Rittenberg, M. B. (1995). Enhancement and destruction of antibody function by somatic mutation: unequal occurrence is controlled by V gene combinatorial associations. *Embo J* **14**, 2784-94.
109. Ramirez-Benitez, M. C. & Almagro, J. C. (2001). Analysis of antibodies of known structure suggests a lack of correspondence between the residues in contact with the antigen and those modified by somatic hypermutation. *Proteins* **45**, 199-206.
110. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & et al. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci U S A* **85**, 5879-83.
111. Sun, F. (1995). The polymerase chain reaction and branching processes. *J Comput Biol* **2**, 63-86.
112. Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B. L. & Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. *J Immunol Methods* **276**, 185-96.
113. Hayhurst, A. & Harris, W. J. (1999). Escherichia coli skp chaperone coexpression improves solubility and phage display of single-chain antibody fragments. *Protein Expr Purif* **15**, 336-43.
114. Ayala, M., Balint, R. F., Fernandez-de-Cossio, L., Canaan-Haden, J. W., Larrick, J. W. & Gavilondo, J. V. (1995). Variable region sequence modulates periplasmic export of a single-chain Fv antibody fragment in Escherichia coli. *Biotechniques* **18**, 832, 835-8, 840-2.
115. Knappik, A. & Pluckthun, A. (1995). Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Eng* **8**, 81-9.
116. Moore, G. L. & Maranas, C. D. (2000). Modeling DNA mutation and recombination for directed evolution experiments. *J Theor Biol* **205**, 483-503.
117. Bond, C. J., Wiesmann, C., Marsters, J. C., Jr. & Sidhu, S. S. (2005). A structure-based database of antibody variable domain diversity. *J Mol Biol* **348**, 699-709.
118. Garcia-Rodriguez, C., Levy, R., Arndt, J. W., Forsyth, C. M., Razai, A., Lou, J., Geren, I., Stevens, R. C. & Marks, J. D. (2007). Molecular evolution of antibody

cross-reactivity for two subtypes of type A botulinum neurotoxin. *Nat Biotechnol* **25**, 107-16.
119. Foote, J. & Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. *Proc Natl Acad Sci U S A* **91**, 10370-4.
120. Mian, I. S., Bradwell, A. R. & Olson, A. J. (1991). Structure, function and properties of antibody binding sites. *J Mol Biol* **217**, 133-51.
121. Jeffrey, P. D., Strong, R. K., Sieker, L. C., Chang, C. Y., Campbell, R. L., Petsko, G. A., Haber, E., Margolies, M. N. & Sheriff, S. (1993). 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc Natl Acad Sci U S A* **90**, 10310-4.
122. Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B. L. & Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nat Biotechnol* **25**, 563-5.
123. Wark, K. L. & Hudson, P. J. (2006). Latest technologies for the enhancement of antibody affinity. *Adv Drug Deliv Rev* **58**, 657-70.
124. Worn, A. & Pluckthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* **305**, 989-1010.
125. Chothia, C., Gelfand, I. & Kister, A. (1998). Structural determinants in the sequences of immunoglobulin variable domain. *J Mol Biol* **278**, 457-79.
126. Waters, M. L. (2002). Aromatic interactions in model systems. *Curr Opin Chem Biol* **6**, 736-41.
127. Carver, F. J., Hunter, C. A., Jones, P. S., Livingstone, D. J., McCabe, J. F., Seward, E. M., Tiger, P. & Spey, S. E. (2001). Quantitative measurements of edge-to-face aromatic interactions by using chemical double-mutant cycles. *Chemistry* **7**, 4854-62.
128. Black, S. D. & Mould, D. R. (1991). Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal Biochem* **193**, 72-82.
129. Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Adv Protein Chem* **14**, 1-63.
130. Shea, J. E. & Brooks, C. L., 3rd. (2001). From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu Rev Phys Chem* **52**, 499-535.
131. Daugherty, P. S., Olsen, M. J., Iverson, B. L. & Georgiou, G. (1999). Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. *Protein Eng* **12**, 613-21.
132. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-8.

Chapter 2: A synthetic antibody library design with a bias towards peptide binders

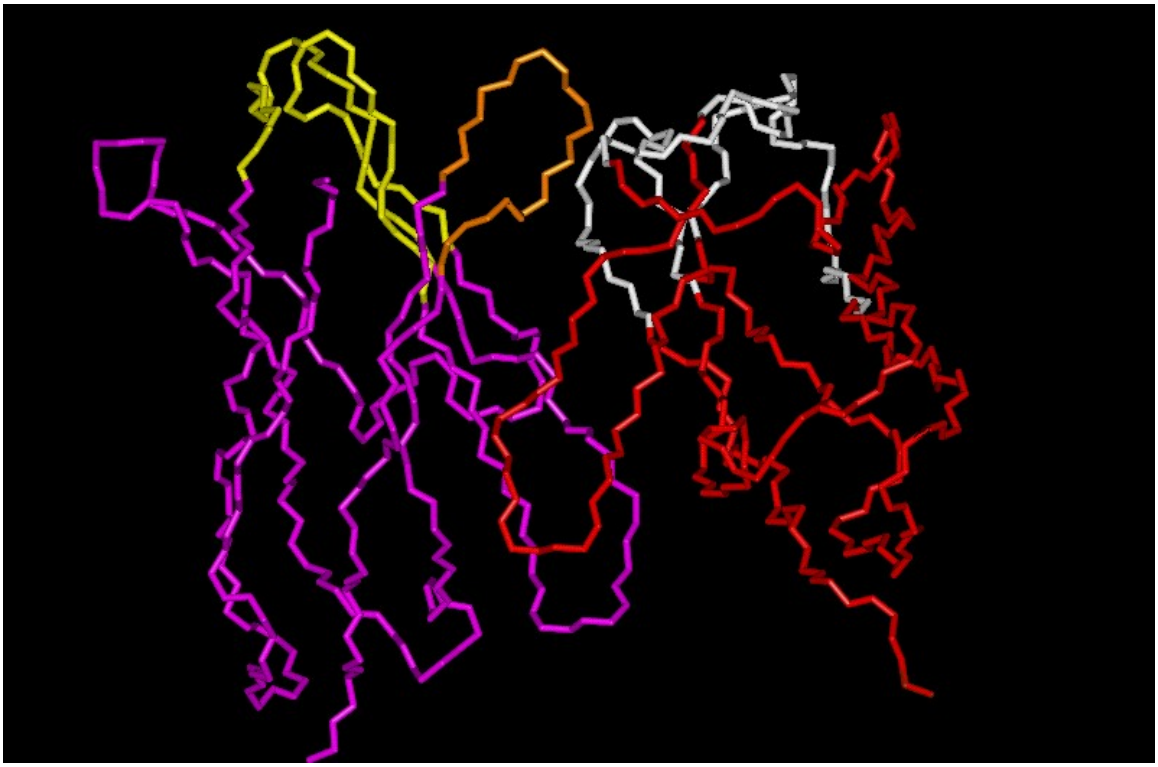
All numbering used in this chapter will use Chothia numbering unless otherwise noted[1].

INTRODUCTION

The discovery of antibodies for therapeutic, diagnostic, analytical or affinity chromatography purposes relies increasingly on the use of large combinatorial libraries encoding repertoires of antibody variants. Antibodies that recognize and bind to a target with the requisite affinity and specificity are isolated from a library by high throughput screening techniques such as phage, bacterial, yeast or ribosome display[2; 3; 4]. Three distinct types of antibody libraries can be employed: libraries that sample the antibody diversity available in immunized animals, so called "naïve repertoire" libraries that mimic the diversity of the antibodies in non-immunized individuals and, finally, synthetic libraries. In the latter, the sequence of one or more antibody scaffolds is diversified so that it encodes all or a subset of the 20 natural amino acids at multiple positions known to be critical for antigen recognition. The theoretical diversity of synthetic antibody libraries in which multiple complementarity determining region (CDR) loops are randomized far exceeds practical limitations in library construction and screening.

In antibodies, the variable heavy (V_H) and variable light (V_L) chains combine to form a symmetrical structure that contacts the antigen *via* residues in the six hypervariable loops (HV; three from each chain)[5]. Each variable domain contains four framework regions (FRs) that form a β-sheet scaffold which orients the hypervariable (HV) loops

(which are also known as CDRs) towards the solvent, forming the antigen-binding site (Figure 2-1).



Amzel and Poljak, Annu Rev Biochem. 1979;48:961-97

Figure 2-1: scFv structure with the CDR loops
Each of the six CDRs is contained in one of the six hypervariable loops. The V_H CDRs are colored yellow (loops 1 & 2) and orange (loop 3) the V_L CDRs are light grey. The V_H FRs are colored purple, the V_L FRs are colored red.

Incorporating information on how the natural immune repertoire has evolved to recognize foreign antigens can facilitate the generation of more useful synthetic antibody libraries[6]. In particular, the growing databases of antibody structures[7; 8; 9; 10; 11] allow for a comprehensive correlation of antibody binding site topography with the various classes of recognized antigens. Binding site topography, in turn, is derived largely from HV loop lengths and canonical conformations[9; 10; 11]. The canonical conformations that contribute to the binding site topography are determined by conserved amino acid residues in the

HV loops and FRs[1; 12]. The geometrical features of the antigen-binding site can be correlated with the type of antigen recognized[11]. Concave, grooved, or barrel-shaped binding sites are typically found in antibodies that bind to small molecules while flattened antigen-binding surfaces are involved in protein recognition. The binding site topographies of peptide binders lie between these two[10] (Figure 2-2). Persson *et al.*[13] recently reported the construction of a synthetic library biased toward the isolation of antibodies against small molecules by incorporating the structural and sequence features known to occur in many anti-hapten antibodies. However, to date, there are no reports describing the creation of libraries tailored to the isolation of anti-peptide antibodies.

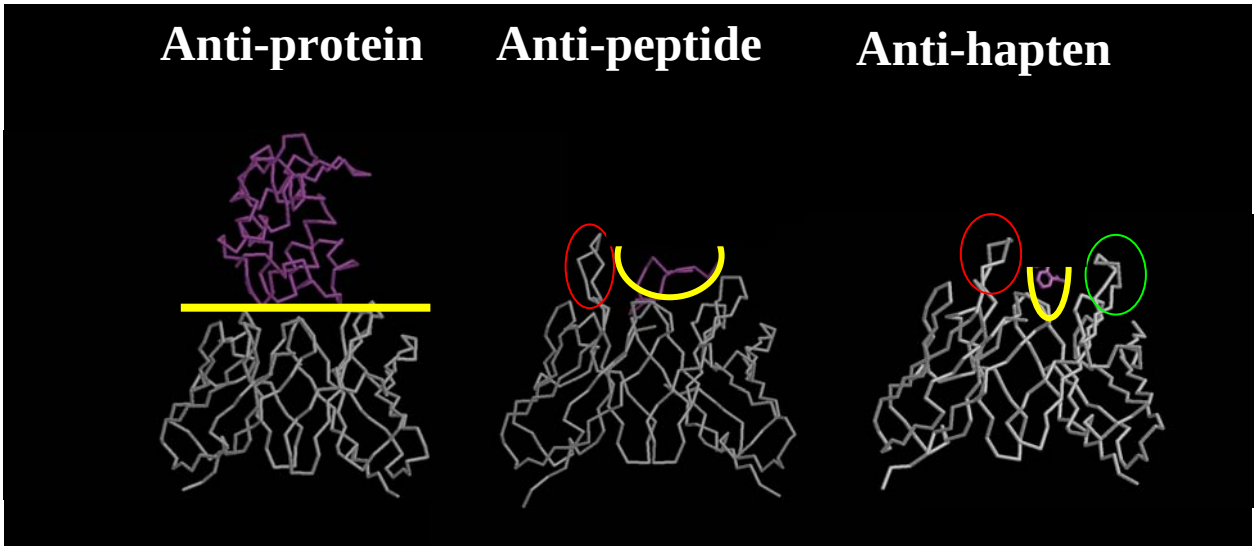
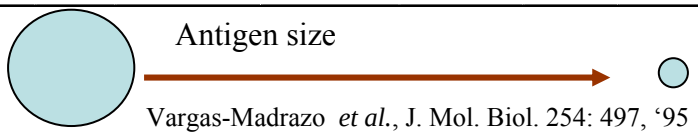


Figure 2-2: Gross topographical features of the three major antigen classes of antibodies.
These models show the gross topographical features of the antigen binding site antibodies that bind to three different classes of antigens. HV loops one of the V_H (circled red on the anti-peptide and anti-hapten antibodies) and loop two of the V_L (circled green on the anti-hapten antibody) contribute the topographical differences. (figure courtesy of Juan Carlos Almagro). Antigens are displayed in purple.

Through analyses of the known structures and sequences of antibodies, a limited number of main chain conformations, a.k.a. *canonical structures*, for five of the six HV

loops, denoted H1 and H2 for V_H, and L1, L2, and L3 for V_L have been described[1; 12]. Each of the canonical loops structures is defined by loop length and invariant residues at specific positions throughout the variable region[1; 12]. Even though the sequences of different antibodies vary, the conformations of the loops will not vary if they belong to the same canonical class. Therefore, if two different antibodies contain HV loops with matching canonical structures, the binding sites of those antibodies will have the same gross structure due to conserved interactions between amino acids at specific sites in the loops and FRs[5]. Differences in loop sequence composition merely modify the surface provided by side chains on the canonical main chain structure without disrupting the topography of the antigen-binding site. The modification of the surface contributes greatly to the specificity of an antibody. Analyses of the three-dimensional structures of antibodies and an even larger number of sequences have established a relationship between the structure and sequence[10; 14] from which canonical structures for five of the six HV loops can be reliably predicted based solely on the amino acid sequence[15](Figure 2-3). The assignment of classes to each of the loops enables a reliable rendering of the binding site topography.

Canonical Structure class H1-H2-L1-L2-L3	Frequency (%)	Protein (169)	Surface Antigen (22)	Polysaccharide (17)	Nucleic acid (42)	**Peptide (19)**	Hapten (112)
1-1-2-1-1	3.2	56[d]	0	0	25	**0**	19
1-1-4-1-1	3.7	10	0	33	13	**0**	44
1-2-1-1-1	7.1	5	5	80	5	**0**	4
1-2-2-1-1	24.5	16	44	0	18	**0**	23
1-2-3-1-1	2.9	57	43	0	0	**0**	0
1-2-4-1-1	14.2	11	4	5	24	**52**	4
1-3-2-1-1	7.9	15	26	0	31	**20**	8
1-3-4-1-1	10	43	0	14	11	**25**	6
1-4-3-1-1	6.8	11	0	0	0	**0**	89
1-4-4-1-1	6.6	4	0	0	41	**0**	55
others	13.1	26	17	17	6	**21**	13



Vargas-Madrazo *et al.*, J. Mol. Biol. 254: 497, '95

Figure 2-3. Canonical structure patterns correspond to gross specificities
The most common canonical structure patterns of antibody HV regions and their frequencies in the Kabat database are listed. The frequency that antibodies containing each of these canonical structure patterns bind to the specified major classes of antigens is also listed. 52% of the antibodies with a canonical structure pattern of 1-2-4-1-1 bind to peptides.

An extrapolation from the canonical structure paradigm demonstrated that certain combinations of canonical structure patterns, *i.e.* binding site topographies from sequences, correlate with types of recognized antigens[10] (Figure 2-3). Several canonical structure patterns correlate with antibodies that are specific for peptides. Analyses of the antigen-contacting residues, also called specificity-determining residues (SDRs)[16], in antibodies of known structure indicate that anti-protein antibodies have a large SDR surface, in contrast to a small SDR surface of anti-hapten antibodies. Anti-peptide antibodies have an SDR surface of a size in between anti-protein and anti-hapten antibodies, with SDRs of high usage located in the interior of the antigen-binding site but not buried in the V_L:V_H interface as in anti-hapten antibodies. These patterns of SDR

usage have successfully been employed[13; 17] to create an antibody library with focused diversity for improved recognition of proteins and haptens.

Antibodies isolated from libraries against peptide mimics of protein epitopes, referred to as *mimotopes*, are often cross-reactive to the parent protein. In addition to mimotope targets, peptides are attractive therapeutic and diagnostic targets due to their varied physiological roles, from signaling molecules as hormones or neurotransmitters to secreted toxins[23; 24]. Due to the conformational flexibility of peptides and the loss of entropy that occurs when they are bound by antibodies, the generation of high affinity antibodies to peptides can be challenging[12]. The availability of the validated scFv libraries biased towards peptide binders reported here is likely to facilitate the isolation of high affinity antibodies to peptides of biochemical or therapeutic interest.

Herein, a novel strategy is introduced to generate synthetic antibody libraries intended to be peptide-specific. By leveraging the canonical structural features of anti-peptide antibodies, it was hypothesized that libraries could be designed with a predisposition to bind peptides. The following sections describe the construction of two libraries based, respectively, on a human (Hu:anti-pep) and on a mouse (M:anti-pep) scaffold antibody. The scaffold of the Hu:anti-pep library is derived from a well-characterized germline V_H-V_L fusion with loop length modifications to the V_L. The diversity is limited at low-usage SDRs and unlimited at high-usage SDRs[17]. The M:anti-pep library utilizes the well-expressed anti-digoxin 26-10 antibody scaffold that contains canonical loop structures that are similar to other anti-peptide antibodies[10]. The M:anti-pep library contains limited diversity at potential SDRs. Following screening by phage display, the M:anti-pep

library yielded antibodies to two different peptides having dissociation constants below 20 nM.

RESULTS

Design and construction of combinatorial M:anti-pep library

Several commonly used, stable, and well-expressed antibodies were evaluated as potential scaffolds for the construction of a library biased towards peptide binders. Features such as canonical structure patterns, loop lengths, SDRs and sequence composition were examined. Vargas-Madrazo *et al*. noted that 52% of the antibodies with the canonical structure pattern of 1-2-4-1-1 (H1-H2-L1-L2-L3) are peptide binders[10]. The canonical structure designations rely heavily on the HV loop lengths. In addition, the dominant HV loop lengths of 46 anti-peptide antibodies in the Kabat database are as follows[9]: L1: 11 (35%) or 16 (55%); L2: 7 (100%); L3: 8 (9.5%) or 9 (88%); H1: 5 (85%); H2: 17 (72%); H3: 7 (22%), 8 (13%), or 12 (13%).

The mouse anti-hapten antibody, 26-10, which binds to the cardiac glucoside digoxin with a nanomolar dissociation constant[25], displays the 1-2-4-1-1 canonical structure pattern typically found in anti-peptide antibodies. Also, the loop lengths of 26-10 conform to those of anti-peptide antibodies in five of the six loops: L1: 16; L2: 7; L3: 9; H1: 5; H2: 17 (Table 2-1). The 26-10 scFv is well-expressed in bacteria, highly soluble, stable, and displays a minimal propensity to form dimers or higher molecular weight aggregates in solution, making it well-suited for library selection experiments. For these reasons, even though the 26-10 antibody was elicited though immunization with the heart

glycoside digoxin, it was selected as the scaffold for construction of an anti-peptide library.

	L1	L2	L3	H1	H2	H3
46 anti-peptide Abs*	11, 16	7	8, 9	5	17	7, 8, 12
26-10	16	7	9	5	17	10
3-23/J_H4-A27md/Jκ3	17	7	9	5	17	6-8

*(Collis *et al.*, 2003)

Table 2-1. HV loop length of 46 anti-peptide antibodies and both library scaffolds
The HV loops lengths of 46 anti-peptide antibodies, the 26-10 Ab, and the Hu:anti-pep germline-based scaffold are shown.

An analysis of the contact residues in eight antibodies that bind to small polypeptides[11] and 18 anti-peptide antibodies[6] was employed to identify specificity determining residues (SDRs) that occur preferentially in anti-peptide antibodies (Table 2-2). All 11 SDRs in the co-crystal of 26-10 with digoxin[26] (Figure 2-4) correspond to SDRs found in anti-peptide antibodies in the aforementioned studies, making these sites candidates for sequence diversification. Positions with solvent-exposed side chains in 26-10 corresponding to SDRs in anti-peptide antibodies were also diversified. To simplify the construction of the library, termed the M:anti-pep library, diversity was introduced only in the V_H. No positions in the V_L were diversified (Figure 2-5).

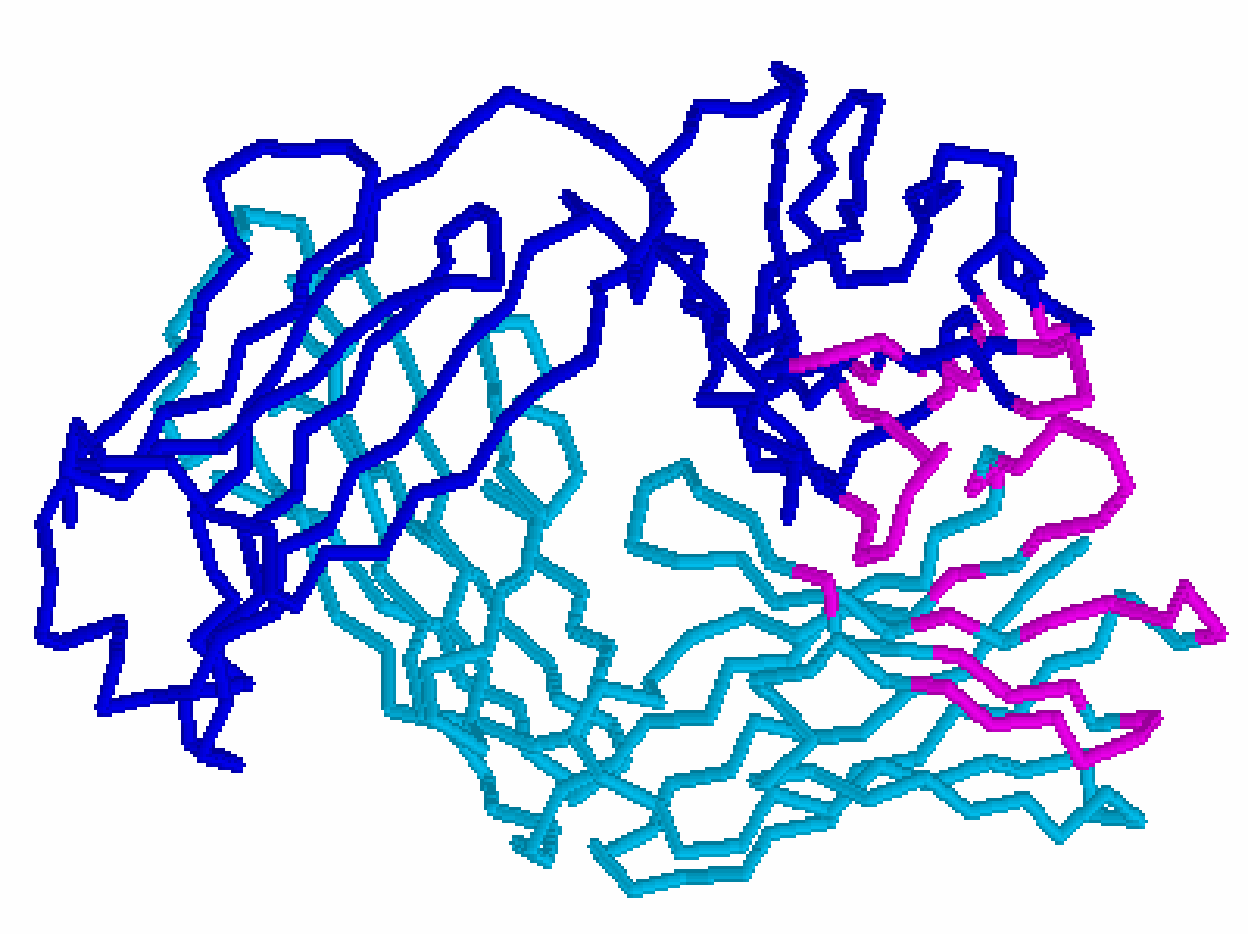


Figure 2-4. SDRs of 26-10.
A model of the 26-10 antibody showing the positions of the SDRs (pink) of the heavy (dark blue) and light chains (light blue).

