Construction of a Synthetic Human V_L Phage Display Library and Isolation of Potential Neuropilin-1-specific V_L Therapeutics from the Library

Artine Keklikian

Thesis Submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the M.Sc. degree in Microbiology & Immunology

Department of Biochemistry, Microbiology & Immunology

Faculty of Medicine

University of Ottawa

ABSTRACT

Antibody phage display technology mimics the natural immune system, and has been widely used for rapid isolation of single-domain antibodies (sdAbs) with various binding specificities and affinities in the micromolar to low nanomolar range. SdAbs are the variable regions of immunoglobulins (e.g., V_H, V_L, V_HH) and serve as potential probes with therapeutic value. The small size, high solubility, high expression and stability, and high specificity and affinity for the cognate antigen, make sdAbs ideal in improving drug delivery and the overall therapeutic value of antibodies. The main objective of this thesis was to construct a large V_L phage display library (~10^{10} diversity); analyze it via sequence analysis, and to subtractively pan the library for isolation of Neuropilin-1 (NRP1)-specific V_Ls. Neuropilin-1 (NRP1), a cell-surface receptor for both vascular endothelial growth factor (VEGF) and class 3 Semaphorins (Sema3A), contributes to neuron cell death through its interaction with Sema3A in stroke patients. Disruption of this NRP1-Sema3A interaction would allow for axonal outgrowth and neuron regeneration in the area of the brain affected by stroke. Construction of the synthetic phage antibody library utilized a single V_L framework with selected positions in the complementarity-determining regions (CDRs) targeted for randomization *in vitro* using synthetic oligonucleotides that introduced sequence degeneracy. Specific V_Ls were then selected from the repertoire through subtractive panning against a cell line endogenously expressing NRP1 (PC12) as well as a negative cell line that does not express NRP1 (HEK293) with competitive elution carried out using a synthetic Sema3A-derived peptide. Fifteen V_L clones were isolated, cloned in *E. coli*, expressed and purified, and of these, nine were determined to be non-aggregating by size exclusion chromatography. Further studies will determine the potential therapeutic use of these V_L sdAbs as agents in recovery from stroke and neuron degeneration.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Jamshid Tanha, for having given me the opportunity to work on a project at the NRC and for the patient guidance, constructive comments and advice he has provided throughout my time as his student. I am also deeply grateful to the entire Antibody Engineering group for their friendship and assistance in introducing me to the various laboratory procedures, and for helping to make my stay at NRC a memorable experience. Many thanks to Amy Aylsworth, Susan Jiang, and Sheng T. Hou, for preparing and providing the cell lines that were used as part of this project. Also, thank you to Rebecca To for cloning of the "V_L-24" template into the Fd-tet vector, Henk van Faassen for assistance with flow cytometry, Dr. Dae Young Kim for protocols on TG1 electrocompetent cell preparation, protein expression and protein purification, and Sonia Leclerc for DNA sequencing. I owe my loving gratitude to my mom and dad for their unending encouragement and support, and also to my three brothers who will always be my best of friends.

TABLE OF CONTENTS

I. LIST OF ABBREVIATIONS

Ab	Antibody
Ag	Antigen
Amp	Ampicillin
AP	Alkaline Phosphatase
BBB	Blood-Brain Barrier
bp	Base Pair(s)
BSA	Bovine Serum Albumin
C_H	Constant Domain of the Heavy Chain
C_L	Constant Domain of the Light Chain
CDR	Complementarity-Determining Region
CFU	Colony-Forming Unit
Chlor	Chloramphenicol
CT	C-terminal
Da	Dalton
D-MEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphate
DRG	Dorsal Root Ganglion
dsDNA	Double-Stranded DNA
DTT	Dithiothreitol
ELISA	Enzyme-Linked Immunosorbent Assay
Fab	Antigen-Binding Fragment