	L1	L2	L3	H1	H2	H3
Position-number of SDRs in set of 24 antibodies (**26-10 SDRs in bold**)	32-20, 30a-15, 30c-12, 30f-5, 30e-3, 30b-8, 27-2, 36-2	50-4, 46-3, 49-3	**91**-24, **94**-20, **96**-17, 92-12, 93-11, 89-2	**33**-16, 31-9, **35**-6, 32-3, 30-3	**47**-10 **50**-20, 52-18, 56-17, 58-17, 53-13, 54-8, 52a-8, 55-7, 57-4	**95**-20, 96-10, 97-8, 98-5, **100**-4, **100b**-4, 92-4, 93-4, 99-3, **100a**-3, 100c-3, 100d-3, 100e-3, 100f-2, 101-2

Table 2-2. SDRs of 24 anti-peptide compared to SDRs of 26-10.
The SDRs of all the antibodies have been structurally determined. The position of each SDR is listed followed by the number of antibodies that use that position as an SDR. The SDRs of 26-10 are shown in bold.

The positional frequencies of amino acids at all of the SDRs located in H1 and H2 of anti-peptide antibodies were determined by aligning 103 heavy chains available in the Kabat database[27]. Amino acid frequencies for each SDR position were determined (Table 2-3). Subsequently, degenerate codons were designed for each SDR of H1 and H2 with an attempt to encode for the natural diversity of amino acids found at each SDR position. In some cases, two or more degenerate codons, and therefore two or more oligonucleotides, were necessary in order to encode most of the expected amino acid diversity (Tables 2-3) while eliminating termination codons and cysteines.

Position	Target diversity Prevalence in natural anti-peptide Abs (%)	Codon	Residues encoded/extra residues	Coverage(%)
H29	F-80 L-10 I-9	HTT	FLI	100
H30	T-57 S-29 K-7 N-3 P-1	AVK	TSNK/R	99
H31	S-35 D-23 T-10 E-9 G-8 N-8 R-4 I-1	RVK	SDTEGNR/KA	99
H32	Y-71 N-8 T-7 F-4 S-4 G-2 A-1 C-1	WHT	FSYTN/I	96
H33	A-21 G-21 W-20-Y-16 T-9	DSG	AGWTS/R	74
	N-6 F-3 I-1 P-1 S-1	WWT	YNFI	26
H35	H-43 N-22 S-21 D-2 E-2 W-2 Y-2 P-1 K-1 L-1 C-1 A-1	HMT	HNSYP/T	90
H47	W-99 Y-1	TGG	WY	100
		TAC		
H50	R-16 G-15 T-13 N-12 E-11	RNK	RGTNEVAS/IMKD	87
	V-10 W-8 Y-5 A-4 S-2	TGG	W	8
		TAT	Y	5
H51	I-94 F-3 S-1 V-1	WYT	IFS/T	99
H52A	P-44 S-14 T-11 L-6 N-6 F-2 Y-1 G-1	HMT	FSLPT/I	91
H53	G-23 Y-21 N-20 A-13 S-12	TAT	Y	21
	K-6 D-3 I-1	RVN	NKSADG/TER	78
54	N-39 G-30 S-13 Y-7 D-6 T-4	RRT	GSDN	89
55	G-58 S-13 Y-11 T-5 V-3 A-2 D-2-E-1-H-1 I-1 K-1 M-1 F-1	RBT	ITSVAG	82
		THT	FSY	12
56	S-20 G-15 Y-14 N-12 D-11	DMT	SYTNAD	72
	A-10 T-5 V-4 E-4 R-3 C-1 Q-1	RKA	VRG/M	22
57	T-80 P-7 V-3 S-2 A-2 I-1 K-1 L-1 Y-1	NYA	LSPITVA	98
58	Y-26 N-23 S-13 K-10 T-8	THT	FSY	44
	H-7 F-4 D-4 E-2 A-1	VMN	TAHNKDE/PQ	56

Table 2-3. M:anti-pep library target diversity of H1, H2.
Positions in bold correspond to the SDRs of 26-10. Positions in italics correspond to the positions of SDRs from 24 anti-peptide antibody co-structures that have side chains which are solvent exposed in 26-10. See Experimental procedures for the degenerate codes.

The length, sequence, and structural variation of the H3 loop implies its importance in antigen recognition and binding, supported by studies that show this loop contacts the antigen more than the other loops and that all 20 amino acids occur in this loop in the natural repertoire of antibodies[9; 28]. The length of the H3 loop was allowed to vary (n=8, 10, and 12) to reflect the length diversity of other anti-peptide antibodies (Table 2-4). In H3, a consensus pattern for the four amino acid positions just prior to position 103 of Ala-Met-Asp-Tyr is often observed in natural antibodies[29]. Similarly, Ala and Arg/Gly are conserved at positions 93 and 94, respectively[30; 31]. These consensus sequence motifs are maintained in the H3 library design. Between four and nine positions of H3 were diversified in 26-10, depending on the length variation introduced (Table 2-4). Primers of appropriate lengths were designed using the NNS coding scheme (S= G or C), which encodes for all 20 amino acids and also one stop codon (TAG). It was hoped that allowing all 20 amino acids in the randomized positions of this H3 loop of the library will result in higher affinity clones[29].

			CDR-H3 positions 95-102												
Oligo	93	94	95	96	97	98	99	100	100a	100b	100c	100d	101	102	Length
NNS-10a	A	*SGT*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	Y	A	M			D	Y	10
NNS-10b	A	*SGT*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*DSG*	A	M			D	Y	10
NNS-12a	A	*SGT*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	Y	A	M	D	Y	12
NNS-12b	A	*SGT*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*NNS*	*DSG*	A	M	D	Y	12
NNS-8a	A	*SGT*	*NNS*	*NNS*	*NNS*	Y	A	M					D	Y	8
NNS-8b	A	*SGT*	*NNS*	*NNS*	*NNS*	*DSG*	A	M					D	Y	8

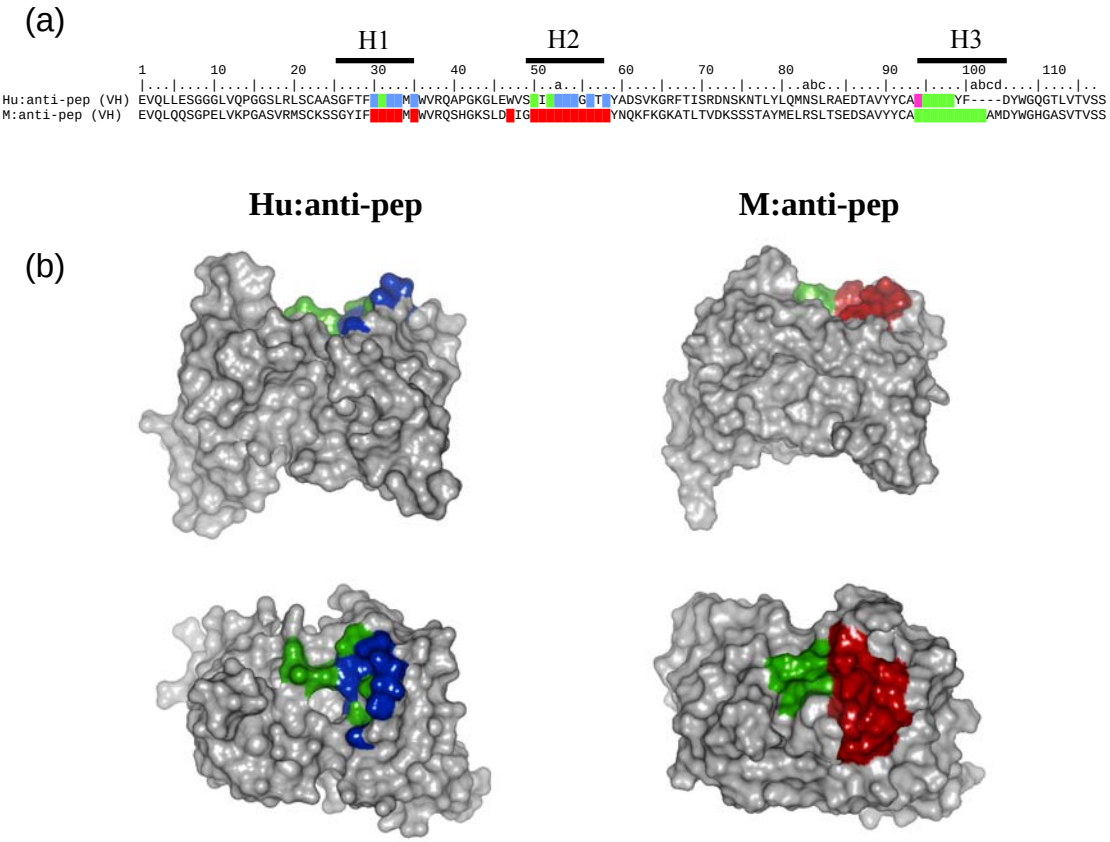
Table 2-4: H3 Diversity.
Sub-libraries encoding H3 regions of different lengths were constructed by introducing diversity at the positions shown in italics. The codon scheme used at those positions is indicated. NNS encodes all 20 amino acids; DSG encodes W/S/A/G/T/R.

Following gene construction and ligation into pAK200 vector[32], a total of 5.1×10^8 independent transformants were obtained. 12 clones, picked at random, were sequenced

and found to contain inserts with the expected diversity. The length diversity of loop H3 was also well represented (n=8, 5 out of 12; n=10, 3 out of 12; n=12, 4 out of 12).

Design and construction of combinatorial Hu:anti-pep library

The design and construction of the Hu:anti-pep library has been described[17]. Similar to the M:anti-pep library discussed above, all diversified SDRs in the Hu:anti-pep library are located in the V_H domain, while the V_L domain is kept constant (Figure 2-5). In brief, the 3-23 V_H human germline gene recombined with the J_H4 minigene, which is well expressed in *E. coli*[33], was used as the V_H scaffold. Diversity was generated at a consensus of anti-protein and anti-peptide SDRs. Unlike the M:anti-pep library, in the Hu:anti-pep library, SDRs of high usage were varied with all 20 amino acids whereas SDRs that make contact less frequently were limited to Tyr, Asp, Ala, or Ser (by using oligonucleotides with the degenerate codon KMY[17]). Tyr and Asp play a dominant role in antigen recognition, while the Ala and Ser residues allow for space and conformation flexibility, functioning in an auxiliary role to facilitate contacts between Tyr or Asp and the antigen[34; 35]. Position 93 of H3 was fixed to an Ala whereas position 94 was allowed to encode the more typical germline amino acid Arg[36] and the amino acid found in 3-23/J_H4[17], Lys. In summary, the H3 apical region (positions 95, 96-98) as well as positions 50 and 52 of H3 and position 31 of H1 encoded all 20 amino acids (NNK codon scheme) whereas positions 30, 32, 33, and 35 of H1, 52a-54, 56, and 58 of H2 encoded Tyr, Asp, Ala, or Ser (KMY codon scheme). The length of the H3 loop was also diversified (n= 6 or 8) to reflect partially the type of length diversity seen in the H3 loop of other anti-peptide antibodies.



Green: full diversity. Blue: YDAS. Red: Tailored diversity (Table 2-3)

Figure 2-5. Structure of the Hu:anti-pep and M:anti-pep VH libraries.
(a) Sequence of the VH chains showing randomized positions. (b) Structures of 3-23/J_H-4 and 26-10 V_H domains that were used respectively as the scaffold for Hu:anti-pep and M:anti-pep libraries showing the positions that were diversified. Green: fully randomized positions incorporating all 20 amino acids; Blue positions diversified to Tyr, Asp, Ala, and Ser; Violet: positions diversified to Arg and Lys; Red positions diversified to incorporate the most common amino acids at each position found in an alignment of anti-peptide antibodies in the Kabat database (see Table 2-4 for details).

The V_H repertoire was cloned with a synthetic V_L constructed by modifying the human germline V_κ gene A27, a highly promiscuous light chain, capable of generating antigen-binding sites that accommodate a wide diversity of antigens[37] that is also easily overexpressed *in vivo*[38; 39; 40]. A27 codes for flat antigen-binding sites and displays a type

6 canonical structure in L1[41]. To modify it to encode a grooved antigen-binding site, the L1 loop from the germline B3 gene, which is a canonical structure type 3 loop (the longest canonical structure of L1), was grafted into A27 (Figure 2-6). The hybrid A27md was fused to the Jκ1 region to create the complete V_L scaffold. The scaffold gene assumes a 1-2-3-1-1 canonical structure pattern for HV loops H1-H2-L1-L2-L3, which is more commonly associated with protein binders[42]. The V_H HV loops have the anti-peptide pattern of 1-2 while V_L HV loops lack the quintessential 4-1-1 observed for peptide binders due to an extra amino acid in L1. The library was cloned into the pHEN vector[38], resulting in 2×10^8 independent transformants. A total of 20 clones were picked at random and sequenced. All clones contained the insert and intended diversity with two clones each containing one stop codon.

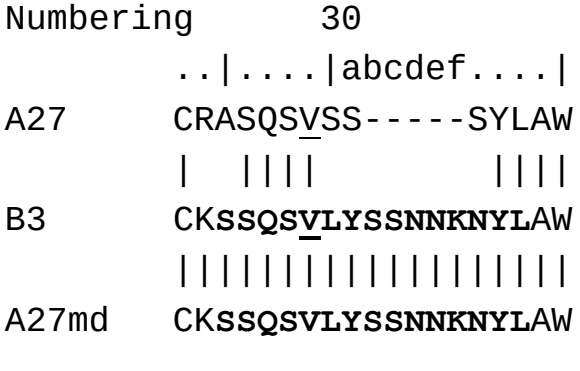


Figure 2-6. Sequence of L1 in A27mdJκ1
The HV loop L1 of germline V_L gene B3, with a long canonical loop structure, was grafted onto germline V_κ gene A27. B3 and A27 are highly homologous in L1[43] with the exception of the additional length in B3. L1 of both genes A27 and B3 share a key canonical structural determining residue (V29, underlined in B3). Other sequence identities are indicated with vertical bars.

Selections of scFvs with 2 different peptide targets

The M:anti-pep and the Hu:anti-pep libraries were screened by phage display against the peptides human angiotensin (AT, 10 amino acids, M.W. 1296 Da) and human neuropeptide Y (NPY, 36 amino acids, M.W. 4271 Da). Both peptides were synthesized with a biotin moiety at the *N*-terminus. Three rounds of panning were performed in each

case, except for the M:anti-pep library, which was panned against NPY for five rounds. Enrichment was monitored by polyclonal phage enzyme-linked immunosorbent assay (ELISA) (Figure 2-7).

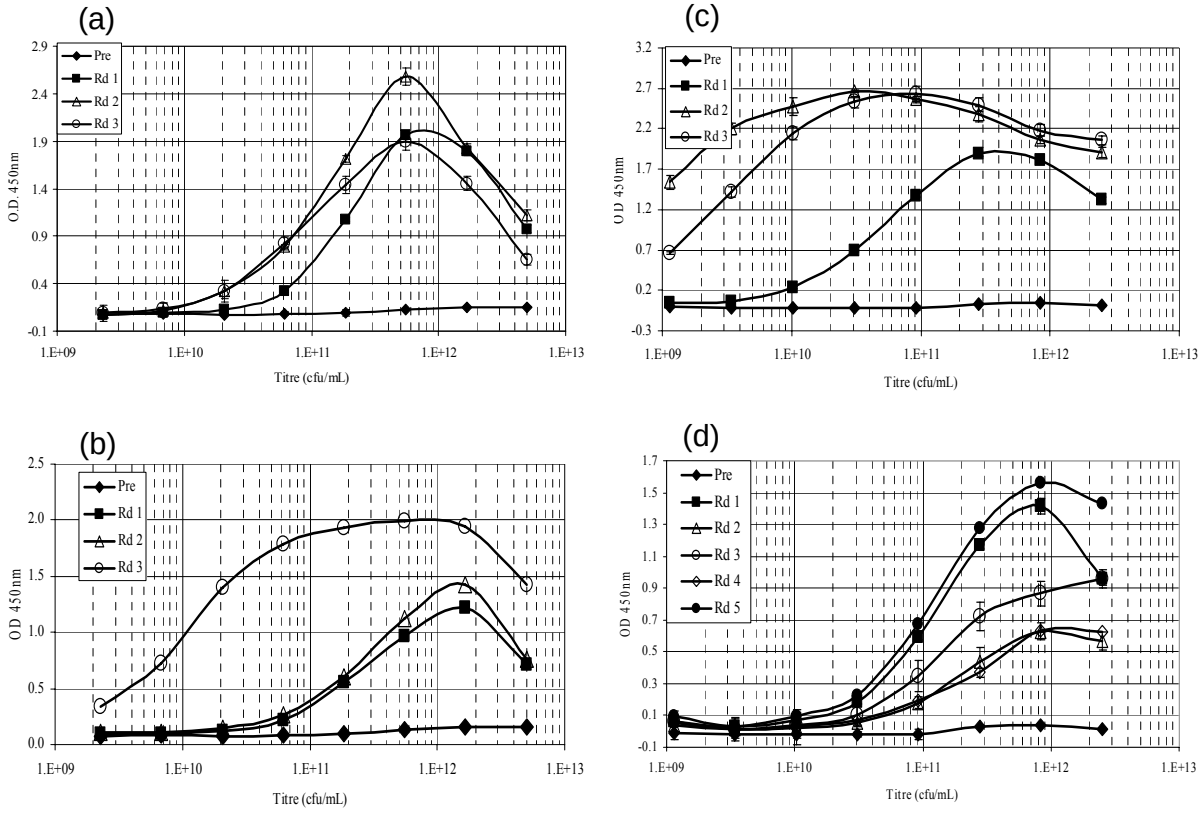


Figure 2-7. Polyclonal phage ELISAs.
Polyclonal phage from the pre-pan through the third or fifth rounds of screening with Hu:anti-pep (a & b) or M:anti-pep (c & d) libraries against AT (a or c) and NPY (b or d).

Individual clones from the last rounds of selection were randomly chosen for additional screening by monoclonal phage ELISA. Specifically, phage antibodies from 96 individual clones selected at random from the last round were assayed for binding to either the selection peptide or an unrelated peptide as a control (Figure 2-8). Following three rounds of selection, 71% of the screened clones from the Hu:anti-pep library panned against NPY showed specificity to NPY, while 87% of the clones from the

Hu:anti-pep panned against AT showed specificity to AT. For the M:anti-pep library, 79% of the clones panned against AT showed specific binding after 2 rounds and 96% after 3 rounds. A total of 99% of the clones from the fifth round of selection of M:anti-pep library against NPY showed specific binding.

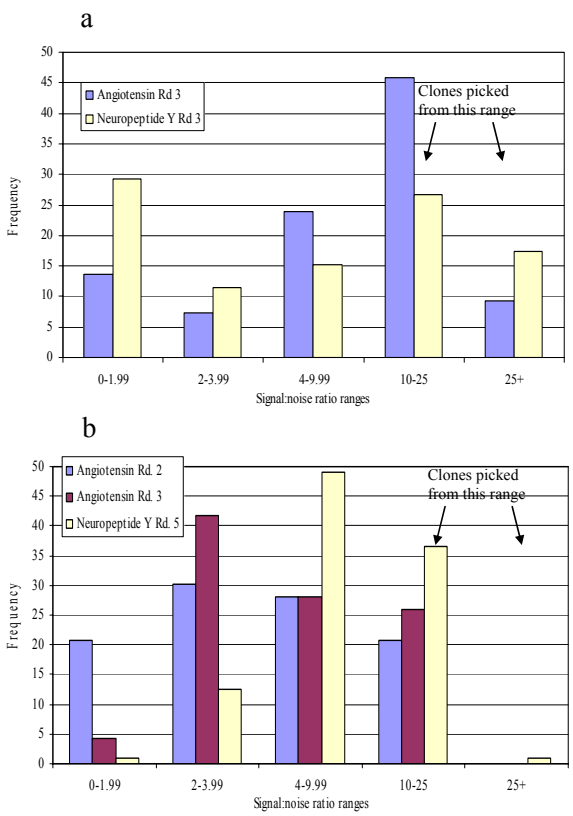


Figure 2-8. Monoclonal phage ELISA screening.
Signal-to-noise ratios obtained in monoclonal phage ELISA assays with AT or NPY with clones isolated from: (a) Hu:anti-pep and (b) M:anti-pep libraries. Results are presented as the frequency of clones exhibiting a certain signal-to-noise range. To determine the signal-to-noise ratio, the binding signals of phage antibodies were determined against both the target peptide and the unrelated peptide. Clones were isolated from the range with signal-to-noise ratios greater than 10 for additional analysis

.

Sequence analysis

Clones exhibiting a signal-to-noise ratio greater than 10 by monoclonal phage ELISAs were sequenced (Tables 2-5 and 2-6). 15 clones were sequenced from both the AT and NPY panning of the Hu:anti-pep library revealing two distinct clones for each peptide (Table 2-5). For the M:anti-pep library, after two rounds of screening against AT, nine of 12 clones were distinct. But after the third round, sequencing of twelve clones yielded only two distinct sequences, one of which, M:ATB5, was already present in round two. Interestingly, 20 of the 24 total M:anti-pep clones specific for AT contain the same H2 DNA sequence as the parental 26-10 scFv. In contrast, the round 5 phage population of M:anti-pep library screened with NPY exhibited a high degree of diversity and contained 13/15 distinct clones (Table 2-6). Not surprisingly, the sequences of all the clones isolated from the M:anti-pep library conformed to the 1-2-4-1-1 canonical structure pattern while all the clones isolated from the Hu:anti-pep library conformed to the 1-2-3-1-1 canonical structure pattern.

(a)

CLONE	CDRH1						CDRH2												CDRH3					H3 length	Frequency
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58		94	95	96	97	98		
3-23/J_H4	F	S	S	Y	A	S	W	A	I	S	G	S	G	G	S	T	Y							12	NA
Hu:AT1	.	D	L	S	S	Y	.	T	.	A	Y	A	S	.	A	.	.		R	R	K	D	I	8	13/15
Hu:AT2	.	S	T	A	D	A	.	I	.	G	S	.	D	.	Y	.	S		K	N	G	-	-	6	2/15

(b)

CLONE	CDRH1						CDRH2												CDRH3					H3 length	Frequency
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58		94	95	96	97	98		
3-23/J_H4	F	S	S	Y	A	S	W	A	I	S	G	S	G	G	S	T	Y							12	NA
Hu:NPY1	.	D	.	A	D	Y	.	W	.	.	A	Y	S	.	.	.	D		K	R	S	S	H	8	6/15
Hu:NPY2	.	A	W	A	D	Y	.	Y	.	G	A	Y	A	.	.	.	S		K	R	T	T	G	8	9/15

Table 2-5. Alignment of clones isolated from the Hu:anti-pep library:
Alignment of the randomized positions in clones isolated after three rounds of panning of the Hu:anti-pep library against (a) AT or (b) NPY. The sequence of the scaffold V_H gene, 3-23/J_H4 is listed at the top. The H3 sequence of the scaffold gene exists only in randomized form and therefore is not included.

(a)

CLONE	CDRH1						CDRH2												CDRH3											H3 length	Frequency
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58		94	95	96	97	98	99	100	100a	100b	100c	100d		
26-10	F	T	D	F	Y	N	Y	Y	I	S	P	Y	S	G	V	T	G		G	S	S	G	N	K	W	A	M	-	-	10	NA
Rd. 2																															
M:ATB5	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.		.	G	T	.	D	.	A	.	.	-	-	10	3/12(13/24)†
M:ATC10	.	N	A	.	S	T	.	.	.	.	.	.	.	.	.	.	.		R	I	G	W	G	A	M	-	-	-	-	8	1/12
M:ATC11	I	S	T	I	F	H	W	A	S	R	Y	E	G	.	D	V	Y		.	N	V	Q	G	T	G	.	.	-	-	10	1/12
M:ATC12	.	.	.	T	F	S	.	.	.	.	.	.	.	.	.	.	.		.	.	G	V	T	A	M	-	-	-	-	8	1/12
M:ATE3	.	N	.	I	S	H	.	G	F	R	N	.	D	S	G	I	Y		.	Q	G	L	L	Q	K	T	G	A	M	12	1/12
M:ATE5	L	.	N	Y	W	S	.	.	.	.	.	.	.	.	.	.	.		.	.	G	K	T	A	M	-	-	-	-	8	2/12
M:ATB4	L	.	A	T	F	H	.	.	.	.	.	.	.	.	.	.	.		.	Q	G	V	T	A	M	-	-	-	-	8	1/12
M:ATC6	L	.	A	T	W	H	W	T	S	R	Y	.	D	F	G	T	D		.	*	N	.	Y	A	M	-	-	-	-	8	1/12
Rd. 3																															
M:ATB5	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.		.	G	T	.	D	.	A	.	.	-	-	10	3/12(13/24)†
M:ATG11	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.		.	.	G	V	T	A	M	-	-	-	-	8	2/12

(b)

CLONE	CDRH1						CDRH2												CDRH3											H3 length	Frequency
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58		94	95	96	97	98	99	100	100a	100b	100c	100d		
26-10	F	T	D	F	Y	N	Y	Y	I	S	P	Y	S	G	V	T	G		G	S	S	G	N	K	W	A	M	-	-	10	NA
Rd. 5																															
M:NPYA2	I	.	G	I	W	P	W	G	S	Q	.	G	G	V	N	I	Y		R	.	I	P	S	V	L	P	T	A	M	12	1/15
M:NPYA3	.	.	K	T	A	S	.	W	S	W	.	.	D	F	G	V	Y		R	T	G	V	M	S	.	.	.	-	-	10	1/15
M:NPYA6	.	S	T	.	W	Y	W	A	S	.	H	G	D	T	Y	.	S		.	L	E	L	Y	A	M	-	-	-	-	8	1/15
M:NPYA9	L	N	.	Y	T	P	W	G	S	R	S	A	.	I	D	I	S		R	D	H	V	Y	A	M	-	-	-	-	8	1/15
M:NPYD8	I	S	K	S	W	H	W	M	T	R	H	.	G	F	G	L	N		.	V	N	Y	A	E	Y	.	.	-	-	10	1/15
M:NPYE3	.	.	R	S	I	T	W	G	F	.	S	A	D	V	A	V	S		R	L	L	M	S	A	M	-	-	-	-	8	1/15
M:NPYE5	I	.	N	S	I	.	W	T	.	R	H	.	.	F	G	V	Q		R	M	A	Q	V	C	R	.	.	-	-	10	1/15
M:NPYE6	L	R	.	Y	G	Y	W	.	T	Y	H	.	D	F	G	A	H		.	C	Q	A	Y	A	M	-	-	-	-	8	1/15
M:NPYG4	I	.	T	N	I	P	W	T	F	R	Y	S	D	V	Y	I	S		R	P	A	M	W	D	Y	.	.	-	-	10	1/15
M:NPYG12	L	S	A	T	W	H	.	A	F	R	S	.	G	S	.	.	Y		.	Q	P	P	S	A	V	G	R	A	M	12	1/15
M:NPYH1	I	.	A	I	A	.	W	R	F	R	Y	T	.	T	Y	V	Y		.	I	K	E	Y	A	M	-	-	-	-	8	1/15
M:NPYH8	I	K	T	Y	W	P	W	W	S	W	H	.	D	S	G	V	Y		.	H	W	F	Q	D	P	M	R	A	M	12	1/15
M:NPYSol1	.	S	T	.	W	S	W	V	S	R	Y	G	D	I	A	A	Y		.	P	G	D	Y	A	M	-	-	-	-	8	2/15
M:NPYSol2	L	K	T	.	T	S	.	V	S	R	S	.	D	.	G	P	Y		.	L	E	W	W	A	M	-	-	-	-	8	1/15

Table 2-6. Alignment of clones isolated from the M:anti-pep library.
(a) Alignment of the randomized positions in clones isolated after two or three rounds of panning of the M:anti-pep library against AT. The sequence of the scaffold VH from 26-10 is shown at the top. * denotes stop codon. †Clone M:ATB5 appeared 3/12 and 10/12 times in rounds 2 and 3, respectively. (b) Alignment of the randomized positions in clones isolated after five rounds of panning of the M:anti-pep library against NPY.

Characterization of soluble scFvs

Based on the monoclonal ELISA sequencing results, human and mouse clones specific for each of AT and NPY were chosen for further characterization. Proteins were expressed in soluble scFv form in *E. coli* and purified by immobilized metal affinity chromatography (IMAC) followed by gel filtration to remove dimers or higher molecular weight aggregates. All nine scFv proteins expressed well with seven of nine resulting in yields of purified, monomeric protein >1 mg/L (Table 2-7). Binding specificity of each purified antibody was assayed by ELISA using either the target peptide or an unrelated peptide (data not shown). The kinetics of antigen binding were determined by surface plasmon resonance on a BIACore 3000 instrument. Briefly, the biotinylated peptides

were immobilized on a streptavidin chip at a surface density sufficient to give 30 (AT) or 75 (NPY) RUs, respectively, and the kinetics of association and dissociation were measured. The antibodies isolated from the Hu:anti-pep library in general exhibited lower affinities relative to the antibodies from the M:anti-pep (Table 2-7). Specifically, three out of four antibodies from the Hu:anti-pep library exhibited dissociation constants (K_D values) for their ligands in the micromolar range. In contrast, with the exception of the anti-NPY antibody fragment M:ATE5, five of six antibodies from the M:anti-pep library displayed K_D values in the 10-100 nM range. This is particularly noteworthy considering that the V_L domain of the M:anti-pep library had not been subjected to diversification.

(a)

Clone	k_a (1/Ms)	k_d (1/s)	K_D (nM)	$\chi 2$	Expression level (mg/L of culture)
Hu:NPY1	177	7.24×10^{-4}	4100	2.73	4.8
Hu:NPY2	9.06×10^{3}	4.40×10^{-4}	48.6	0.589	0.3
M:NPYSol1	5.06×10^{3}	3.81×10^{-4}	75.3	0.731	1.4
M:NPYSol2	1.66×10^{4}	3.03×10^{-4}	18.2	0.489	0.4

(b)

Clone	k_a (1/Ms)	k_d (1/s)	K_D (nM)	$\chi 2$	E×pression level (mg/L of culture)
Hu:AT1	1.24×10^{4}	0.0139	1130	1.55	1.6
Hu:AT2	2.08×10^{3}	0.0112	5400	3.61	1.3
M:ATG11	1.13×10^{5}	5.94×10^{-3}	52.6	0.44	3.1
M:ATE5	1.27×10^{3}	4.62×10^{-3}	3630	0.364	2.8
M:ATB5	2.56×10^{5}	3.62×10^{-3}	14.2	0.505	1.5

Table 2-7. BIACore analysis and expression yields of scFv antibody fragments specific to NPY or AT.
Freshly prepared monomeric scFv protein was used to determine the k_a and k_d by surface plasmon resonance on a BIAcore 3000 instrument using biotinylated-NPY (a) or biotinylated-AT (B) immobilized on different SA sensor chips, as described in the experimental protocols section. The expression yield was calculated from the amount of purified scFv protein as determined by a BCA plate assay.

CONCLUSIONS

Examination of the available structures of anti-peptide antibodies has revealed that they contain similar topographical features, based mostly on loop lengths and canonical structures. It is reasonable to propose that the similar topography facilitates recognition

of a general class of antigens, in this instance peptides[10; 11]. Shared sequence-based features include SDRs common to peptide binders and also the amino acid preferences displayed at these SDR positions.

The available information on the structure and sequence features of anti-peptide antibodies was exploited to construct a mouse (M:anti-pep) and a human (Hu:anti-pep) library. The anti-digoxin 26-10 scaffold of the M:anti-pep library was selected primarily based on the structural features it shares with anti-peptide antibodies. These shared structural features include: (i) peptide-specific canonical structure pattern (1-2-4-1-1) for the HV regions and (ii) loop lengths nearly identical to those of anti-peptide antibodies. Additional similarities between 26-10 and anti-peptide antibodies were observed in the SDR usage, indicating that the side chains at the putative SDRs of 26-10 are well positioned to interact with peptide antigens. 26-10 is also known to bind to peptides[44], perhaps because of its anti-peptide-like binding pocket topography and SDR side chain orientation. Other desirable features of the 26-10 antibody as a scaffold for synthetic library construction include its high expression level in *E. coli* and its thermodynamic stability[45].

The 3-23/ J_H4 V_H chain and the modified V_κ germline chain A27md segment used for the construction of the Hu:anti-pep library is also well-expressed[38]. This scaffold, without the minor modification to the V_L region, is capable of recognizing a wide variety of antigens[17; 37] and has also been validated in several other phage display pannings[13; 34; 35]. In contrast to 26-10, which is a murine antibody isolated from a hybridoma that had undergone somatic hypermutation[46], the scaffold of the Hu:anti-pep library is a fusion of germline genes that have not undergone any prior selection.

Both libraries incorporated variation only in the V_H chain. The canonical structure patterns of the V_H chains for both libraries are identical. By leaving the V_L of the M:anti-pep library intact, the library maintained the 4-1-1 anti-peptide-like pattern for the V_L. The difference between the libraries' canonical structure pattern in the V_L chain (3-1-1 for Hu:anti-pep) is subtle: CDRL1 is one amino acid longer in the 3-1-1 canonical structure compared to the 4-1-1 arrangement. While the 17 amino acid long L1 loop of the Hu:anti-pep scaffold is not typically associated with peptide binders, an examination of the canonical structures available in the PDB database indicates that the 3-1-1 binding pocket is likely concave enough to accommodate peptides.

The M:anti-pep library incorporated only the diversity observed by aligning known anti-peptide antibody sequences. Residues were targeted for diversity based on two considerations: their use as SDRs in anti-peptide antibodies and 26-10, and on surface exposure. Since the targeted diversity was based on that of existing antibodies, the number of structurally deficient clones should be lower in this library[47]. Although not buried, position H47 is occupied by Trp in 99% of reported anti-peptide antibodies, whereas it is a Tyr in 26-10. For this reason, the diversity at that position in the M:anti-pep library included only Tyr in addition to Trp. Length diversity was included in the M:anti-pep library for H3, also to reflect the range of H3 lengths of natural anti-peptide antibodies.

The number of randomized positions in the Hu:anti-pep library (17) was lower than in the M:anti-pep library (21-25, depending on H3 length). Additionally, the Hu:anti-pep had a simpler diversification scheme: all 20 amino acids were incorporated at "high use" SDRs and only 4 amino acids at "low use" SDRs. Varying the positions that correspond

to frequently used SDRs to all 20 amino acids enabled a large number of possible amino acids to be explored at the positions critical to peptide recognition. The SDR positions identified as infrequently used SDRs in other anti-peptide antibodies[6; 11] had diversity restricted to Tyr, Asp, Ala, or Ser in order to limit the total sequence space of the library. This limited diversity at the low usage SDRs has been validated for high affinity interactions with protein antigens[17; 35]. Due to the limited role the side chains at these positions play in antigen recognition, restricting the diversity to these four amino acids was anticipated to have only minimal adverse effects on the activity of library clones.

Anti-peptide antibodies to angiotensin and neuropeptide Y were isolated from each of the two libraries following phage panning in solution. AT and NPY are representative of peptides having different lengths and solution structures. Although AT is short, at only 10 amino acids, a solution structural analysis[48] reveals a persistent short β-sheet in the middle of the peptide. At 21.8 Å in length, AT is only slightly longer than digoxin, the hapten recognized by the scaffold antibody 26-10 (~19 Å). The solution structure of the larger peptide, NPY (36 amino acids, M.W. 4271 da), with a length of 51 Å[49], shows a persistent six turn α-helix at the *C*-terminus. Thus, AT is close to the size range of the 26-10 binding peptides (15 mers around 1800 da, 26mers around 2900 da) isolated by Ball *et al*. [44], whereas NPY is much longer.

Isolated clones from each library bound specifically only to their target peptide. Two high affinity clones, M:ATB5 and M:ATG11 (K_D values of 14.2 nM and 52.6 nM, respectively) were isolated from the M:anti-pep library after two (M:ATB5) and three (M:ATG11) rounds of selection against AT. Interestingly, three AT-binding clones from the M:anti-pep library had the same H2 sequence, which is identical to the H2 sequence

of the scaffold. Since this sequence is present in both the aforementioned high affinity clones and also in the low affinity clone M:ATE5, it is possible that it was isolated because it was slightly overrepresented in the library as a result of PCR artifacts and may not reflect the optimal sequence for antigen-binding. However, since the two NPY-specific clones did not contain the 26-10 H2 loop, it is likely that its selection in the M:ATB5, M:ATE5 and M:ATG11 antibodies reflects a possible structural role in stabilizing the binding pocket or enhancing specificity for smaller peptides such as AT.

The sequences of the two higher affinity M:anti pep-AT clones, M:ATB5 and M:ATG11, are identical except for the H3 region. The H3 length of M:ATG11 (K_D=52.6 nM) is 8 amino acids, while H3 of M:ATB5 (K_D=14.2 nM) is 10 amino acids. The side chains at the apex of H3 are not accessible to the binding pocket[26] when the H3 loop length is 10 amino acids or longer, rendering any interaction between the apical positions of H3 of M:ATE5 and AT unlikely. A key SDR in the parental 26-10 antibody, H100Trp, is replaced with smaller amino acids, namely Ala in M:ATB5 and Thr in M:ATG11 at position H98 which is equivalent to position H100 when H3 length is eight. The presence of smaller amino acids could enlarge the binding pocket, allowing the AT peptide to bind in an orientation analogous to that of digoxin in 26-10. The sequence of the low affinity M:anti-pep clone, M:ATE5, is different from the two higher affinity clones in H1 as well as in the H3 loop. The H3 loop length is eight amino acids and the sequence is almost identical to that of the higher affinity M:ATB5. Lys at position H97 of M:ATE5 is likely oriented towards the top of the binding pocket and could hinder entry of AT into the binding pocket compared to the much smaller Val or Gly at this

position in M:ATG11 or M:ATB5, respectively. Consistent with this notion, M:ATE5 exhibits a slower association rate.

Binders to NPY from the M:anti-pep library were isolated after five rounds of panning. Although the sequence diversity in round five was still high, additional rounds of screening resulted in a low polyclonal phage ELISA signal indicating de-enrichment (data not shown). Two clones showing the highest ELISA signal, M:NPYSol1 and M:NPYSol2, exhibited K_D values of 75.2 and 18.3nM, respectively. The respective amino acid sequences are distinct, but both exhibit an eight amino acid H3 loop length. A shorter H3 loop may serve to form a more shallow grove capable of accommodating the helical structure of NPY [10]. The solvent exposed apex of the H3 loop in M:NPYSol2 has two Trp residues. Such solvent exposed hydrophobic patches are often destabilizing in antibodies[50]. However, because of the short H3 of M:NPY2, these two Trp residues might actually be part of the binding pocket giving the hydrophobic side chains an opportunity to interact with NPY.

After three rounds of panning the Hu:anti-pep library, two distinct clones for each peptide target were obtained. The two clones specific for AT, Hu:AT1 and Hu:AT2 exhibited only micromolar K_D values. Despite differences in the length of H3 among the two clones, the k_ds were very similar: Hu:AT1 k_d=0.0139 s^{-1}: Hu:AT2 k_d=0.0112 s^{-1}. The H3 length of the lower affinity clone, Hu:AT2, is shorter than any of the other four anti-AT clones from either library. This might result in a more planar paratope surface which would not be able to interact as completely with the peptide-sized antigens[10]. The two NPY binders isolated from the Hu:anti-pep library differed markedly in their affinities. Interestingly, for clone Hu:NPY1, an unusually slow association rate (k_a=5.09

$\times 10^3$) was primarily responsible for the low overall affinity (K_D=4.1 μM) relative to the second clone, Hu:NPY2, which exhibited a K_D=48.6 nM. The slow association rate constant indicates that Hu:NPY1 may have to undergo a rate-limiting conformational change in order to bind antigen. Alternatively, the two aspartic acids at positions 31 (in the H1 loop) and 58 (H2) could be negatively impacting antigen binding. Recent computational studies have underscored the significance of electrostatic interactions in antibody: antigen recognition[51]. In addition, Hu:NPY2 contains a Trp at position 31 (in H1) compared to a Ser in the lower affinity Hu:NPY1 antibody. The sequence differences of the H3 loops of these two antibodies are relatively minor. The most notable difference is the presence of a His at position 98 of Hu:NPY1 compared to Gly in Hu:NPY2. Hu:NPY1 may have been selected in part because it is very well-expressed resulting in a yield of 4.8 mg/L of purified protein, more than 15-fold better than what was obtained with Hu:NPY2 (0.3 mg/L).

Possibly due to the flexible nature of peptides in the unbound state and the loss of entropy in the antibody-bound state[52], previous attempts to isolate high-affinity anti-peptide antibodies from synthetic libraries have led to mixed results. The presence of few or no reported hits against peptide antigens suggests a lack of peptide specificity in previous peptide libraries[38; 53; 54]. Screening of highly complex synthetic libraries based on human germline sequences resulted in isolated anti-peptide antibodies with K_D values > 1μM[53; 55]. Better anti-peptide antibody K_D values in the high-nanomolar range were reported for a library constructed by incorporating natural CDR diversity within a single scaffold[56]. Anti-peptide antibody affinities ranging from 67-420 nM were reported for a library constructed by incorporating natural CDR diversity within a single scaffold[56]. On

the other hand, as might be expected, immune repertoire libraries have resulted in antibodies with far greater affinity (K_D values in the picomolar range) to peptides[57]. Most notably, affinity maturation using ribosome display of an antibody isolated from an immune repertoire library has even resulted in an anti-peptide antibody with a K_D of 1pM[58].

The synthetic libraries reported here gave rise to antibodies with nanomolar dissociation constants against two structurally different peptides without the use of animal immunizations, affinity maturation, or light chain diversification. Analysis of selected clones provided a preliminary indication that the M:anti-pep library design is more suitable for the isolation of high-affinity antibodies to peptides relative to the Hu:anti-pep library. It is important to keep in mind that in both libraries the light chain was not subjected to diversification and consequently most, if not all, of the binding energy for AT or NPY likely arose from contacts with SDRs in the V_H. All the sequenced clones (Tables 3 and 4) from both libraries maintained the same canonical structure patterns as the respective scaffold scFvs. Encouraged by this early success in creating a fertile anti-peptide synthetic antibody library, work is currently underway to create next generation libraries that include V_L diversification.

EXPERIMENTAL PROTOCOLS

Library Construction

M:anti-pep Library: The anti-digoxin scFv 26-10[25; 46] was first cloned into the pAK200 vector[32]. TAA stop codons were introduced into each of the three V_H HV loops and the resulting gene, "scFv 26-10 stop" was used as the template for the construction of

the library. The stop codons are repaired during the library construction. Six or 10 putative SDR positions in H1 and H2, respectively, and between 4-9 positions in H3 were diversified depending on the length variation (n) introduced (n= 8, 10, or 12).

Diversity was introduced by separate PCR amplification of each of the three CDRs in the V_H using oligonucleotide primers (Table 2-8) containing degenerate codons using the codon diversification scheme proposed by Sidhu, *et al*[47] (Tables 2-3 and 2-4). Length and sequence diversity of H3 was generated by three separate PCR reactions using oligonucleotides encoding 8, 10, or 12 amino acids. H1 was amplified using an upstream vector specific primer and two reverse library primers, H2 was amplified using six forward library primers and 1 reverse primer, and each of the H3 loops was amplified using two forward library primers and one downstream vector specific reverse primer that also amplified the V_L region. High Fidelity polymerase (Roche Diagnostics, Indianapolis, IN, USA) was used in all PCR reactions to limit mutations to those encoded by the oligonucleotides.