FBS	Fetal Bovine Serum
Ff	Filamentous Bacteriophage
FPLC	Fast Protein Liquid Chromatography
FR	Framework Region
HEK293	Human Embryonic Kidney-293
HEPES	*N*-(2-hydroxyethyl)-piperazine-*N*-2-ethanesulfonic acid
IG	Intergenic
Ig	Immunoglobulin
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
kbp	Kilo-Base Pair(s)
kDa	KiloDaltons
mAb	Monoclonal Antibody
MAM	Meprin/A5-protein/PTPmu
MW	Molecular Weight
NAR	New Antigen Receptor
NIP	Neuropilin Interacting Protein
NRP	Neuropilin
OD_{280}	Optical Density at 280 nm
OD_{600}	Optical Density at 600 nm
pIII	Minor Coat Protein of Filamentous Phage
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline

PC12	Pheochromocytoma-12
PCR	Polymerase Chain Reaction
PE	Phycoerythrin
PEG	Polyethylene Glycol
pI	Isoelectric Point
PMSG	Phenylmethylsulphonyl Fluoride
PS	Packaging Signal
pVIII	Major Coat Protein of Filamentous Phage
RF	Replicative Form
scFv	Single-Chain Variable Fragment
sdAb	Single Domain Antibody
SDS	Sodium Dodecyl Sulfate
SEC	Size Exclusion Chromatography
Sema3A	Semaphorin-3A
SOC	Super Optimal Catabolite
SPR	Surface Plasmon Resonance
ssDNA	Single-Stranded DNA
Tet	Tetracyclin
Tm	Melting Temperature
VEGF	Vascular Endothelial Cell Growth Factor
V_H	Heavy Chain Variable Domain of Conventional Abs
V_HH	Variable domain of Heavy-Chain Abs
V_L	Light Chain Variable Domain of Conventional Abs

III. LIST OF TABLES

III. LIST OF FIGURES

IV. INTRODUCTION

Antibodies (Abs), or immunoglobulins (Igs), are gamma globulin proteins generated by the immune system of vertebrates that identify and bind to a vast array of antigens (Ags). The high affinity and binding specificity for a target Ag, make Abs valuable agents in biomedical applications. *In vitro* studies examining Ag-specific Abs began with the introduction of hybridoma technology in 1975 whereby stable "hybridoma" cell lines were produced via fusion of Ag-stimulated B-cells with myeloma cells, resulting in the production of Ag-specific monoclonal Abs (mAbs) (Kohler and Milstein, 1975). Advancements in molecular biology and immunology have since allowed for ameliorated and efficient methods in the development and isolation of mAb fragments. For example, the use of phage display technology in conjunction with structure-function databases (Winter et al., 1994; Harding et al., 2004) offers a novel approach in developing specific mAb fragments.