Name	Sequence	Concentration
pAKscFvAmpF	5'-CATGAAATACCTATTGCCTACGGC	1mM
H1aR	5'-AGTCTAGAGACTTACCATGAGACTGGCGAACCCAAKDCATCSHADWMBYMBTAADAATGTACCCTGAGGA	1mM
H1bR	5'-AGTCTAGAGACTTACCATGAGACTGGCGAACCCAAKDCATAWWADWMBYMBTAADAATGTACCCTGAGGA	500nM
H2aF	5'-GGTAAGTCTCTAGACTGGATCGGGRNKWYTMRKHMTRVNRRTRBTDMTNYAVMNTACAACCAGAAGTTT	500nM
H2bF	5'-GGTAAGTCTCTAGACTACATCGGGRNKWYTMRKHMTTATRRTTHTRKANYATHTTACAACCAGAAGTTT	250nM
H2cF	5'-GGTAAGTCTCTAGACTGGATCGGGRNKWYTMRKHMTRVNRRTRBTDMTNYATHTTACAACCAGAAGTTT	500nM
H2dF	5'-GGTAAGTCTCTAGACTGGATCGGGRNKWYTMRKHMTTATRRTTHTRKANYAVMNTACAACCAGAAGTTT	250nM
H2eF	5'-GGTAAGTCTCTAGACTACATCGGGTGGWYTTGGHMTTATRRTTHTRKANYATHTTACAACCAGAAGTTT	25nM
H2fF	5'-GGTAAGTCTCTAGACTGGATCGGGTATWYTTATHMTTATRRTTHTRKANYAVMNTACAACCAGAAGTTT	25nM
H2R	5'-GCAATAGTATACCGCGGAGTCCTCAGA	1mM
H3n=8a	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSTATGCCATGGATTATTGGGGT	500nM
H3n=8b	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSDSGGCCATGGATTATTGGGGT	500nM
H3n=10a	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSNNSNNSTATGCCATGGATTATTGGGGT	500nM
H3n=10b	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSNNSNNSDSGGCCATGGATTATTGGGGT	500nM
H3n=12a	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSNNSNNSNNSNNSTATGCCATGGATTATTGGGGT	500nM
H3n=12b	5'-GCGGTATACTATTGCGCCSGTNNSNNSNNSNNSNNSNNSNNSDSGGCCATGGATTATTGGGGT	500nM
pAKscFvAmpR	5'-AAAATCACCGGAACCAGAA	1mM

Table 2-8. Oligonucleotide primers used for the construction of the M:anti-pep library.
The concentration of primers used in initial HV fragment PCR is also given. scFv 26-10 stop in pAK200 was used as template for the construction of all fragments. pAKscFvAmpF was used to amplify the H1 from the pAK vector and to amplify the final overlap-extension library insert. pAKscFvAmpR used to amplify H3 from the pAK vector and to amplify the final overlap-extension library insert. SDR diversity guided by degenerate codons and varying concentrations of each primer. H3 length diversity provided by additional NNS codons in H3 primers. In primer nomenclature, "R" refers to primers used to amplify in the reverse direction from the sense strand, while "F" refers to primers used to amplify in the forward direction from the anti-sense strand.

Each CDR fragment was agarose-gel purified and extracted using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). The respective fragments were assembled into the complete scFv by overlap-extension PCR. This reaction included 150ng of DNA from each of the genes having diversified CDR regions region (*i.e.* H1, H2, $H3_8$, $H3_{10}$ and $H3_{12}$). The full length scFv fragment was PCR purified using the QIAquick PCR Purification Kit and digested with *Sfi*I (New England Biolabs, Ipswich, MA, USA) at 50°C for six hours, gel purified and ligated, using T4 ligase (New England Biolabs) overnight at 16°C, into the complementary sites of *Sfi*I-cut phagemid, pAK200[32] using a molar ratio of 2:1 (insert to vector). Ligation reactions were desalted on nitrocellulose (Millipore, Billerica, MA, USA) and 200ng of DNA were electroporated into *E. coli*

XL1-Blue (Stratagene, La Jolla, CA, USA) in 10 separate reactions, recovered, plated on TYE-2% glucose supplemented with 34μg/mL chloramphenicol and incubated overnight at 30°C[32; 38] resulting in library of 5.1×10^8 independent transformants.

Hu:anti-pep Library: The scaffold of the Hu:anti-pep library comprises the V_H of 3-23(DP-47)/J_H4[38], and the V_L of A27md and $J_{\kappa1}$. The construction of the V_H repertoire was described earlier[17]. V_H synthetic fragment was purified, digested with *Xho*I and *Sfi*I, and ligated into the phagemid vector pHEN1-V_LA27md/Jκ3 as previously described[17]. The library DNA was amplified as above and following electroporation, transformants were selected by plating on plates containing 100 μg/L ampicillin instead of chloramphenicol. 2×10^8 independent transformants were obtained.

Degenerate DNA is represented using the IUB code: N=A,C,G,T; V=G,A,C; B=G,T,C; H=A,T,C; D=G,A,T; K=G,T; S=G,C; W=A,T; M=A,C; Y=C,T; R=A,G.

Phage panning

Library cells were scraped and used to inoculate 50mL 2×YT-2% glucose containing chloramphenicol resulting in an OD_{600} of 0.1. The cells were grown at 30°C to OD_{600} =0.6 and infected with a 100-fold concentration of VCS-M13 helper phage (Stratagene, La Jolla, CA, USA). Subsequently, the cells were incubated at 37°C for one hour, spun down, and resuspended in 100mL of 2×YT containing 0.1% glucose, chloramphenicol as above and kanamycin at 50 μg/L and incubated overnight at 30°C. Phage particles were harvested from the cultured supernatant following two rounds of precipitation with 20% PEG_{6000} /2.5M NaCl (Tomlinson Protocol, Geneservice, UK), stored at 4°C and panned within 18 hours.

Panning was performed in solution with biotinylated-human angiotensin I (AT) and biotinylated human neuropeptide Y (NPY) (Genscript, Piscataway, NJ, USA) using streptavidin-coated paramagnetic beads (Invitrogen, Carlsbad, CA, USA) essentially as described[59]. Prior to panning the phage with 10μM biotin-peptide, non-specific phage antibodies were depleted on 100μL naked streptavidin beads without peptide. The remaining phage antibodies were then incubated with 10μM biotin-peptide in 1 mL PBS before the phage antibody-biotin-peptide complexes were captured with 100μL streptavidin beads. The beads were washed vigorously and bound phage were eluted by incubating the beads with 1mL 100mM triethylamine (TEA) for 10 min while rotating. The beads were separated and the TEA-phage mixture was added immediately to 0.5 mL 1M Tris, pH 7.4 for neutralization. 9mL of *E. coli* XL1-Blue (OD_{600}= 0.7) culture were infected, the cells were plated on selective media as above, colonies were scraped and used to inoculate liquid cultures for phage preparation as above. The rescued phage particles were used in subsequent selection rounds.

Phage ELISAs

To evaluate the enrichment of selected phage pools, the binding of polyclonal phage populations to the target peptide was monitored by polyclonal ELISA. Costar 3590 (Corning) plates were coated overnight[59] with 100μL/well biotin-BSA (2μg/mL in PBS) (Sigma-Aldrich, St. Louis, MO, USA) at 4°C and washed three time with PBS-0.1% Tween20 (EMD) between reagents and samples. 100μL/well streptavidin (a 10μg/mL solution in PBS-0.5% gelatin, Jackson Laboratories) was then added for one hour at room temperature followed by 100μL/well of biotinylated-peptide (2μg/mL in PBS) for one

hour at room temperature. The plates were blocked with 200μL/well PBS-2% milk for one hour. 100μL/well of normalized titers of phage antibodies from each round of panning were added to the plate in PBS-2% milk, in triplicate. After one hour at room temperature bound phage was detected with 100μL/well of anti-M13-horseradish peroxidase (HRP) conjugate (1:5000) (Amersham-Pharmacia Biosciences, Piscataway, NJ, USA) followed by development with 50μL/well TMB substrate (Dako) quenching with 50μL/well 1N sulfuric acid. Plates were read at OD_{450} and the signals from background binding were subtracted from the peptide binding.

To analyze individual clones, randomly selected clones from rounds of panning showing maximal enrichment by polyclonal ELISA were cultured overnight in microtiter plates (Tomlinson Protocol, Geneservice, UK). 50μL of phage supernatant from each well, was collected and added to ELISA plates containing either the target peptide or an unrelated peptide. The plates were coated and developed as described above for the polyclonal analysis. A ratio of the absorbance reading for each clone's ability to bind to the target peptide and to the unrelated peptide was determined.

Expression of soluble scFvs

Infected cells collected from the last round of panning of the M:anti-peptide library as described, were cultured in 100mL TB (Difco) containing 2% glucose and chloramphenicol to an OD_{600} of 1.5. Plasmid DNA was prepared, digested with *Sfi*I, gel purified, ligated into pAK400[32] and the DNA was transformed into *E. coli* Jude-1[60] and plated on TYE with 2% glucose and chloramphenicol. Cloning into pAK400 allows expression of scFvs with a C-terminal His tag without fusion to the pIII.

Phage from the isolated clones of the Hu:anti-pep library were infected into HB2151 *E. coli*, an F', non-suppressor strain (Maxim Biotech, Rockville, MD, USA) to promote recognition of the amber codon between the scFv coding region and gene III, enabling expression of his-tagged scFv from pHEN without the pIII fusion.

Selected clones from both libraries were grown in 400 mL TB containing 2% glucose, the appropriate antibiotic and one drop of anti-foam C (Sigma-Aldrich) at 25°C. After 3 hours, 1mM IPTG was added and incubation was continued for an additional 4 hours. The cells were pelleted (5000g × 10min), and the periplasmic fraction was isolated. Briefly, the cell pellet was resuspended in 10mL ice-cold 100mM Tris-HCl, pH 7.4 containing 0.75M Sucrose and 1mL 10mg/mL lysozyme (Sigma-Aldrich). 20mL of 1mM EDTA was added, the cells were incubated for 15 minutes and then 1.4mL 0.5M $MgCl_2$ was added for another 15 min. The periplasmic fraction was separated from the spheroplasts by centrifuging the lysed mixture at 12,000g for 30 min. An equal volume of immobilized metal affinity chromatography buffer (IMAC; 20mM Na_2HPO_4, 500mM NaCl, 10mM imidazole, 0.1% Tween-20, pH 7.4) was added followed by 0.5mL washed Ni-NTA Agarose (Qiagen) for 1 hour. The resin was washed twice in IMAC buffer and elution was carried out with 4mL 250mM imidazole in IMAC buffer. Monomeric scFv was further purified by size exclusion gel filtration chromatography on a Superdex 200 column (Amersham-Pharmacia Biosciences, Uppsala, Sweden).

BIACore Analysis

Antigen binding kinetics of monomeric scFvs were determined *via* surface plasmon resonance analysis using a BIAcore 3000 (BIAcore-GE Healthcare, Piscataway, NJ,

USA). Biotinylated peptides were captured on SA sensor chip with the reference cell containing streptavidin without peptide. Five different concentrations of each scFv were injected in triplicate over the sensor chips at a flow rate of 30μL/min for 2 min for k_a measurements. k_d was assessed during a 10 min dissociation stage. The chip was regenerated by a one min injection of 10mM glycine, pH 2.0 followed by a one minute injection of HBS-EP at a flow rate of 50μL/min. Response curves were generated by subtracting the response of the reference cell from that of the experimental cell and the data were analyzed using BIAevaluation v. 3.2. Binding kinetics were calculated by simultaneously fitting the k_a and k_d data to a 1:1 Langmuir binding model. Only data fits with a χ^2 value less than 2.5 were used.

REFERENCES

1. Chothia, C. & Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* **196**, 901-17.
2. Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B. L. & Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nat Biotechnol* **25**, 563-5.
3. Sidhu, S. S. & Fellouse, F. A. (2006). Synthetic therapeutic antibodies. *Nat Chem Biol* **2**, 682-8.
4. Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. *Nat Biotechnol* **23**, 1105-16.
5. Amzel, L. & Poljak, R. (1979). Three-Dimensional Structure of Immunoglobulins. *Annual Review of Biochemistry* **48**, 961-997.
6. Almagro, J. C. (2004). Identification of differences in the specificity-determining residues of antibodies that recognize antigens of different size: implications for the rational design of antibody repertoires. *J Mol Recognit* **17**, 132-43.
7. Van Regenmortel, M. H. (2002). Reductionism and the search for structure-function relationships in antibody molecules. *J Mol Recognit* **15**, 240-7.
8. Wilson, I. A. & Stanfield, R. L. (1994). Antibody-antigen interactions: new structures and new conformational changes. *Curr Opin Struct Biol* **4**, 857-67.
9. Collis, A., Brouwer, B. & Martin, A. (2003). Analysis of the antigen combining site: correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. *J Mol Biol* **325**, 337-354.
10. Vargas-Madrazo, E., Lara-Ochoa, F. & Almagro, J. C. (1995). Canonical structure repertoire of the antigen-binding site of immunoglobulins suggests

strong geometrical restrictions associated to the mechanism of immune recognition. *J Mol Biol* **254**, 497-504.
11. MacCallum, R. M., Martin, A. C. R. & Thornton, J. M. (1996). Antibody-antigen Interactions: Contact Analysis and Binding Site Topography. *J Mol Biol* **262**, 732-745.
12. Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R. & et al. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* **342**, 877-83.
13. Persson, H., Lantto, J. & Ohlin, M. (2006). A focused antibody library for improved hapten recognition. *J Mol Biol* **357**, 607-20.
14. Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). Structural repertoire of the human VH segments. *J Mol Biol* **227**, 799-817.
15. Almagro, J. C., Vargas-Madrazo, E., Zenteno-Cuevas, R., Hernandez-Mendiola, V. & Lara-Ochoa, F. (1995). VIR: a computational tool for analysis of immunoglobulin sequences. *Biosystems* **35**, 25-32.
16. Padlan, E. A., Abergel, C. & Tipper, J. P. (1995). Identification of specificity-determining residues in antibodies. *Faseb J* **9**, 133-9.
17. Almagro, J. C., Quintero-Hernandez, V., Ortiz-Leon, M., Velandia, A., Smith, S. L. & Becerril, B. (2006). Design and validation of a synthetic VH repertoire with tailored diversity for protein recognition. *J Mol Recognit* **19**, 413-22.
18. Barbas, C. F., 3rd, Amberg, W., Simoncsits, A., Jones, T. M. & Lerner, R. A. (1993). Selection of human anti-hapten antibodies from semisynthetic libraries. *Gene* **137**, 57-62.
19. Brett, P. J., Tiwana, H., Feavers, I. M. & Charalambous, B. M. (2002). Characterization of oligopeptides that cross-react with carbohydrate-specific antibodies by real time kinetics, in-solution competition enzyme-linked immunosorbent assay, and immunological analyses. *J Biol Chem* **277**, 20468-76.
20. Charalambous, B. M. & Feavers, I. M. (2000). Peptide mimics elicit antibody responses against the outer-membrane lipooligosaccharide of group B neisseria meningitidis. *FEMS Microbiol Lett* **191**, 45-50.
21. Charalambous, B. M. & Feavers, I. M. (2001). Mimotope vaccines. *J Med Microbiol* **50**, 937-9.
22. Tiwana, H., Clow, K. J., Hall, C., Feavers, I. M. & Charalambous, B. M. (2005). The immunogenicity of a conformationally restricted peptide mimetic of meningococcal lipooligosaccharide. *Scand J Immunol* **62**, 385-92.
23. Stresser, D. M. & Kupfer, D. (1999). Monospecific antipeptide antibody to cytochrome P-450 2B6. *Drug Metab Dispos* **27**, 517-25.
24. Sobel, J. H., Trakht, I., Pileggi, N. & Qi Wu, H. (1998). Antipeptide monoclonal antibodies to defined fibrinogen Aalpha chain regions: anti-Aalpha 487-498, a structural probe for fibrinogenolysis. *Blood* **91**, 1590-8.
25. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & et al. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci U S A* **85**, 5879-83.

26. Jeffrey, P. D., Strong, R. K., Sieker, L. C., Chang, C. Y., Campbell, R. L., Petsko, G. A., Haber, E., Margolies, M. N. & Sheriff, S. (1993). 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc Natl Acad Sci U S A* **90**, 10310-4.
27. Kabat, E. A. & Wu, T. T. (1971). Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann N Y Acad Sci* **190**, 382-93.
28. Xu, J. L. & Davis, M. M. (2000). Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* **13**, 37-45.
29. Lee, C. V., Liang, W. C., Dennis, M. S., Eigenbrot, C., Sidhu, S. S. & Fuh, G. (2004). High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. *J Mol Biol* **340**, 1073-93.
30. Morea, V., Tramontano, A., Rustici, M., Chothia, C. & Lesk, A. M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J Mol Biol* **275**, 269-94.
31. Shirai, H., Kidera, A. & Nakamura, H. (1999). H3-rules: identification of CDR-H3 structures in antibodies. *FEBS Lett* **455**, 188-97.
32. Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R. & Pluckthun, A. (1997). 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.
33. Kobayashi, N., Soderlind, E. & Borrebaeck, C. A. (1997). Analysis of assembly of synthetic antibody fragments: expression of functional scFv with predefined specificity. *Biotechniques* **23**, 500-3.
34. Fellouse, F. A., Barthelemy, P. A., Kelley, R. F. & Sidhu, S. S. (2006). 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.
35. Fellouse, F. A., Wiesmann, C. & Sidhu, S. S. (2004). Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc Natl Acad Sci U S A* **101**, 12467-72.
36. Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T. & Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J Exp Med* **188**, 2151-62.
37. Rojas, G., Almagro, J. C., Acevedo, B. & Gavilondo, J. V. (2002). Phage antibody fragments library combining a single human light chain variable region with immune mouse heavy chain variable regions. *J Biotechnol* **94**, 287-98.
38. Griffiths, A., Williams, S., Hartley, O., Tomlinson, I., Waterhouse, P., Crosby, W., Kontermann, R., Jones, P., Low, N., Allison, T., Prospero, T., Hoogenboom, H., Nissim, A., Cox, J., Harrison, J., Zaccolo, M., Gherardi, E. & Winter, G. (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J* **13**, 3245–3260.
39. Ohlin, M. & Borrebaeck, C. A. (1996). Characteristics of human antibody repertoires following active immune responses *in vivo*. *Mol Immunol* **33**, 583-92.
40. de Wildt, R. M., Hoet, R. M., van Venrooij, W. J., Tomlinson, I. M. & Winter, G. (1999). Analysis of heavy and light chain pairings indicates that receptor editing shapes the human antibody repertoire. *J Mol Biol* **285**, 895-901.

41. Tomlinson, I. M., Cox, J. P., Gherardi, E., Lesk, A. M. & Chothia, C. (1995). The structural repertoire of the human V kappa domain. *Embo J* **14**, 4628-38.
42. Vargas-Madrazo, E., Almagro, J. C. & Lara-Ochoa, F. (1995). Structural repertoire in VH pseudogenes of immunoglobulins: comparison with human germline genes and human amino acid sequences. *J Mol Biol* **246**, 74-81.
43. Almagro, J. C., Hernandez, I., Ramirez, M. C. & Vargas-Madrazo, E. (1998). Structural differences between the repertoires of mouse and human germline genes and their evolutionary implications. *Immunogenetics* **47**, 355-63.
44. Ball, W. J., Jr., Wang, Z., Malik, B., Kasturi, R., Dey, P., Short, M. K. & Margolies, M. N. (2000). Selection of peptidic mimics of digoxin from phage-displayed peptide libraries by anti-digoxin antibodies. *J Mol Biol* **301**, 101-15.
45. Ribnicky, B., Van Blarcom, T. & Georgiou, G. (2007). A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. *J Mol Biol* **369**, 631-9.
46. Hunter, M. M., Margolies, M. N., Ju, A. & Haber, E. (1982). High-affinity monoclonal antibodies to the cardiac glycoside, digoxin. *J Immunol* **129**, 1165-72.
47. Sidhu, S. S., Li, B., Chen, Y., Fellouse, F. A., Eigenbrot, C. & Fuh, G. (2004). Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. *J Mol Biol* **338**, 299-310.
48. Spyroulias, G. A., Nikolakopoulou, P., Tzakos, A., Gerothanassis, I. P., Magafa, V., Manessi-Zoupa, E. & Cordopatis, P. (2003). Comparison of the solution structures of angiotensin I & II. Implication for structure-function relationship. *Eur J Biochem* **270**, 2163-73.
49. Monks, S. A., Karagianis, G., Howlett, G. J. & Norton, R. S. (1996). Solution structure of human neuropeptide Y. *J Biomol NMR* **8**, 379-90.
50. Worn, A. & Pluckthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* **305**, 989-1010.
51. Lippow, S. M., Wittrup, K. D. & Tidor, B. (2007). Computational design of antibody-affinity improvement beyond *in vivo* maturation. *Nat Biotechnol*.
52. Zahnd, C., Spinelli, S., Luginbühl, B., Amstutz, P., Cambillau, C. & Plückthun, A. (2004). Directed *in Vitro* Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. *J Biol Chem* **279**, 18870-18877.
53. Krebs, B., Rauchenberger, R., Reiffert, S., Rothe, C., Tesar, M., Thomassen, E., Cao, M., Dreier, T., Fischer, D., Hoss, A., Inge, L., Knappik, A., Marget, M., Pack, P., Meng, X. Q., Schier, R., Sohlemann, P., Winter, J., Wolle, J. & Kretzschmar, T. (2001). High-throughput generation and engineering of recombinant human antibodies. *J Immunol Methods* **254**, 67-84.
54. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. & Johnson, K. S. (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat Biotechnol* **14**, 309-14.
55. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. & Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* **296**, 57-86.

56. Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C., Danielsson, L., Carlsson, R. & Borrebaeck, C. A. (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol* **18**, 852-6.
57. Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H. R. & Pluckthun, A. (1998). Ribosome display efficiently selects and evolves high-affinity antibodies *in vitro* from immune libraries. *Proc Natl Acad Sci U S A* **95**, 14130-5.
58. Luginbuhl, B., Kanyo, Z., Jones, R. M., Fletterick, R. J., Prusiner, S. B., Cohen, F. E., Williamson, R. A., Burton, D. R. & Pluckthun, A. (2006). Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation. *J Mol Biol* **363**, 75-97.
59. Chames, P., Hoogenboom, H. R. & Henderikx, P. (2002). Selection of antibodies against biotinylated antigens. *Methods Mol Biol* **178**, 147-57.
60. Griswold, K. E., Kawarasaki, Y., Ghoneim, N., Benkovic, S. J., Iverson, B. L. & Georgiou, G. (2005). Evolution of highly active enzymes by homology-independent recombination. *Proc Natl Acad Sci U S A* **102**, 10082-7.

Chapter 3: Analysis of Consensus Mutations in High Error Rate Libraries of scFv Antibodies

INTRODUCTION

Enhancements to a protein's function, expression level, stability, ligand-binding affinity, and other properties can be engineered by sequence randomization using a variety of mutagenesis strategies followed by selections to isolate improved variants. Random mutagenesis, commonly performed by error-prone PCR, is one of the most efficient strategies to sample a large sequence space. Error-prone PCR is a technique used to introduce random mutations in an amplified stretch of DNA by lowering the fidelity of the polymerase[1]. The mutation frequency induced by decreasing polymerase fidelity can be controlled by varying a number of parameters such the $[MnCl_2]$[2], $[MgCl_2]$[3], the amount of polymerase[1], or using forcing concentration of one of the di-deoxy nucleotides in the reaction[1; 4]. Traditionally, random mutagenesis rates have been kept low in directed evolution efforts to avoid an overabundance of deleterious mutations causing the inactivation of too many library proteins[5; 6]. Such deleterious mutations from high-error rates in combinatorial libraries are associated with large-scale perturbations in the protein conformation or function and with the introduction of stop codons. As a result, the number of functional clones decreases as the mutational load is increased[7; 8]. Several studies describe an exponential loss in library activity relative to increasing mutation rates[4; 6]. As the error-rate increases, the number of mutations per sequence in the libraries also increases. The throughput of most of the methods used to screen libraries for improvements is often not high enough to sample efficiently the vast

sequence space of high error-rate libraries. In summary, because high error-rates result in a library with a small functional fraction contained in a large sequence space that is difficult to screen, the use of high rates of mutagenesis is not considered to be an effective method of generating diversity in libraries for directed evolution. On the other hand, high rates of mutagenesis represent a tractable strategy for sparse sampling of protein sequence space. Previous reports indicate that the exponentially inverse relationship between mutation rate and library activity begins breaking down when high mutation rates are applied to a library, indicating that there are numerous evolutionary paths to improved fitness[7; 9]. High error-rate libraries are more likely to enable the creation of coupled mutations, *i.e.* pairs of amino acid substitutions that interact with one another epistatically to enhance function and to be identified[7; 9; 10]. Some of these interacting mutations may directly increase function. Others cooperate to provide protein enhancements in a more indirect manner. For example, a certain mutation might destabilize global protein folding, causing it to not be selected despite any enhancements that it makes to the protein's function. However, if that mutation can be coupled with a second mutation that is able to counteract the destabilizing effect of the first mutation, the enhanced and stable protein can now be selected along with its coupled mutations[9].

The variable domains of antibodies have been used in several previous studies to examine the effects of different mutagenesis strategies on protein function[11; 12; 13; 14; 15; 16]. Antibodies are well-suited to mutagenesis studies because residue replacement is highly tolerated throughout the protein, especially in the complementarity determining regions (CDRs), indicating plasticity in terms of folding, specificity, and affinity[16; 17]. Conservative (*e.g.* F→Y) mutations in antibodies variable domains are highly tolerated[18].

In earlier studies, the high affinity anti-digoxin single chain fragment (scFv) derived from the 26-10 monoclonal anti-digoxin antibody was used as the parent gene for the creation of two high mutation rate libraries[10]. The libraries were screened by displaying the scFvs as tethered Lpp-OmpA-scFv fusions on the outer membrane of *E. coli*, incubated with fluorescently labeled digoxin, and sorting using fluorescence-activated cell sorting (FACS). Despite the high-error rate in these libraries, the fraction of active clones was higher than what would be expected based on extrapolations of the linear relationship of loss in library activity due to high error-rate mutagenesis[7; 9; 10]. Digoxin binding clones were readily isolated from this library by FACS screening as discussed above. Sequencing revealed the presence of an average between 4.3 (L2) and 5.4 (L3) amino acid changes per selected gene, suggesting that even in a highly mutated library, there are epistatic mutations that retain or improve antibody function. The sequence alignment of 91 clones isolated from the two high-error rate 26-10 scFv libraries revealed the appearance of three consensus mutations around CDRH1: V_H24(S→F), V_H29(F→S), and V_H34(M→V).

The following sections describe an attempt to determine the role of these high frequency mutations in improving antibody affinity. Specifically, we sought to evaluate whether the V_H24(S→F), V_H29(F→S), and V_H34(M→V) couple to confer higher cell fluorescence due to enhanced expression, affinity, or stability and whether the effect of combining the individual consensus amino acid substitutions is additive.

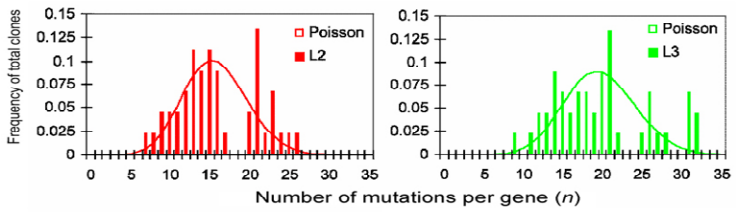
RESULTS

Library construction and analysis

An scFv derived from the monoclonal antibody, 26-10[19], which binds with high affinity to the cardiac glycoside digoxin, was used as the template to create two error-prone PCR (EP-PCR) based libraries of scFvs (L2 and L3) as described previously[10]. Briefly the scFv proteins were displayed on the surface of *E. coli* via a fusion to an outer membrane targeting sequence comprised of Lpp-OmpA[10; 20]. The fusion protein was expressed from the P_{BAD} arabinose promoter in the *E. coli* strain, LMG194. The mean number of mutations (*m*) for each of the two libraries analyzed was determined by sequencing 45 clones from L2 and 46 clones from L3. A mean number of mutations, m_{L2} = 15.8 (2.17% of the gene) and m_{L3} = 19.8 (2.72%) was thus estimated. The average number of amino acid mutations was, respectively: 10.5 (4.33%) and 13.1 (5.42%). Figure 3-1 shows the frequency distribution of mutations per gene (*n*) among the sequenced clones from each library overlaid with the predicted Poisson distribution.

The types of substitution mutations introduced at the genetic level from each library were analyzed (Table 3-1). Mutations at A-T base pairs occur more frequently than at G-C base pairs in L2, while the inverse is true with L3. In L2, the A→T, A→G, and G→A mutations (along with the complementary mutations) occur more frequently than A→C, G→C, and G→T substitutions. In L3, the trend is reversed with G→A being more dominant, while A→T and A→C are still prevalent.

Following incubation with a BODIPY-digoxin conjugate, cells were analyzed by flow cytometry and the most fluorescent clones were isolated[21]. The library screening and pre-screening analysis was performed by Dr. Richard Loo.



Presorted Library	Number of clones	Avg. number of DNA mutations	Percentage of DNA sequence space mutated	Avg. number of amino acid mutations per gene	Percentage of amino acids mutated
L2	46	15.8	2.17	10.5	4.33
L3	45	19.8	2.72	13.1	5.42

Figure 3-1. Presort analysis of both high error-rate libraries.
The distributions of number of mutations per sequenced clone from each library with an overlaid Poisson distribution are shown in the two charts. The number of mutations at both the genetic and amino acid level with the average percent of the gene mutated is also shown. The 26-10 scFv gene is 753 bp in length.

	L2 (2.17%) Library		L3 (2.72%) Library	
Type of substitution	Number	Frequency	Number	Frequency
A→T, T→A	172	23.7%	103	12.0%
A→C, T→G	7	1.0%	7	0.8%
A→G, T→C	336	46.4%	194	22.6%
G→A, C→T	188	25.9%	508	59.3%
G→C, C→G	11	1.5%	28	3.3%
G→T, C→A	11	1.5%	17	2.0%
A→N, T→N	515	71.0%	304	35.5%
G→N, C→N	210	29.0%	553	64.5%

Table 3-1. Types of mutations.
Mutational spectrum showing the occurrence and frequencies of all the types of substitutions in the clones sequenced from each library prior to the selections. The genetic mutation rate for each library is provided in parentheses.

The frequency of fluorescent clones expressing functional scFv antibodies capable of binding to digoxin was found to be much larger than was anticipated based on low error-rate library studies. In low error-rate libraries, the number of functional clones is predicted to decrease exponentially as a function of the average number of mutations[4] according to a Poisson model where P is the probability that a sequence will have m number of mutations, $\langle m_{nt} \rangle$ is the mean number of nucleotide mutations per sequence, and e is the base of the natural logarithm[4; 9]:

$$P\left(m;\left\langle m_{nt}\right\rangle\right)=\frac{\left\langle m_{nt}\right\rangle^{m} e^{-\left\langle m_{nt}\right\rangle}}{m!}$$

However, in high-error rates libraries, the frequency of mutations follows a PCR-modeled distribution instead of a Poisson distribution[22]. As a result, the predicted fraction of functional clones in such libraries is much higher than what is predicted from a Poisson distribution[9]. A consequence of modeling the process of mutation generation

by PCR is a much larger variance relative to the Poisson distribution which is no longer equal to the mean of the distribution[9]. The model, developed by Drummond, *et al*.[9] reveals that the frequency of active clones in high-error rate libraries is far greater than that predicted by libraries that have a Poisson distribution of mutations (*i.e.* low-error rate libraries). This is the case for of the L2 and L3 scFv libraries analyzed here.

Library	Observed fraction active	Expected fraction active*	Observed-to-expected ratio
L2	1.2×10^{-3}	1.7×10^{-5}	71
L3	4.1×10^{-4}	1.6×10^{-6}	256

Table 3-2. Library activity.
Prior to selections, the number of observed active clones in each library determined by flow cytometry is compared to the expected number of active clones, based on an extrapolation of error-rate vs. activity of lower error-rate libraries. *Expected values were extrapolated from low error-rate libraries which are characterized by a Poisson distribution of mutations

Analysis of clones exhibiting increased cell fluorescence

L2 and L3 were screened by FACS for binding to BODIPY-digoxin. In the first and second rounds, positive clones were isolated after incubating the cell population with 100 nM BODIPY-digoxin. The most fluorescent cells were isolated using the recovery mode of the flow cytometer that collects cells based on strong fluorescence but also results in the enrichment of nearby coincident cells. In rounds 3-5, in addition to being labeled with 100 nM BODIPY-digoxin, the cells were also co-incubated with 2 μM digoxin as a competitor to select for clones with scFvs that exhibit slower hapten dissociation rates.

46 colonies were randomly picked from the final round of enrichment of L2 and 45 from L3 and sequenced. A sequence alignment indicated the appearance of consensus mutations at several positions. Specifically in 62% of the sequenced clones from library L2 and 35% from L3, the serine at position V_H24 was changed to a phenylalanine (Table 3-3). Positions 25, 29, 30, 32, and 34 of the V_H, are also mutational hotspots. Table 3-3 shows the frequency that each mutation occurs simultaneously with the mutation Ser→Phe mutation at position V_H 24. In both libraries, the Ser→Phe mutation at position V_H 24 is produced in all clones possessing this mutation through the same base-pair substitution to the wild-type codon (TCC→TTC). Likewise, the consensus mutations at V_H positions 25, 29, 32, and 34 of both libraries are produced by only one, repeating base-pair substitution. On the other hand, the amino acid mutations at V_H positions 6 and 30 are produced by two and three mutant codons, respectively.

Library	Percent mutated: SerH24	Second mutated residue	Percent mutated: SerH24 + 2nd residue
L2	62	GlnH6	19
		SerH25	24
		PheH29	24
		ThrH30	29
		PheH32	33
		MetH34	18
L3	35	SerH25	19
		MetH34	15

Table 3-3. Commonly occurring mutations.
Common mutations observed in an alignment of sequences from 91 (45 sequences from L2 and 46 from L3) randomly isolated clones after five rounds of selection. The frequency of the most common mutation in each library, S→F at position V_H24, is given. The frequencies that other consensus mutations appear simultaneously with the S→F mutation at V_H24 in the sequence alignments are also provided.

Construction and characterization of single amino acid variants of the 26-10 scFv antibody

Various combinations of the consensus mutations discussed above (at positions V_H 24, 25, 29, 30, 32, and 34) were constructed using site-directed mutagenesis. The resulting genes were cloned into the pB30D vector and transformed into *E. coli* LMG194. Cells were incubated with 100 nM BODIPY-digoxin and fluorescence was measured by FACS. The clones with the highest fluorescence contained the following mutations: V_H24(S→F), V_H29(F→S), and V_H34(M→V) (data not shown). Based on the fluorescence rank-ordering, the clones described in Table 3-4 were selected for more detailed biochemical analysis:

Name of clone	Position(s) mutated	Mutation
24	V_H24	Ser→Phe
29	V_H29	Phe→Ser
34	V_H34	Met→Val
2429	V_H24	Ser→Phe
	V_H29	Phe→Ser
2434	V_H24	Ser→Phe
	V_H34	Met→Val
2934	V_H29	Phe→Ser
	V_H34	Met→Val
242934	V_H24	Ser→Phe
	V_H29	Phe→Ser
	V_H34	Met→Val

Table 3-4. Variant clones created by site-directed mutagenesis.
The variants listed were created through site-directed mutagenesis of the 26-10 antibody to incorporate the listed mutations. These variants reflect the most commonly selected mutations from the selections of both libraries. The name of each clone and the position(s) and identity of each mutation are listed.

The variable regions of the genes encoding the mutations shown in Table 3-4 were cloned into the pMoPac16 vector[23] between the *pel*B leader and the human constant kappa domain (huCκ)[24]. Fusion of the variable regions to the huCκ generates single chain antibodies (scAb) which are more soluble and less prone to *C*-terminal proteolysis than scFvs[23]. There is a hexa-histidine sequence (His-tag) fused to the *C*-terminus of the huCκ domain for ease of detection and purification. pMoPac16 also expresses the Skp foldase chaperone that enhances solubility of the antibody fragments without affecting the expression levels[23]. 440 mL cultures of each clone were grown in shake flasks and protein synthesis was induced with 1 mM IPTG at room temperature for four hours. The periplasmic fraction, containing His-tagged scAb protein, of each culture was collected after osmotic shock. The scAbs were purified from the periplasmic fractions by immobilized metal affinity chromatography (IMAC). Size exclusion FPLC was then used to purify the monomeric scAb from dimeric scAb.

The kinetics of binding and dissociation of the freshly purified, monomeric scFvs to digoxin were measured by SPR on a BIAcore 3000 instrument. Briefly, digoxin, was conjugated to bovine serum albumin (BSA) *via* oxidation of the terminal sugar residues of digoxin followed by covalent attachment to BSA through reductive amination[11]. The kinetics for all of the mutant clones, as well as the wild-type, are listed in Table 3-5. That statistical analysis on all the data proved significant. Six of the seven variant clones, except for 34, show similar or moderately lower K_Ds (between 1.5-fold and 3.5-fold lower) than the parent gene, 26-10 (4.41 nM), indicating that the mutations are neutral or enhancing to antigen affinity. The on-rates, k_a, of all the variants that contain the selected F→S mutation at V_H29 increased from wild-type (1.85×10^5 M^{-1} S^{-1}) resulting in the largest enhancement in K_D. The mutations S→F at V_H24 and M→V at V_H34 confer a minor decrease (*i.e.* enhancement) relative to the wild-type in the off-rate, k_d. Even though these two mutations appear to enhance the k_d, in one case they worsen the k_a (*e.g.* M→V mutation at position V_H34 alone) resulting in a diminished K_D relative to the wild-type. The mutations S→F at V_H24 and M→V at V_H34 appear to cooperate in a non-additive manner to lower the k_d while increasing the k_a, leading to an enhanced K_D.

Mutant clone	k_a (1/Ms)	k_d (1/s)	K_D (nM)	X^2
26-10	1.85	8.16	4.41	0.9
24	2.19	6.27	2.86	1.3
29	6.39	8.43	1.32	0.7
34	1.03	5.96	5.79	0.9
2429	3.07	5.54	1.80	1.2
2434	1.71	4.94	2.89	0.7
2934	3.04	6.19	2.04	1.8
242934	3.36	4.26	1.27	0.9

Table 3-5. Kinetics of 26-10 and mutants determined with BIAcore.
The kinetics of monomeric purified protein as determined by BIAcore analysis using direct binding are listed. X^2 is a measure of how well the data are modeled. Any X^2 value less than 2.5 is considered highly accurate.

The consensus mutations were selected using directed evolution methods that previously have been used to isolate antibodies with higher affinities than the parent antibody[10; 25]. The lack of any exceptional affinity enhancements amongst the mutants suggests that other criteria could have been used to drive the selection of these three consensus mutations. The FACS-based selection method relies on enhanced fluorescence of a clone relative to the wild-type, to isolate individual clones from the original libraries. The fluorescence is directly related to the amount of fluorescently-labeled antigen bound to the cell-surface displayed antibodies. Typically, the affinity of the displayed antibody is the primary determinant of the amount of bound antigen and the resulting fluorescent signal.

However, the expression level of the antibody displayed on the surface of each cell can also influence the amount of bound antigen and therefore the resulting fluorescent signal. In light of the results that show the mutations at positions V_H24, V_H29, and V_H34 are selected despite the lack of any exceptional affinity improvements, it is important to

determine if these variants show enhanced expression levels. Variations in expression level of antibodies, which may explain the selection of certain mutations, are strongly influenced by the amino acid sequences of the antibody variable regions[26]. Single amino acid substitutions can dramatically alter the expression levels of antibodies[27]. The expression levels of the variants and wild-type protein were determined from *E. coli* cultures of each clone that were normalized to the same optical density to ensure that the same number of cells was used in protein preparations. After purification of scAb from each clone, the amount of monomer, dimer, and total scAb were determined (Figure 3-2). Although only preliminary (this was not repeated), clones 29, 2934, and 242934 displayed modestly improved expression levels, compared to wild-type (28% higher, 26% higher, and 27% higher, respectively). In fact, the only scFv that doesn't express as well or better than 26-10 was clone 24.

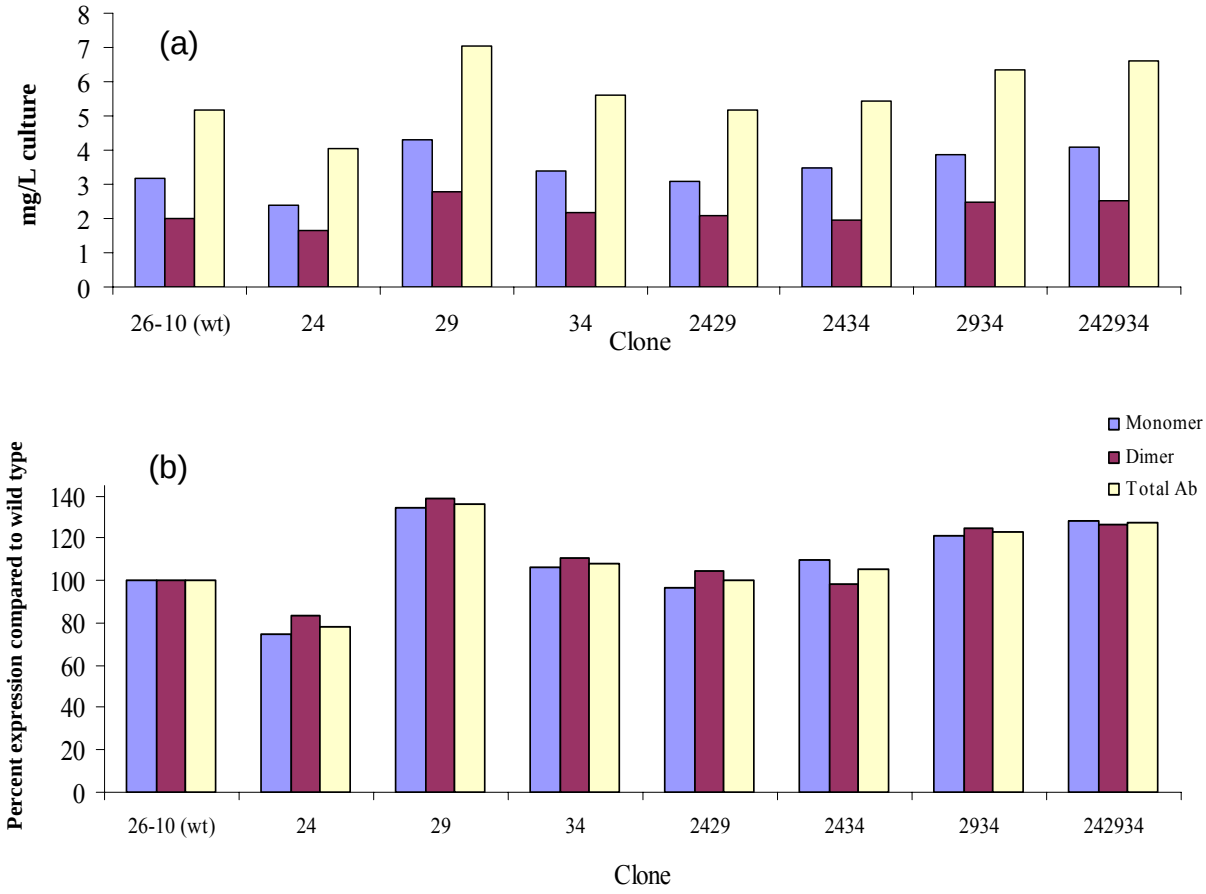


Figure 3-2. Expression levels of 26-10 and mutants.
(a) Expression levels of the wild-type and mutant scAbs (all combinations of mutations involving V_H24(S→F), V_H29(F→S), V_H34(M→V)) determined as mg protein/liter of culture. (b) Percent expression levels of the mutant scAbs compared to the wild-type.