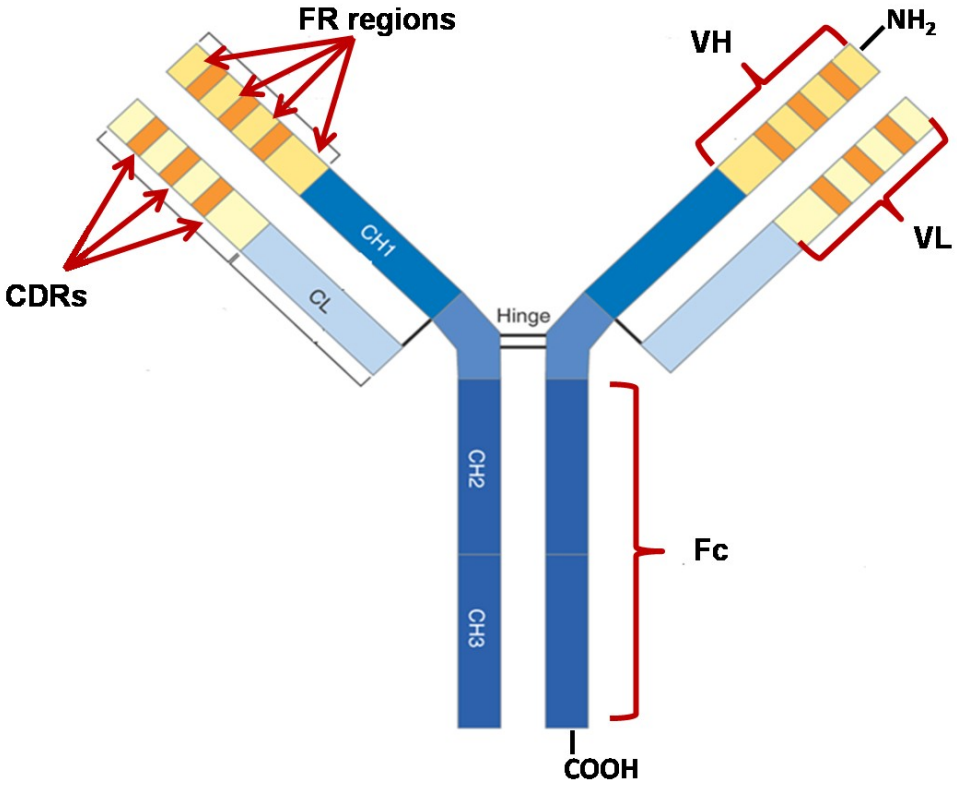
1. Immunoglobulin structure and function

The humoral immune response primarily involves production of Igs, or Abs, by plasma cells in response to the detection of foreign molecules. There are 5 major classes or isotypes of Igs: IgG, IgA, IgM, IgD, and IgE. The most abundant class of these Igs in the blood is IgG (~73%), which has a molecular weight (MW) of 150 kDa. IgG is found in plasma and external secretions, and is expressed on the surface of the B-cell membrane. IgG is composed of 4 different polypeptide chains: two heavy chains (approximately 440 amino acids each) and two light (approximately 220 amino acids each) chains. Light chains are characterized by having one variable and one constant domain (V_L and C_L, respectively), while heavy chains consist of one variable and three constant domains (V_H, C_{H1}, C_{H2} and C_{H3}, respectively) (Figure 1); each of

these chains is a combination of α-helices and β-pleated sheets, with the four chains linked together by disulphide bonds (Cys-Cys) and numerous non-covalent interactions.

The epitope, or antigenic determinant, is the site on an Ag that is recognized by a small portion of the Ab (15-22 amino acids) called the paratope (Goldsby et al., 2003). Epitopes are normally derived from non-self proteins, although host sequences can also be recognized. The region defining the binding specificity of an Ab is referred to as the variable domain and is located at the N terminal end of each chain. Variable domains are divided into three complementarity-determining regions (CDRs), or hypervariable loops, and four framework regions (FRs) (Figure 1). The CDRs confer both Ab specificity and affinity for a target Ag and form the major portion of the Ag-binding site; it is in fact the CDR3 of both V_L and V_H that is mostly responsible for Ag contacts in Ab-Ag complexes (Chothia et al., 1989; Benhar, 2007). The CDR3 of the heavy chain (V_H-CDR3) is also the most diverse loop in composition and length (with an estimated potential diversity of 10^{23} sequences) and is most central to the Ag-binding site of all CDRs (Sanz, 1991). FRs separate the CDRs and consist of stable amino acid sequences, forming a beta-sheet structure which serves as a scaffold to hold the CDRs in position to contact Ag (Figure 1). The fragment crystallizable (Fc) region consists of the heavy chain constant domains (C_{H2} + C_{H3}) and mediates effector functions (e.g., activation of the complement pathway, binding of neutrophils and macrophages for promotion of phagocytosis). Abs can be expressed on the surface of filamentous phage such as Fd-tet or M13 in the form of sdAbs (separate V_H and V_L domains), which retain binding characteristics of the parental Ab.