The functional thermostability of each of the isolated antibody variants was assessed *via* ELISA. Aliquots of each concentration-normalized scAb protein were incubated at 25°C for 24, 48, and 72 hours. The binding ability of "25°C incubated" scAb and fresh scAb from each clone was assayed in triplicate using microtitre plates with immobilized digoxin-BSA. Anti-HuCκ-HRP followed by development with TMB substrate was used to detect the amount of bound antibody. This ELISA assays the effect that 25°C incubations have on the activity of the scAb, *i.e.* a measure of functional stability of the

antibody. As seen in Figure 3-3, the antibodies 29 and 2434 show greatly improved short term stability after 24 hours over the wild-type antibody. All other proteins, except for 2429 and 34, show some stability improvements over wild-type after 48 hours at 25°C. After 72 hours at 25°C, the only antibody that does not have improved stability is 2429. 2434 proved to be less stable over the longer term than 29, 24, and 242934, but is more stable than wild-type. Antibody 29 turned out to be the most stable after 72 hours at 25°C.

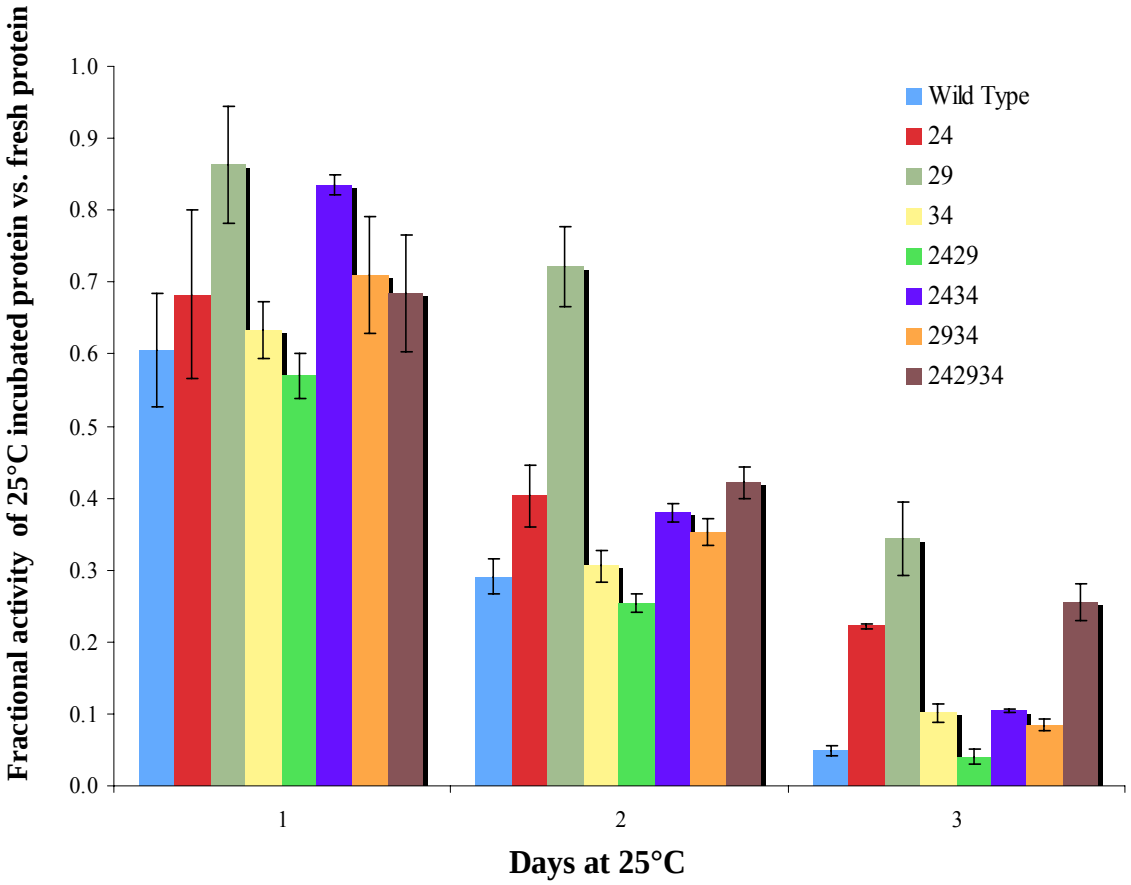


Figure 3-3. Functional stability of 26-10 and mutants.
ELISA was used to determine the binding of each purified monomeric scFv after leaving it at 25°C for up to three days. The binding signals were compared to the signals of fresh protein.

DISCUSSION

The mutational distribution in both high error-rate libraries rejects a Poisson distribution in favor of a PCR-modeled distribution agreeing with the predictions of Arnold and coworkers[9] for high-error rate libraries of the 26-10 scFv. The variance of each library exceeds the mean number of DNA mutations, indicating that library activity follows a PCR-modeled distribution. Based on this observation, it is not surprising that our libraries have more functional variants than if they followed Poisson statistics as previously reported[9].

The PCR-modeled distribution of mutations, which is applied to high-error rate libraries, accesses unusual nucleotide subsitutions[9]. Consistent with this observation, the nucleotide substitutions in the current study of high error-rate libraries display very different mutational spectra when compared to previous studies of low mutation rates[4]. The results of the types of mutations observed, while not wholly consistent between libraries (*e.g.* A→N mutations prevail in L2, G→N mutations prevail in L3), show a prevalence in both libraries for A→T, A→G, and G→A mutations, in agreement with previous reports[4; 28]. Curiously, the most commonly reported and predicted mutations[4; 28] for high mutation rate libraries, A→T, while common, are not the most prevalent in this study. The A→G and G→A mutations are the most prevalent in this, including at the selected mutational hotspots of V_H24, V_H29, and V_H34.

PCR efficiency declines inversely to the rate of mutation because the first generation mutants amplified from the parent gene are persistently copied throughout subsequent cycles and are more likely to retain the wild-type function[9; 29]. Thus, the number of active clones in a high-error rate library is greater than that predicted by the exponential

decline seen in libraries that follow Poisson statistics (*i.e.* low-error rate libraries). In the case of our libraries, PCR efficiency is likely attenuated due to high error-rates, helping to explain why the actual activity of the libraries is 71-fold (L2) and 256-fold (L3) greater than the expected activity.

Coupled substitutions at positions that play little-to-no direct role (*i.e.* contact) in the actual antigen binding can provide antibody enhancements[29], demonstrating the usefulness of libraries generated *via* high error-rates of mutation to discover unpredictable enhancing mutations[9; 27]. Previous studies have suggested that mutations selected from combinatorial antibody libraries based on high-affinity templates often occur in regions away from the contact residues[30] and are selected due to enhancing effects on the conformation of the antibody[31] or by optimizing the positions of the contact residue side chains to interact with the antigen[32]. None of the three positions (V_H24, V_H29, and V_H34) that are commonly mutated in the selected clones correspond to an antigen contact residues in the parent antibody[33], suggesting that in addition to enhancing antigen interactions directly, the mutations at these positions might be selected due to indirect effects. Other selection effects might include global stabilization of the overall protein[34; 35; 36], stabilizing the binding pocket during the initial stages of antigen recognition, or increasing the protein's expression level[29; 37]. The mutations at positions V_H24 and V_H29 of CDRH1 from 26-10 were selected in previous high-error rate mutagenesis selections of 26-10[10]. Since CDRH1 is not considered to play as important of a role in antigen recognition as the other CDRs for the 26-10 antibody[16; 33], selected mutations in this region support further the idea that they are selected based on indirect effects. Both the expression (clones 29, 2934 and 242934) and stability (clones 24, 29,

2434, 2934, & 242934) were found to be modestly enhanced for the mutant scFvs, characteristics that might help explain why these mutations were selected.

The K_D is improved (ranging from 1.5 – 3.5 fold higher) over wild-type for nearly all the common mutations: *i.e.* in clones 24, 29, 2429, 2434, 2934, 242934. There appear to be two paths to the highest affinity, $K_D \approx 1.3 \times 10^{-9}$ M, shared by proteins with either a single mutation at position 29 or coupled mutations at positions 24, 29, and 34. One path, the F→S mutation at V_H29, demonstrates that by increasing the k_a alone, the protein's affinity is increased 3.5-fold. The other path, involving mutations at all three hotspots (V_H24, V_H29, and V_H34), while lacking the optimal k_a seen in the FV_H29S protein, uses enhancements to the k_d to reach $K_D \approx 1.3 \times 10^{-9}$ M. Significantly, these three mutations do not appear to synergize, that is cooperate, to enhance the affinity beyond that seen with the single mutation at position V_H29.

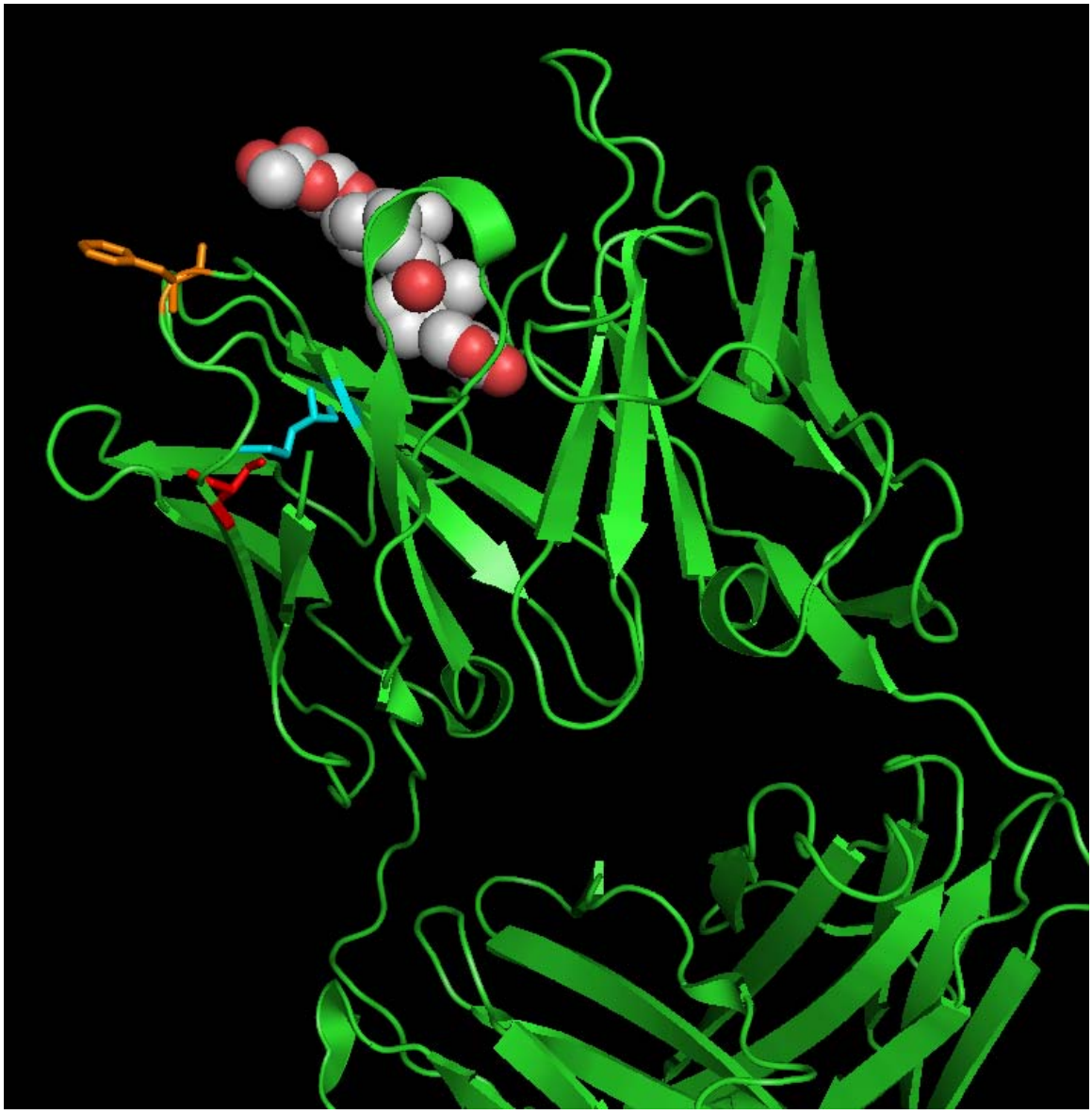


Figure 3-4. Structural model of 26-10 binding pocket.
This ribbon-structure, using the PDB coordinates for file 1IGJ, shows the binding pocket of 26-10 with its antigen, digoxin (grey and red space-filling balls). The side chains of each of V_H positions 24, 29, and 34 are shown in red, orange, and blue, respectively. Note that none of the side chains are oriented towards the binding pocket (figure courtesy of Clint Leysath).

Position V_H24, positioned just prior to the beginning of CDRH1[12; 20], is typically associated with neutral residues[37; 38], *e.g.* a Ser in the wild-type 26-10 scFv, and occasionally hydrophobic residues[37]. Since the side chain at this position points into hydrophobic core of the 26-10 antibody[39] (Figure 3-4), a Ser to the hydrophobic Phe

mutation should be well-tolerated by the protein and by the CDRH1 without drastically perturbing its structure. Variation at this position is accommodated by slight differences in the conformation of the loop region that is outside of the common core of the scFv[37]. Although rare at this position, Phe does naturally occur here in other antibodies[125] and it has been presumed that a hydrophobic residue here will act as a buried anchor point for the CDR1 loop[29; 37]. Through its insertion into the backbone of the scFv, the substitution of a Phe at this position may stabilize the loop of CDRH1 increasing the rigidity of the backbone and the CDR, a role often played by hydrophobic residues on the borders of CDRs[39]. The importance of an anchor at position V_H24 is accentuated in 26-10 because as described below, position V_H29, which often serves as loop anchor, is actually solvent exposed in 26-10 (Figure 3-4), disabling its typical anchoring role. The S→F mutation, selected at position V_H24, enhances the stability of the protein left at 25°C unless coupled with the F→S mutation at position V_H29. Molecular simulations of the V_H24 (S→F) mutation (data not shown) indicate that the aromatic ring of a Phe at V_H24 is near enough the aromatic rings of Tyr V_H27 and/or Phe V_H32 to form either edge-to-face or offset stacked geometry typical of the π-π interactions between aromatic residues[40]. These types of aromatic interactions have been calculated to stabilize proteins[41] offering a explanation for the increase in functional stability for mutant protein 24.

The side chain of another position of a selected hotspot position, V_H34 (M→V), is imbedded within the backbone of the parent antibody (Figure 3-4) and likely serves as a CDRH1 anchor[29; 37; 39]. Of equal hydrophobicity[42], the selected M→V mutation at this imbedded position should also be well-tolerated. Although a molecular simulation of this mutation indicates no change in structure, the insertion into the backbone of the smaller

Val at this position encounters less steric hindrance compared with the wild-type Met, particularly with respect to the side chains of V_H34 and V_H24[33]. As an anchor point for the loop of CDRH1, it is possible that a Val at V_H34 instead of the bulkier Met may further stabilize CDRH1. In addition to enhancing stability, the selected mutations at positions V_H34 and V_H24 apparently affect the nearby contact residues (YV_H33 and NV_H35) in such a way as to slow the rate of dissociation of digoxin (decrease the k_d).

The apex (H28-H31) of the loop of CDRH1 maintains the appropriate hydrophobic residues predicted by the canonical structure paradigm for the loop (described by Chothia *et al.*[39]). However, in the crystal structure, the side chains at positions 28 (Ile) and 29 (Phe) point away from the core of the antibody and are solvent exposed[33; 39]. In most antibodies, the side chain at V_H29 is hydrophobic[39; 43], and is inserted into the protein core constraining the number of loop conformations to the canonical structures[29; 39], but this is *not* the case for the 26-10 scFv[33] (Figure 3-4). The unusual conformations of positions 28 and 29 contribute to the crystallographic disorder of the region spanning positions V_H27-30[16; 33]. Because the wild-type Phe at V_H29 is solvent exposed, and not imbedded in the backbone, its influence over the conformation of CDRH1 is likely limited. Therefore, mutating the Phe→Ser ought to have almost no effect on the conformation of this loop. In the structures of four proteins that have homologous sequences to the CDRH1 region of 26-10, (PDB IDs 1KB5, 1PLG, 1EGJ, and 1YNT) the side chains of the positions homologous to 28 and 29 are indeed pointed toward the core, away from the solvent (personal communication from PD Jeffrey). In the other antibodies, the Phe at position V_H29 typically packs down against CDRH2. In 26-10, the interplay between the two loops of CDRH1 and CDRH2 at this point is attenuated (but

not eliminated), due to the conformation of V_H29, potentially leading to greater flexibility of CDRH2. For example, with the unusual conformation of the Phe V_H29 directed away from CDRH2, the Tyr at position V_H53 at the apex of CDRH2 is less hindered than in the four other homologous sequences. The substitution of Ser at V_H29, would likely promote even fewer electrostatic constraints on Tyr V_H53 enabling more conformational flexibility of CDRH2. CDRH2 would have more flexibility enabling its one contact residue Tyr V_H50 to be slightly closer to the antigen, thereby improving the k_a of the H29 mutants, as observed in this study.

Solvent exposed, hydrophobic regions similar to the unusual hydrophobic patch on CDRH1 of 26-10 (Ile V_H28 and Phe V_H29) are associated with unstable proteins[44] and antibodies[36]. In general, disrupting a solvent exposed patch with a polar side chain (*i.e.* Ser) will make a protein more soluble and stable[45]. Consistent with this observation, when Phe V_H29 is mutated to the non-hydrophobic, polar Ser, the protein is functionally stabilized in this study, in all mutants except for clone 2429. It is difficult to offer an explanation for the poor stability of 2429, especially considering that both mutations alone offer improvements in functional stability. Equally confounding is the result that when added to 2429, 34, a rather innocuous mutation as explained below, improves the stability of the protein.

The selected mutation at V_H34, M→V, is not a significant biochemical change: a large hydrophobic residue replaced with a smaller, hydrophobic residue. Not surprisingly, this mutation, alone, does little to stabilize the protein. However, when used in combination with 24, the protein does become modestly more stable than wild-type. As the side chains of V_H24 and V_H34 are practically opposed to one another (Figure 3-4),

the smaller Val may allow more room for the bulky Phe at V_H24, a new feature that may lead to greater stability. This mutation might allow for the Phe at V_H24 to bury further and form stronger π-π interactions, as described above.

Based on the structure of the complex between 26-10 and digoxin[33] (Figure 3-4), it is unlikely that these selected mutations enable any new contacts between the mutated residues and the antigen. However, new contacts at other positions or tighter contacts to wild-type contact residues cannot be ruled out. Taken individually, these substitutions may affect the conformation of the non-binding pocket face of CDRH1 and affect the conformational flexibility of CDRH2. Together, these substitutions may alter the anchor points of CDRH1 leading to slightly improved affinity, stability, and expression of the variant clones.

It is noteworthy that the positions identified as consensus mutations (V_H24, V_H29, V_H34) map to a region of the crystal structure that contains high B-factors, indicating disorder in this region[33]. Although no functional link has been established in this work, the data presented can be interpreted as being consistent with the notion that the coupled mutations somehow provide increased order at this position of the structure, providing an entropic advantage that enables higher affinity antibody binding. Our results support previous studies that described the CDRH1 as being plastic, consistent with the disordered crystal structure[16].

High mutation rate libraries may be the most direct method to identify coupled or synergistic mutations, both in and out of the binding pocket, that lead to an enhanced protein[9]. In the final analysis, high-error rate mutagenesis was used to create two separate libraries based on the gene 26-10. Both libraries were able to produce enhanced

antibodies at a higher frequency than predicted based on an extrapolation of lower error-rate library activity *vs*. error-rate studies. High-error rate libraries are useful to unlock the constraints put upon *in vivo* derived Abs by the germ line V sequences, as in the case of 26-10. Separate screenings of both libraries isolated clones that, when sequenced, displayed identical mutational hotspots at positions V_H24, V_H29, and V_H34.

At the end of these studies, it must be stated that the data were not as conclusive as one might have hoped. Although suggestive of modest effects, it is not clear that the identified consensus mutations, when made individually and within groups, tell the entire story with regard to affinity/expression/stability enhancement of the 26-10 scFv subjected to high error-rates. Importantly, these consensus mutations were always found in the context of additional mutations that were more stochastic in nature. It might be that these additional mutations hold the key to unraveling the full consensus mutation story, a story whose telling must await much future work.

EXPERIMENTAL PROTOCOLS

Strains, Plasmids, and Reagents

Plasmids pB30D and pMoPac16 are described above and elsewhere[21; 46; 47]. *E. coli* strain LMG194 was used for all library screening by Dr. Loo, as previously described[10] while the ABLE C strain was used for expression of purified scFvs, as described below. Oligonucleotide primers were synthesized by Invitrogen (Carlsbad, CA) and New England Biolabs restriction enzymes were used. Digoxin-BODIPY was synthesized as described [131].

Creating the common mutants

The 26-10 antibody was mutated using site-directed mutagenesis *via* splicing by overlap extension PCR described in detail by Horton *et al*.[48] to create the mutants described in Table 3-4. After the PCR, each antibody gene was digested with *Sfi*I, gel purified, ligated into *Sfi*I cut pMoPac16, and transformed into *E. coli* ABLE C. The transformed cells were plated on LB Media (Difco, BD Biosciences, Sparks, MD) supplemented with 2% glucose and ampicillin (100μg/mL). The sequences of the antibody genes were confirmed with bi-directional DNA sequencing of plasmids prepared (Qiagen, Valencia, CA) from randomly selected transformed colonies.

scAb protein expression and purification

scAb antibody proteins were produced in soluble form and purified from the periplasm as described[21]. Briefly, *E. coli* ABLE C cells containing the appropriate pMoPac derivatives were cultured from single colonies in 2mL of Terrific Broth (TB) (Difco, BD Biosciences, Sparks, MD) supplemented with 2% glucose and 200μg/mL ampicillin for 10h at 37°C. This entire culture was then used to inoculate a 40mL culture of TB-2% Gluc + AMP_{200} and the cells were incubated overnight at 25°C. The next day, each 40mL culture was added to 400mL of TB + AMP_{200} which incubated for 2 hours at 37°C. The cultures were shifted to 25°C for 30 minutes and IPTG (Sigma-Aldrich, St. Louis MO) was added to a final concentration of 1mM. The cells were incubated at 25°C for 4 hours and pelleted (5000g × 10min). The periplasmic fraction was isolated by resuspending the cell pellets in 10mL ice-cold 100mM Tris-HCl, pH 7.4 containing 0.75M Sucrose and 1mL 10mg/mL lysozyme (Sigma-Aldrich). 20mL of 1mM EDTA

was added, the cells were incubated for 15 minutes and then 1.4mL 0.5M $MgCl_2$ was added for another 15 min. The periplasmic fraction was separated from the spheroplasts by centrifuging the mixture at 12,000g for 30 min. An equal volume of immobilized metal affinity chromatography buffer (IMAC; 20mM Na_2HPO_4, 500mM NaCl, 10mM imidazole, 0.1% Tween-20, pH 7.4) was added followed by 0.5mL washed Ni-NTA Agarose (Qiagen) for 1 hour. The resin was washed twice in IMAC buffer and elution was carried out with 4mL 250mM imidazole in IMAC buffer. Monomeric scFv was further purified by size exclusion gel filtration chromatography on a Superdex 200 column (Amersham-Pharmacia Biosciences, Uppsala, Sweden).