Figure 1: Monomeric structure of an immunoglubulin. The conventional immunoglobulin monomer (or Ab) consists of two light chains, each characterized by having one variable and one constant domain, V_L and C_L, respectively) and two heavy chains, each composed of one variable and three constant domains (V_H, C_{H1}, C_{H2} and C_{H3}, respectively). Variable domains are further subdivided into 3 complementarity-determining regions (CDRs) and 4 framework regions (FRs). CDRs are responsible for providing both specificity and affinity of an Ab for its target Ag, and FRs, which separate the CDRs, form a beta-sheet structure that functions as a scaffold to position the CDRs to properly contact Ag (adapted from Rother et al., 2007).


FR regions
VH
NH2
CDRs
CH1
CL
VL
Hinge
CH2
CH3
Fc
COOH

a. Types of antibody fragments/domains

The generation of Ag-specific mAb fragments derived from conventional full-size Abs first occurred in the late 1980s soon after the genetic mechanism associated with Ab gene rearrangement was elucidated. Subsequent advancements made in subcloning and expression strategies, in addition to the introduction of polymerase chain reaction (PCR) technology, allowed for the expression of Ab fragments within bacterial hosts (Skerra and Plückthun, 1988; Winter and Milstein, 1991). Conventional methods of producing mAb fragments involves isolating mRNA from hybridoma, spleen cells, or lymph node leukocytes, followed by reverse transcription PCR (RT-PCR) to form complementary DNA (cDNA) and PCR amplification with gene-specific primers to obtain a complete Ab gene sequence (Maynard and Georgiou, 2000). A wide assortment of Ab fragments have been isolated, characterized, and applied in a number of biological systems, including the Ag-binding fragment (Fab), the single-chain variable fragment (scFv), and single domain Ab (sdAb).

b. Fab fragment

The Fab, or antigen-binding fragment, (Better et al., 1988) is defined as a large (MW of approximately 50-55 kDa) heterodimeric molecule consisting the light chain (i.e., C_L + V_L domains) of a parent mAb paired with a second chain of V_H + C_{H1} (Jeffrey et al., 1993) (Figure 2). Fabs are more stable and less prone to proteolytic degradation than other Ab fragments (e.g., scFv). The high stability of Fabs is attributed to a covalent disulphide bond which aids in retaining tertiary structural characteristics and provides non-covalent interactions between the paired chains (Maynard and Georgiou, 2000). Despite their large size in comparison with other Ab fragment types, Fabs can still be used in standard Ab engineering techniques to improve