Expression levels of functional scFv

Polyacrylamide gel electrophoresis using 4-20% gradient Tris-glycine gels (Cambrex, East Rutherford, New Jersey) was used to analyze 20μL of each 0.5mL antibody containing fraction from the size exclusion gel filtration chromatography under reducing conditions. The gels were stained with GelCode Blue, Pierce Biotech, Rockford, IL) to distinguish which fractions contained scAb monomer and which contained dimer. A bicinchoniric acid (BCA, Pierce Biotech) assay was used to determine protein concentrations of both the monomer and dimer scFv. These concentrations were extrapolated linearly to give the expression levels of monomeric and dimeric scAb in terms of mg/L of culture.

Protein stability analysis

50μL/well of 4μg/mL of digoxin-BSA conjugate was immobilized on high-binding Costar 3590 microtitre plates (Corning). Standard ELISA blocking and washing procedures were used throughout[49]. Normalized concentrations of mutant and wild type scAb were left at 25°C for 24, 48, and 72 hours and added to the wells along with fresh scAb to determine the decay of each scAb's binding ability. Binding of antibody fragments was detected with 100μL/well of 1:30,000 dilution of anti-His-HRP (Sigma). The plates were developed with aTMB substrate kit (Pierce Biotech), quenched with 5N sulfuric acid and read at 450nm. The optical density at 450nm of each 25°C incubated scAb was compared to the optical density at 450nm of the respective fresh scAb. The results for the stability of each scAb are expressed as a fraction of its binding ability compared to binding ability of the fresh scAb.

Surface plasmon resonance (SPR) kinetic analysis of single chain antibodies

250 response units (RUs) of the digoxin-BSA was immobilized on the carboxy-dextran surface of CM5 sensor chip (BIAcore, Piscataway, NJ) using EDC/NHS activated amine coupling[21]. Hank's buffered saline (HBS) was used as the running and dilution buffer for all kinetics experiments and the flow rate was kept constant at 100μL/min. scAb proteins in triplicate in one independent trial per clone over a concentration range of 100pM to 100nM were injected onto the surface of the chip to determine the association and dissociation rates of the proteins. Antibodies were eluted and the chip was regenerated with 50% ethylene glycol-HBS, pH 10.0 using a flow rate of 30μL/min. BIAeval software calculated the association/dissociation rates and

ultimately the equilibrium constant (K_D) of the antibody proteins to digoxin. k_a was determined from a plot of the [ln(dR/dt)]/t versus concentration while k_d values were detemined from a plot of ln(R_O/R) versus time.

REFERENCES

1. Cadwell, R. C. & Joyce, G. F. (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**, 28-33.
2. Beckman, R. A., Mildvan, A. S. & Loeb, L. A. (1985). On the fidelity of DNA replication: manganese mutagenesis in vitro. *Biochemistry* **24**, 5810-7.
3. Eckert, K. A. & Kunkel, T. A. (1991). DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Appl* **1**, 17-24.
4. Shafikhani, S., Siegel, R. A., Ferrari, E. & Schellenberger, V. (1997). Generation of large libraries of random mutants in Bacillus subtilis by PCR-based plasmid multimerization. *Biotechniques* **23**, 304-10.
5. Arnold, F. H. (1998). Enzyme engineering reaches the boiling point. *Proc Natl Acad Sci U S A* **95**, 2035-6.
6. Suzuki, M., Christians, F. C., Kim, B., Skandalis, A., Black, M. E. & Loeb, L. A. (1996). Tolerance of different proteins for amino acid diversity. *Mol Divers* **2**, 111-8.
7. Zaccolo, M. & Gherardi, E. (1999). The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J Mol Biol* **285**, 775-83.
8. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science* **247**, 1306-10.
9. Drummond, D. A., Iverson, B. L., Georgiou, G. & Arnold, F. H. (2005). Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol* **350**, 806-16.
10. Daugherty, P. S., Chen, G., Iverson, B. L. & Georgiou, G. (2000). Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proc Natl Acad Sci U S A* **97**, 2029-34.
11. Burks, E. A., Chen, G., Georgiou, G. & Iverson, B. L. (1997). In vitro scanning saturation mutagenesis of an antibody binding pocket. *Proc Natl Acad Sci U S A* **94**, 412-7.
12. Chen, G., Dubrawsky, I., Mendez, P., Georgiou, G. & Iverson, B. L. (1999). In vitro scanning saturation mutagenesis of all the specificity determining residues in an antibody binding site. *Protein Eng* **12**, 349-56.
13. Brown, M., Rittenburg, M. B., Chen, C. & Roberts, V. A. (1996). Tolerance of single, but not multiple, amino acid replacements in antibody VH CDR 2: a

means of minimizing B cell wastage from somatic hypermutation? *J Immunol* **156**, 3285-91.
14. Ito, W., Iba, Y. & Kurosawa, Y. (1993). Effects of substitutions of closely related amino acids at the contact surface in an antigen-antibody complex on thermodynamic parameters. *J Biol Chem* **268**, 16639-47.
15. Wong, Y. W., Kussie, P. H., Parhami-Seren, B. & Margolies, M. N. (1995). Modulation of antibody affinity by an engineered amino acid substitution. *J Immunol* **154**, 3351-8.
16. Short, M. K., Jeffrey, P. D., Kwong, R. F. & Margolies, M. N. (1995). Contribution of antibody heavy chain CDR1 to digoxin binding analyzed by random mutagenesis of phage-displayed Fab 26-10. *J Biol Chem* **270**, 28541-50.
17. Chen, C., Roberts, V. A., Stevens, S., Brown, M., Stenzel-Poore, M. P. & Rittenberg, M. B. (1995). Enhancement and destruction of antibody function by somatic mutation: unequal occurrence is controlled by V gene combinatorial associations. *Embo J* **14**, 2784-94.
18. Ramirez-Benitez, M. C. & Almagro, J. C. (2001). Analysis of antibodies of known structure suggests a lack of correspondence between the residues in contact with the antigen and those modified by somatic hypermutation. *Proteins* **45**, 199-206.
19. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & et al. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci U S A* **85**, 5879-83.
20. Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. & Curtiss, R., 3rd. (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* **15**, 29-34.
21. Daugherty, P. S., Chen, G., Olsen, M. J., Iverson, B. L. & Georgiou, G. (1998). Antibody affinity maturation using bacterial surface display. *Protein Eng* **11**, 825-32.
22. Sun, F. (1995). The polymerase chain reaction and branching processes. *J Comput Biol* **2**, 63-86.
23. Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B. L. & Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. *J Immunol Methods* **276**, 185-96.
24. Hayhurst, A. & Harris, W. J. (1999). Escherichia coli skp chaperone coexpression improves solubility and phage display of single-chain antibody fragments. *Protein Expr Purif* **15**, 336-43.
25. Harvey, B. R., Georgiou, G., Hayhurst, A., Jeong, K. J., Iverson, B. L. & Rogers, G. K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. *Proc Natl Acad Sci U S A* **101**, 9193-8.
26. Ayala, M., Balint, R. F., Fernandez-de-Cossio, L., Canaan-Haden, J. W., Larrick, J. W. & Gavilondo, J. V. (1995). Variable region sequence modulates periplasmic

export of a single-chain Fv antibody fragment in Escherichia coli. *Biotechniques* **18**, 832, 835-8, 840-2.
27. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. & Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* **296**, 57-86.
28. Moore, G. L. & Maranas, C. D. (2000). Modeling DNA mutation and recombination for directed evolution experiments. *J Theor Biol* **205**, 483-503.
29. Bond, C. J., Wiesmann, C., Marsters, J. C., Jr. & Sidhu, S. S. (2005). A structure-based database of antibody variable domain diversity. *J Mol Biol* **348**, 699-709.
30. Garcia-Rodriguez, C., Levy, R., Arndt, J. W., Forsyth, C. M., Razai, A., Lou, J., Geren, I., Stevens, R. C. & Marks, J. D. (2007). Molecular evolution of antibody cross-reactivity for two subtypes of type A botulinum neurotoxin. *Nat Biotechnol* **25**, 107-16.
31. Foote, J. & Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. *Proc Natl Acad Sci U S A* **91**, 10370-4.
32. Mian, I. S., Bradwell, A. R. & Olson, A. J. (1991). Structure, function and properties of antibody binding sites. *J Mol Biol* **217**, 133-51.
33. Jeffrey, P. D., Strong, R. K., Sieker, L. C., Chang, C. Y., Campbell, R. L., Petsko, G. A., Haber, E., Margolies, M. N. & Sheriff, S. (1993). 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc Natl Acad Sci U S A* **90**, 10310-4.
34. Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B. L. & Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nat Biotechnol* **25**, 563-5.
35. Wark, K. L. & Hudson, P. J. (2006). Latest technologies for the enhancement of antibody affinity. *Adv Drug Deliv Rev* **58**, 657-70.
36. Worn, A. & Pluckthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* **305**, 989-1010.
37. Chothia, C., Gelfand, I. & Kister, A. (1998). Structural determinants in the sequences of immunoglobulin variable domain. *J Mol Biol* **278**, 457-79.
38. Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). Structural repertoire of the human VH segments. *J Mol Biol* **227**, 799-817.
39. Chothia, C. & Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* **196**, 901-17.
40. Waters, M. L. (2002). Aromatic interactions in model systems. *Curr Opin Chem Biol* **6**, 736-41.
41. Carver, F. J., Hunter, C. A., Jones, P. S., Livingstone, D. J., McCabe, J. F., Seward, E. M., Tiger, P. & Spey, S. E. (2001). Quantitative measurements of edge-to-face aromatic interactions by using chemical double-mutant cycles. *Chemistry* **7**, 4854-62.
42. Black, S. D. & Mould, D. R. (1991). Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal Biochem* **193**, 72-82.

43. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J Biol Chem* **252**, 6609-16.
44. Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Adv Protein Chem* **14**, 1-63.
45. Shea, J. E. & Brooks, C. L., 3rd. (2001). From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu Rev Phys Chem* **52**, 499-535.
46. Hayhurst, A. & Georgiou, G. (2001). High-throughput antibody isolation. *Curr Opin Chem Biol* **5**, 683-9.
47. Daugherty, P. S., Olsen, M. J., Iverson, B. L. & Georgiou, G. (1999). Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. *Protein Eng* **12**, 613-21.
48. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-8.
49. Chen, G., Hayhurst, A., Thomas, J. G., Harvey, B. R., Iverson, B. L. & Georgiou, G. (2001). Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol* **19**, 537-42.

Chapter 4: Conclusion and future directions

CONCLUSION

The work presented in this dissertation describes different strategies of creating diversity in single antibody scaffold genes which are ultimately used to construct antibody libraries. These libraries were screened for either the *de novo* isolation of antibodies to two different peptides or for enhancements to a previously isolated antibody (26-10) which binds with high-affinity to digoxin. The first two approaches, described in Chapter 2, used rationally designed randomization with the objective of introducing diversity biased towards peptide binding. These approaches were validated by isolating medium to high-affinity, specific single chain antibody fragments (scFv) using phage display libraries. Specific antibodies were isolated from the M:anti-pep library with equilibrium dissociation constants (K_Ds) as low as 14 nM to angiotensin (AT) and 18 nM to neuropeptide Y (NPY), while the K_Ds of the highest affinity antibodies isolated from the from the H:anti-pep were 11 μM to AT and 49 nM to NPY.

One of the main objectives in the design of both of the anti-peptide libraries was to use scaffolds and diversity that resulted in a library of sequences that were more relevant to binding peptides than some of the other "single-pot" antibody libraries that often provide antibodies to peptides with weak or non-specific binding. Three main factors were considered in the design of each library: 1) which antibody gene to use as the library scaffold, 2) which positions would be targeted for diversification, 3) and what type of diversity would be incorporated at each position. Structural features of anti-peptide antibodies were identified and used to select the 26-10 antibody as the scaffold gene for the M:anti-pep library. These same structural considerations were not used in

the selection of the scaffold gene for the Hu:anti-pep library. The scaffold of this library is based on human germline variable genes and has been commonly used in other synthetic antibody libraries. Positions in the V_H regions of both library scaffolds were chosen for diversification based on their high usage as peptide contacting residues in other anti-peptide antibodies. In the M:anti-pep library, these positions were diversified using site-directed mutagenesis to reflect the natural diversity at each position in an alignment of other anti-peptide antibodies. In the Hu:anti-pep library, these positions were diversified using site-directed mutagenesis to incorporate all 20 amino acids. The targeted mutagenesis was employed to limit the theoretical diversity, thus limiting the amount of sequence space to be surveyed in the library selections. Also, by targeting only potential antigen contact residues, the anti-peptide structural features were left intact and deleterious effects on expression and stability of the library proteins were likely limited. The results indicate that the more rigorous approach of the M:anti-pep library provides antibodies of higher-affinity. In the case of either library, the results are impressive in that neither starting scaffold antibody binds to the target peptides and the potential contact residues of only one of the two variable regions have been diversified.

Rather than using targeted mutagenesis and tailored diversity, the approach used for creating diversity to enhance the 26-10 scFv, described in Chapter 3, employed whole gene randomization *via* error-prone PCR at error rates which are typically viewed as detrimentally high. Despite error-rates that were expected to result in non-functional libraries, the bacterial surface displayed libraries were active and contained clones that were easily isolated using FACS-based selections. The sequences of clones isolated from the libraries contained several selected consensus mutations along with dozens of more

widely distributed mutations. Clones with combinations of the consensus mutations were analyzed to determine if the selected mutations cooperate to enhance the affinity, expression, and stability of the antibody. The results indicate that some of the mutations may provide higher affinity, higher expression levels, or enhanced stability. If lower-error rates were employed, the sequences in the library would be less likely to have multiple mutations. Therefore, coupled, enhancing mutations would be more difficult to identify than with high-error rate libraries.

FUTURE DIRECTIONS

The isolated anti-peptide clones could be easily affinity matured by randomizing positions in the V_L regions that correspond to peptide contacting residues in other anti-peptide antibodies and screening these V_L libraries against the respective peptides. The anti-peptide libraries can be used as a source of monoclonal antibodies to countless other peptides. By screening the libraries against additional peptides, the sequences of all isolated clones can be compared to refine the positions to be diversified and the types of diversity to introduce in any second generation libraries. Positions with conserved mutations need not be diversified in subsequent libraries because they might be important for maintaining a peptide-binding structure, expression levels, or overall stability. This type of refinement could increase the fraction of well-expressed clones relevant to peptide binding in the libraries. By limiting the amount of total sequence space in the library, fewer rounds of selection might be required to isolate high-affinity binders. The model could be further refined by determining and comparing the atomic co-structures of both high and low affinity binders bound to the peptide antigens This type of analysis

would provide insight into which contacts are the most important for binding and if there are any mutations which position the hypervariable loops for high-affinity interactions.

Variations of the original M:anti-pep and Hu:anti-pep libraries, such as V_L libraries (with tailored diversity in only the V_L region with a fixed V_H) or V_H/V_L libraries (with tailored diversity in both the V_L and V_H regions) could also be created and screened. Antibodies from these libraries could be affinity matured by shuffling the V_H regions of the highest affinity antibodies from V_H region libraries with their counterparts from the V_L region libraries and screening them for higher affinity antibodies.

Bibliography

Almagro, J.C. (2004). Identification of differences in the specificity-determining residues of antibodies that recognize antigens of different size: implications for the rational design of antibody repertoires. J Mol Recognit *17*, 132-143.

Almagro, J.C., Hernandez, I., Ramirez, M.C., and Vargas-Madrazo, E. (1998). Structural differences between the repertoires of mouse and human germline genes and their evolutionary implications. Immunogenetics *47*, 355-363.

Almagro, J.C., Quintero-Hernandez, V., Ortiz-Leon, M., Velandia, A., Smith, S.L., and Becerril, B. (2006). Design and validation of a synthetic VH repertoire with tailored diversity for protein recognition. J Mol Recognit *19*, 413-422.

Almagro, J.C., Vargas-Madrazo, E., Zenteno-Cuevas, R., Hernandez-Mendiola, V., and Lara-Ochoa, F. (1995). VIR: a computational tool for analysis of immunoglobulin sequences. Biosystems *35*, 25-32.

Alzari, P.M., Lascombe, M.B., and Poljak, R.J. (1988). Three-dimensional structure of antibodies. Annual review of immunology *6*, 555-580.

Amzel, L., and Poljak, R. (1979). Three-Dimensional Structure of Immunoglobulins. Annual Review of Biochemistry *48*, 961-997.

Arnold, F.H. (1998). Enzyme engineering reaches the boiling point. Proc Natl Acad Sci U S A *95*, 2035-2036.

Ayala, M., Balint, R.F., Fernandez-de-Cossio, L., Canaan-Haden, J.W., Larrick, J.W., and Gavilondo, J.V. (1995). Variable region sequence modulates periplasmic export of a single-chain Fv antibody fragment in Escherichia coli. Biotechniques *18*, 832, 835-838, 840-832.

Ball, W.J., Jr., Wang, Z., Malik, B., Kasturi, R., Dey, P., Short, M.K., and Margolies, M.N. (2000). Selection of peptidic mimics of digoxin from phage-displayed peptide libraries by anti-digoxin antibodies. J Mol Biol *301*, 101-115.

Barbas, C.F., 3rd, Amberg, W., Simoncsits, A., Jones, T.M., and Lerner, R.A. (1993). Selection of human anti-hapten antibodies from semisynthetic libraries. Gene *137*, 57-62.

Beckman, R.A., Mildvan, A.S., and Loeb, L.A. (1985). On the fidelity of DNA replication: manganese mutagenesis in vitro. Biochemistry *24*, 5810-5817.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S., and Whitlow, M. (1988). Single-chain antigen-binding proteins. Science *242*, 423-426.

Black, S.D., and Mould, D.R. (1991). Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. Anal Biochem *193*, 72-82.

Boder, E.T., and Wittrup, K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol *15*, 553-557.

Bond, C.J., Wiesmann, C., Marsters, J.C., Jr., and Sidhu, S.S. (2005). A structure-based database of antibody variable domain diversity. J Mol Biol *348*, 699-709.

Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A., and Sauer, R.T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. Science *247*, 1306-1310.

Brett, P.J., Tiwana, H., Feavers, I.M., and Charalambous, B.M. (2002). Characterization of oligopeptides that cross-react with carbohydrate-specific antibodies by real time kinetics, in-solution competition enzyme-linked immunosorbent assay, and immunological analyses. J Biol Chem *277*, 20468-20476.

Brown, M., Rittenburg, M.B., Chen, C., and Roberts, V.A. (1996). Tolerance of single, but not multiple, amino acid replacements in antibody VH CDR 2: a means of minimizing B cell wastage from somatic hypermutation? J Immunol *156*, 3285-3291.

Burks, E.A., Chen, G., Georgiou, G., and Iverson, B.L. (1997). In vitro scanning saturation mutagenesis of an antibody binding pocket. Proc Natl Acad Sci U S A *94*, 412-417.

Cadwell, R.C., and Joyce, G.F. (1992). Randomization of genes by PCR mutagenesis. PCR Methods Appl *2*, 28-33.

Carver, F.J., Hunter, C.A., Jones, P.S., Livingstone, D.J., McCabe, J.F., Seward, E.M., Tiger, P., and Spey, S.E. (2001). Quantitative measurements of edge-to-face aromatic interactions by using chemical double-mutant cycles. Chemistry *7*, 4854-4862.

Casadevall, A. (2002). Passive antibody administration (immediate immunity) as a specific defense against biological weapons. Emerging infectious diseases *8*, 833-841.

Casadevall, A., and Scharff, M.D. (1995). Return to the past: the case for antibody-based therapies in infectious diseases. Clin Infect Dis *21*, 150-161.

Chames, P., Hoogenboom, H.R., and Henderikx, P. (2002). Selection of antibodies against biotinylated antigens. Methods Mol Biol *178*, 147-157.

Charalambous, B.M., and Feavers, I.M. (2000). Peptide mimics elicit antibody responses against the outer-membrane lipooligosaccharide of group B neisseria meningitidis. FEMS Microbiol Lett *191*, 45-50.

Charalambous, B.M., and Feavers, I.M. (2001). Mimotope vaccines. J Med Microbiol *50*, 937-939.

Chen, C., Roberts, V.A., Stevens, S., Brown, M., Stenzel-Poore, M.P., and Rittenberg, M.B. (1995). Enhancement and destruction of antibody function by somatic mutation: unequal occurrence is controlled by V gene combinatorial associations. Embo J *14*, 2784-2794.

Chen, G., Dubrawsky, I., Mendez, P., Georgiou, G., and Iverson, B.L. (1999). In vitro scanning saturation mutagenesis of all the specificity determining residues in an antibody binding site. Protein Eng *12*, 349-356.

Chen, G., Hayhurst, A., Thomas, J.G., Harvey, B.R., Iverson, B.L., and Georgiou, G. (2001). Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). Nat Biotechnol *19*, 537-542.

Chothia, C., Gelfand, I., and Kister, A. (1998). Structural determinants in the sequences of immunoglobulin variable domain. J Mol Biol *278*, 457-479.

Chothia, C., and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. Embo J *5*, 823-826.

Chothia, C., and Lesk, A.M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol *196*, 901-917.

Chothia, C., Lesk, A.M., Gherardi, E., Tomlinson, I.M., Walter, G., Marks, J.D., Llewelyn, M.B., and Winter, G. (1992). Structural repertoire of the human VH segments. J Mol Biol *227*, 799-817.

Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., *et al.* (1989). Conformations of immunoglobulin hypervariable regions. Nature *342*, 877-883.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D., and Winter, G. (1991). Making antibody fragments using phage display libraries. Nature *352*, 624-628.

Collis, A., Brouwer, B., and Martin, A. (2003). Analysis of the antigen combining site: correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. J Mol Biol *325*, 337-354.

Daugherty, P.S., Chen, G., Iverson, B.L., and Georgiou, G. (2000). Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. Proc Natl Acad Sci U S A *97*, 2029-2034.

Daugherty, P.S., Chen, G., Olsen, M.J., Iverson, B.L., and Georgiou, G. (1998). Antibody affinity maturation using bacterial surface display. Protein Eng *11*, 825-832.

Daugherty, P.S., Olsen, M.J., Iverson, B.L., and Georgiou, G. (1999). Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. Protein Eng *12*, 613-621.

de Haard, H.J., van Neer, N., Reurs, A., Hufton, S.E., Roovers, R.C., Henderikx, P., de Bruine, A.P., Arends, J.W., and Hoogenboom, H.R. (1999). 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.

de Wildt, R.M., Hoet, R.M., van Venrooij, W.J., Tomlinson, I.M., and Winter, G. (1999). Analysis of heavy and light chain pairings indicates that receptor editing shapes the human antibody repertoire. J Mol Biol *285*, 895-901.

Drummond, D.A., Iverson, B.L., Georgiou, G., and Arnold, F.H. (2005). Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. J Mol Biol *350*, 806-816.

Dumont, Y., Martel, J.C., Fournier, A., St-Pierre, S., and Quirion, R. (1992). Neuropeptide Y and neuropeptide Y receptor subtypes in brain and peripheral tissues. Progress in neurobiology *38*, 125-167.

Eckert, K.A., and Kunkel, T.A. (1991). DNA polymerase fidelity and the polymerase chain reaction. PCR Methods Appl *1*, 17-24.

El-Zaim, H.S., Chopra, A.K., Peterson, J.W., Vasil, M.L., and Heggers, J.P. (1998). Protection against exotoxin A (ETA) and Pseudomonas aeruginosa infection in mice with ETA-specific antipeptide antibodies. Infect Immun *66*, 5551-5554.

Elzaim, H.S., Chopra, A.K., Peterson, J.W., Goodheart, R., and Heggers, J.P. (1998). Generation of neutralizing antipeptide antibodies to the enzymatic domain of Pseudomonas aeruginosa exotoxin A. Infect Immun *66*, 2170-2179.

Feldhaus, M.J., Siegel, R.W., Opresko, L.K., Coleman, J.R., Feldhaus, J.M., Yeung, Y.A., Cochran, J.R., Heinzelman, P., Colby, D., Swers, J., *et al.* (2003). Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nat Biotechnol *21*, 163-170.

Fellouse, F.A., Barthelemy, P.A., Kelley, R.F., and Sidhu, S.S. (2006). Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. J Mol Biol *357*, 100-114.

Fellouse, F.A., Wiesmann, C., and Sidhu, S.S. (2004). Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. Proc Natl Acad Sci U S A *101*, 12467-12472.

Foote, J., and Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. Proc Natl Acad Sci U S A *91*, 10370-10374.

Garcia-Rodriguez, C., Levy, R., Arndt, J.W., Forsyth, C.M., Razai, A., Lou, J., Geren, I., Stevens, R.C., and Marks, J.D. (2007). Molecular evolution of antibody cross-reactivity for two subtypes of type A botulinum neurotoxin. Nat Biotechnol *25*, 107-116.

Georgiou, G., Stathopoulos, C., Daugherty, P.S., Nayak, A.R., Iverson, B.L., and Curtiss, R., 3rd (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. Nat Biotechnol *15*, 29-34.

Green, L.L., Hardy, M.C., Maynard-Currie, C.E., Tsuda, H., Louie, D.M., Mendez, M.J., Abderrahim, H., Noguchi, M., Smith, D.H., Zeng, Y.*, et al.* (1994). Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. Nature genetics *7*, 13-21.

Griffiths, A., Williams, S., Hartley, O., Tomlinson, I., Waterhouse, P., Crosby, W., Kontermann, R., Jones, P., Low, N., Allison, T.*, et al.* (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. Embo J *13*, 3245–3260.

Griswold, K.E., Kawarasaki, Y., Ghoneim, N., Benkovic, S.J., Iverson, B.L., and Georgiou, G. (2005). Evolution of highly active enzymes by homology-independent recombination. Proc Natl Acad Sci U S A *102*, 10082-10087.

Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H.R., and Pluckthun, A. (1998). Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. Proceedings of the National Academy of Sciences of the United States of America *95*, 14130-14135.

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

Hanes, J., Schaffitzel, C., Knappik, A., and Pluckthun, A. (2000). Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. Nat Biotechnol *18*, 1287-1292.

Harvey, B.R., Georgiou, G., Hayhurst, A., Jeong, K.J., Iverson, B.L., and Rogers, G.K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. Proc Natl Acad Sci U S A *101*, 9193-9198.

Hawkins, R.E., Russell, S.J., and Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. J Mol Biol *226*, 889-896.

Hayhurst, A., and Georgiou, G. (2001). High-throughput antibody isolation. Curr Opin Chem Biol *5*, 683-689.

Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B.L., and Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. J Immunol Methods *276*, 185-196.

Hayhurst, A., and Harris, W.J. (1999). Escherichia coli skp chaperone coexpression improves solubility and phage display of single-chain antibody fragments. Protein Expr Purif *15*, 336-343.

Hoet, R.M., Cohen, E.H., Kent, R.B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H.*, et al.* (2005). Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. Nat Biotechnol *23*, 344-348.

Hoogenboom, H.R. (2005). Selecting and screening recombinant antibody libraries. Nat Biotechnol *23*, 1105-1116.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene *77*, 61-68.

Hunter, M.M., Margolies, M.N., Ju, A., and Haber, E. (1982). High-affinity monoclonal antibodies to the cardiac glycoside, digoxin. J Immunol *129*, 1165-1172.

Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J., and Lerner, R.A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science *246*, 1275-1281.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R.*, et al.* (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc Natl Acad Sci U S A *85*, 5879-5883.

Ito, W., Iba, Y., and Kurosawa, Y. (1993). Effects of substitutions of closely related amino acids at the contact surface in an antigen-antibody complex on thermodynamic parameters. J Biol Chem *268*, 16639-16647.

Jeffrey, P.D., Strong, R.K., Sieker, L.C., Chang, C.Y., Campbell, R.L., Petsko, G.A., Haber, E., Margolies, M.N., and Sheriff, S. (1993). 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. Proc Natl Acad Sci U S A *90*, 10310-10314.

Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Pluckthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. Proc Natl Acad Sci U S A *98*, 75-80.

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

Kabat, E.A., Padlan, E.A., and Davies, D.R. (1975). Evolutionary and structural influences on light chain constant (CL) region of human and mouse immunoglobulins. Proc Natl Acad Sci U S A *72*, 2785-2788.

Kabat, E.A., and Wu, T.T. (1971). Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. Ann N Y Acad Sci *190*, 382-393.

Kabat, E.A., Wu, T.T., and Bilofsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. J Biol Chem *252*, 6609-6616.

Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. Adv Protein Chem *14*, 1-63.

Keller, M.A., and Stiehm, E.R. (2000). Passive immunity in prevention and treatment of infectious diseases. Clinical microbiology reviews *13*, 602-614.

Khazaeli, M.B., Conry, R.M., and LoBuglio, A.F. (1994). Human immune response to monoclonal antibodies. J Immunother Emphasis Tumor Immunol *15*, 42-52.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A., and Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. J Mol Biol *296*, 57-86.

Knappik, A., and Pluckthun, A. (1995). Engineered turns of a recombinant antibody improve its in vivo folding. Protein Eng *8*, 81-89.

Kobayashi, N., Soderlind, E., and Borrebaeck, C.A. (1997). Analysis of assembly of synthetic antibody fragments: expression of functional scFv with predefined specificity. Biotechniques *23*, 500-503.

Kohler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. Nature *256*, 495-497.

Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H.R., and Pluckthun, A. (1997). 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.

Krebs, B., Rauchenberger, R., Reiffert, S., Rothe, C., Tesar, M., Thomassen, E., Cao, M., Dreier, T., Fischer, D., Hoss, A.*, et al.* (2001). High-throughput generation and engineering of recombinant human antibodies. J Immunol Methods *254*, 67-84.

Lang, A.B., Cryz, S.J., Jr., Schurch, U., Ganss, M.T., and Bruderer, U. (1993). Immunotherapy with human monoclonal antibodies. Fragment A specificity of polyclonal and monoclonal antibodies is crucial for full protection against tetanus toxin. J Immunol *151*, 466-472.

Lee, C.V., Liang, W.C., Dennis, M.S., Eigenbrot, C., Sidhu, S.S., and Fuh, G. (2004). High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. J Mol Biol *340*, 1073-1093.

Lippow, S.M., Wittrup, K.D., and Tidor, B. (2007). Computational design of antibody-affinity improvement beyond in vivo maturation. Nat Biotechnol.

Lonberg, N. (2005). Human antibodies from transgenic animals. Nat Biotechnol *23*, 1117-1125.

Lonberg, N., Taylor, L.D., Harding, F.A., Trounstine, M., Higgins, K.M., Schramm, S.R., Kuo, C.C., Mashayekh, R., Wymore, K., McCabe, J.G.*, et al.* (1994). Antigen-specific human antibodies from mice comprising four distinct genetic modifications. Nature *368*, 856-859.

Low, N.M., Holliger, P.H., and Winter, G. (1996). Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. J Mol Biol *260*, 359-368.

Luginbuhl, B., Kanyo, Z., Jones, R.M., Fletterick, R.J., Prusiner, S.B., Cohen, F.E., Williamson, R.A., Burton, D.R., and Pluckthun, A. (2006). Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation. J Mol Biol *363*, 75-97.

MacCallum, R.M., Martin, A.C.R., and Thornton, J.M. (1996). Antibody-antigen Interactions: Contact Analysis and Binding Site Topography. J Mol Biol *262*, 732-745.
Marks, J.D., and Bradbury, A. (2004). Selection of human antibodies from phage display libraries. Methods Mol Biol *248*, 161-176.
Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D., and Winter, G. (1991). By-passing immunization. Human antibodies from V-gene libraries displayed on phage. J Mol Biol *222*, 581-597.
Martin, A.C., and Thornton, J.M. (1996). Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. J Mol Biol *263*, 800-815.
Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T., and Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. The Journal of experimental medicine *188*, 2151-2162.
Mattheakis, L.C., Bhatt, R.R., and Dower, W.J. (1994). An in vitro polysome display system for identifying ligands from very large peptide libraries. Proc Natl Acad Sci U S A *91*, 9022-9026.
Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B.L., and Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. Nat Biotechnol *25*, 563-565.
Medzhitov, R., and Janeway, C., Jr. (2000a). Innate immune recognition: mechanisms and pathways. Immunological reviews *173*, 89-97.
Medzhitov, R., and Janeway, C., Jr. (2000b). Innate immunity. The New England journal of medicine *343*, 338-344.
Mian, I.S., Bradwell, A.R., and Olson, A.J. (1991). Structure, function and properties of antibody binding sites. J Mol Biol *217*, 133-151.
Milstein, C. (1999). The hybridoma revolution: an offshoot of basic research. Bioessays *21*, 966-973.
Mirick, G.R., Bradt, B.M., Denardo, S.J., and Denardo, G.L. (2004). A review of human anti-globulin antibody (HAGA, HAMA, HACA, HAHA) responses to monoclonal antibodies. Not four letter words. Q J Nucl Med Mol Imaging *48*, 251-257.
Moczydlowski, E., Lucchesi, K., and Ravindran, A. (1988). An emerging pharmacology of peptide toxins targeted against potassium channels. The Journal of membrane biology *105*, 95-111.
Monks, S.A., Karagianis, G., Howlett, G.J., and Norton, R.S. (1996). Solution structure of human neuropeptide Y. Journal of biomolecular NMR *8*, 379-390.
Moore, G.L., and Maranas, C.D. (2000). Modeling DNA mutation and recombination for directed evolution experiments. J Theor Biol *205*, 483-503.
Morea, V., Lesk, A.M., and Tramontano, A. (2000). Antibody modeling: implications for engineering and design. Methods *20*, 267-279.
Morea, V., Tramontano, A., Rustici, M., Chothia, C., and Lesk, A.M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. J Mol Biol *275*, 269-294.
Ohlin, M., and Borrebaeck, C.A. (1996). Characteristics of human antibody repertoires following active immune responses in vivo. Mol Immunol *33*, 583-592.

Orlandi, R., Gussow, D.H., Jones, P.T., and Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc Natl Acad Sci U S A *86*, 3833-3837.
Padlan, E.A. (1977). Structural implications of sequence variability in immunoglobulins. Proc Natl Acad Sci U S A *74*, 2551-2555.
Padlan, E.A. (1979). Evaluation of the structural variation among light chain variable domains. Mol Immunol *16*, 287-296.
Padlan, E.A., Abergel, C., and Tipper, J.P. (1995). Identification of specificity-determining residues in antibodies. Faseb J *9*, 133-139.
Padlan, E.A., and Davies, D.R. (1975). Variability of three-dimensional structure in immunoglobulins. Proc Natl Acad Sci U S A *72*, 819-823.
Persson, H., Lantto, J., and Ohlin, M. (2006). A focused antibody library for improved hapten recognition. J Mol Biol *357*, 607-620.
Phillips, M.I., Speakman, E.A., and Kimura, B. (1993). Levels of angiotensin and molecular biology of the tissue renin angiotensin systems. Regulatory peptides *43*, 1-20.
Presta, L.G. (2006). Engineering of therapeutic antibodies to minimize immunogenicity and optimize function. Advanced drug delivery reviews *58*, 640-656.
Queen, C., Schneider, W.P., Selick, H.E., Payne, P.W., Landolfi, N.F., Duncan, J.F., Avdalovic, N.M., Levitt, M., Junghans, R.P., and Waldmann, T.A. (1989). A humanized antibody that binds to the interleukin 2 receptor. Proc Natl Acad Sci U S A *86*, 10029-10033.
Raag, R., and Whitlow, M. (1995). Single-chain Fvs. Faseb J *9*, 73-80.
Ramachandiran, V., Kramer, G., and Hardesty, B. (2000). Expression of different coding sequences in cell-free bacterial and eukaryotic systems indicates translational pausing on Escherichia coli ribosomes. FEBS letters *482*, 185-188.
Ramirez-Benitez, M.C., and Almagro, J.C. (2001). Analysis of antibodies of known structure suggests a lack of correspondence between the residues in contact with the antigen and those modified by somatic hypermutation. Proteins *45*, 199-206.
Regoli, D., Boudon, A., and Fauchere, J.L. (1994). Receptors and antagonists for substance P and related peptides. Pharmacological reviews *46*, 551-599.
Ribnicky, B., Van Blarcom, T., and Georgiou, G. (2007). A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. J Mol Biol *369*, 631-639.
Rojas, G., Almagro, J.C., Acevedo, B., and Gavilondo, J.V. (2002). Phage antibody fragments library combining a single human light chain variable region with immune mouse heavy chain variable regions. J Biotechnol *94*, 287-298.
Roth, J. (1973). Peptide hormone binding to receptors: a review of direct studies in vitro. Metabolism: clinical and experimental *22*, 1059-1073.
Rowley, M.J., O'Connor, K., and Wijeyewickrema, L. (2004). Phage display for epitope determination: a paradigm for identifying receptor-ligand interactions. Biotechnol Annu Rev *10*, 151-188.
Rusche, J.R., Javaherian, K., McDanal, C., Petro, J., Lynn, D.L., Grimaila, R., Langlois, A., Gallo, R.C., Arthur, L.O., Fischinger, P.J., *et al.* (1988). Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. Proc Natl Acad Sci U S A *85*, 3198-3202.

Shafikhani, S., Siegel, R.A., Ferrari, E., and Schellenberger, V. (1997). Generation of large libraries of random mutants in Bacillus subtilis by PCR-based plasmid multimerization. Biotechniques *23*, 304-310.

Shea, J.E., and Brooks, C.L., 3rd (2001). From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. Annu Rev Phys Chem *52*, 499-535.

Shirai, H., Kidera, A., and Nakamura, H. (1999). H3-rules: identification of CDR-H3 structures in antibodies. FEBS letters *455*, 188-197.

Short, M.K., Jeffrey, P.D., Kwong, R.F., and Margolies, M.N. (1995). Contribution of antibody heavy chain CDR1 to digoxin binding analyzed by random mutagenesis of phage-displayed Fab 26-10. J Biol Chem *270*, 28541-28550.

Sidhu, S.S., and Fellouse, F.A. (2006). Synthetic therapeutic antibodies. Nat Chem Biol *2*, 682-688.

Sidhu, S.S., Li, B., Chen, Y., Fellouse, F.A., Eigenbrot, C., and Fuh, G. (2004). Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. J Mol Biol *338*, 299-310.

Slade, H.B. (1994). Human Immunoglobulins for intravenous use and hepatitis C viral transmission. Clinical and diagnostic laboratory immunology *1*, 613-619.

Sobel, J.H., Trakht, I., Pileggi, N., and Qi Wu, H. (1998). Antipeptide monoclonal antibodies to defined fibrinogen Aalpha chain regions: anti-Aalpha 487-498, a structural probe for fibrinogenolysis. Blood *91*, 1590-1598.

Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A.M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C.*, et al.* (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. Nat Biotechnol *18*, 852-856.

Spyroulias, G.A., Nikolakopoulou, P., Tzakos, A., Gerothanassis, I.P., Magafa, V., Manessi-Zoupa, E., and Cordopatis, P. (2003). Comparison of the solution structures of angiotensin I & II. Implication for structure-function relationship. European journal of biochemistry / FEBS *270*, 2163-2173.

Stemmer, W.P. (1994). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Natl Acad Sci U S A *91*, 10747-10751.

Stresser, D.M., and Kupfer, D. (1999). Monospecific antipeptide antibody to cytochrome P-450 2B6. Drug metabolism and disposition: the biological fate of chemicals *27*, 517-525.

Sun, F. (1995). The polymerase chain reaction and branching processes. J Comput Biol *2*, 63-86.

Suzuki, M., Christians, F.C., Kim, B., Skandalis, A., Black, M.E., and Loeb, L.A. (1996). Tolerance of different proteins for amino acid diversity. Molecular diversity *2*, 111-118.

Swanson, J.A., and Hoppe, A.D. (2004). The coordination of signaling during Fc receptor-mediated phagocytosis. Journal of leukocyte biology *76*, 1093-1103.

Tiwana, H., Clow, K.J., Hall, C., Feavers, I.M., and Charalambous, B.M. (2005). The immunogenicity of a conformationally restricted peptide mimetic of meningococcal lipooligosaccharide. Scand J Immunol *62*, 385-392.

Tomlinson, I.M., Cox, J.P., Gherardi, E., Lesk, A.M., and Chothia, C. (1995). The structural repertoire of the human V kappa domain. Embo J *14*, 4628-4638.

van Dijk, M.A., and van de Winkel, J.G. (2001). Human antibodies as next generation therapeutics. Curr Opin Chem Biol *5*, 368-374.
Van Regenmortel, M.H. (2002). Reductionism and the search for structure-function relationships in antibody molecules. J Mol Recognit *15*, 240-247.
Vargas-Madrazo, E., Almagro, J.C., and Lara-Ochoa, F. (1995a). Structural repertoire in VH pseudogenes of immunoglobulins: comparison with human germline genes and human amino acid sequences. J Mol Biol *246*, 74-81.
Vargas-Madrazo, E., Lara-Ochoa, F., and Almagro, J.C. (1995b). Canonical structure repertoire of the antigen-binding site of immunoglobulins suggests strong geometrical restrictions associated to the mechanism of immune recognition. J Mol Biol *254*, 497-504.
Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J., and Johnson, K.S. (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. Nat Biotechnol *14*, 309-314.
Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. Nature *341*, 544-546.
Wark, K.L., and Hudson, P.J. (2006). Latest technologies for the enhancement of antibody affinity. Advanced drug delivery reviews *58*, 657-670.
Waters, M.L. (2002). Aromatic interactions in model systems. Curr Opin Chem Biol *6*, 736-741.
White-Scharf, M.E., Potts, B.J., Smith, L.M., Sokolowski, K.A., Rusche, J.R., and Silver, S. (1993). Broadly neutralizing monoclonal antibodies to the V3 region of HIV-1 can be elicited by peptide immunization. Virology *192*, 197-206.
Wilson, I.A., and Stanfield, R.L. (1994). Antibody-antigen interactions: new structures and new conformational changes. Curr Opin Struct Biol *4*, 857-867.
Winter, G., Griffiths, A.D., Hawkins, R.E., and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. Annual review of immunology *12*, 433-455.
Wong, Y.W., Kussie, P.H., Parhami-Seren, B., and Margolies, M.N. (1995). Modulation of antibody affinity by an engineered amino acid substitution. J Immunol *154*, 3351-3358.
Worn, A., and Pluckthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. J Mol Biol *305*, 989-1010.
Xu, J.L., and Davis, M.M. (2000). Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity *13*, 37-45.
Xu, L., Aha, P., Gu, K., Kuimelis, R.G., Kurz, M., Lam, T., Lim, A.C., Liu, H., Lohse, P.A., Sun, L., *et al.* (2002). Directed evolution of high-affinity antibody mimics using mRNA display. Chem Biol *9*, 933-942.
Yamaguchi, Y., Kim, H., Kato, K., Masuda, K., Shimada, I., and Arata, Y. (1995). Proteolytic fragmentation with high specificity of mouse immunoglobulin G. Mapping of proteolytic cleavage sites in the hinge region. J Immunol Methods *181*, 259-267.
Zaccolo, M., and Gherardi, E. (1999). The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. J Mol Biol *285*, 775-783.

Zahnd, C., Spinelli, S., Luginbühl, B., Amstutz, P., Cambillau, C., and Plückthun, A. (2004). Directed in Vitro Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. J Biol Chem *279*, 18870-18877.

Zhang, W., Feng, J., Li, Y., Guo, N., and Shen, B. (2005). Humanization of an anti-human TNF-alpha antibody by variable region resurfacing with the aid of molecular modeling. Mol Immunol *42*, 1445-1451.

Vita

Christian Wessel Cobaugh was born in St. Louis, Missouri, on December 30, 1971, the son of Charles Heilmann Cobaugh and Julia Wessel Cobaugh. After graduating from Ladue Horton Watkins High School, St. Louis, in 1990, Christian studied for two years at Duke University, Durham, North Carolina, before transferring to the University of Texas at Austin. Christian graduated from the University of Texas at Austin in May of 1994 receiving a Bachelor of Arts in Anthropology. Christian was awarded his Master of Science in Molecular Pharmacology from Northwestern University, Chicago, Illinois in 1999 and worked as a developmental scientist for JD Searle Pharmaceuticals. Christian began the graduate program in Cell and Molecular Biology at the University of Texas at Austin in August, 2001. He currently resides in Redwood City, California with his wife, Amy, and daughter, Celeste and works as a staff scientist for KaloBios, Inc., Palo Alto, California.

Permanent address: 3722 Red Oak Way, Redwood City, California, 94061

This dissertation was typed by the author. All figures and tables were created by the author unless noted otherwise.