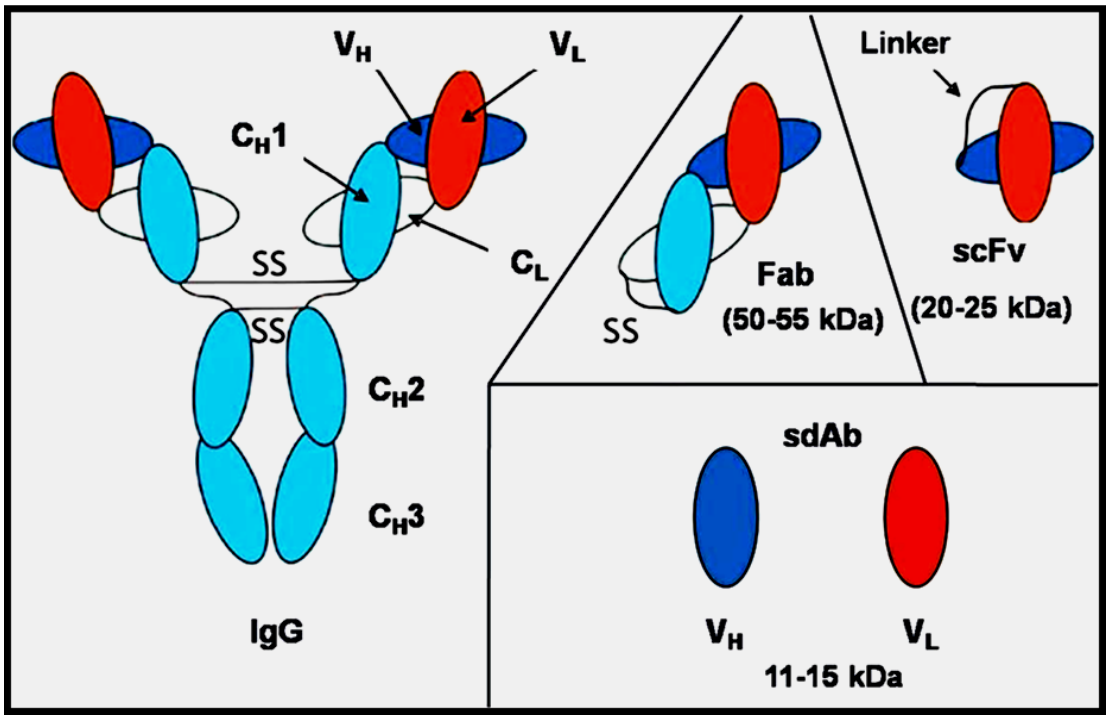
aspects of biological function, affinity for target Ag, etc. (Maynard and Georgiou, 2000; Holliger and Hudson, 2005; Filpula, 2007).

c. scFv fragment

The single-chain variable fragment (scFv) is defined as an Ab fragment in which the V_H and V_L of a parent mAb are connected by a flexible peptide linker (Bird et al., 1989; Huston et al., 1988) (Figure 2). It appears that there is a high degree of variability in terms of binding affinity, stability, and expression efficiency for the scFv, in contrast to larger Ab fragment formats (e.g., Fab). For example, one group reported a scFv with a binding affinity approximately two times less than the parent mAb to its target Ag, while another group using the same target Ag found no difference in affinity between their scFv and mAb (Choi et al., 2004; Wang et al., 2007). Post-translational folding and subsequent protein function may be affected by the expression vector used, growth conditions, and the type of host used. Other limitations to scFv stability include the formation of insoluble inclusion bodies and a high occurrence of linker peptide degradation (Choi et al., 2004). The length of the peptide linker is generally 15 amino acids with a $(Gly_4Ser)_3$ configuration (Huston et al., 1996) and can influence the valency of scFv fragments (Robinson and Sauer, 1998; Turner et al., 1997). For example, it has been shown that three-residue linkers tend to form scFv dimers while linkers with one or two residues favour the formation of trimers (Atwell et al., 1999).

Figure 2: Antibody fragments derived from the immunoglobulin monomer. Fab (antigen-binding fragment), and scFv (single chain variable fragment). Variable light (V_L) and heavy (V_H) domains are illustrated as white and black ovals, respectively. Constant regions of heavy (C_{H1-3}) and light (C_L) chains are denoted as shaded and dotted ovals, respectively. ABD = antigen-binding domain.

The V_H and V_L domains represent the variable domains and are responsible for binding target Ags. The scFv is the combination of V_H and V_L domains, connected by a linker peptide. The constant region of the IgG is composed of C_{H1}, C_{H2}, and C_{H3} domains for the heavy chain and C_L for the light chain. These regions provide immunological effector functions to the IgG. The combination of $V_H + C_{H1}$ and $V_L + C_L$ is known as the Fab fragment, while the combination of $C_{H2} + C_{H3}$ comprise the Fc domain. scFv, Fab, Fv, as well as the single domains V_H and V_L, represent the different Ab fragments commonly used in selection experiments (adapted from Azzazy and Highsmith, 2002).


V_H
V_L
C_H1
SS
SS
C_L
C_H2
C_H3
IgG
SS
Fab
(50-55 kDa)
Linker
scFv
(20-25 kDa)
sdAb
V_H
V_L
11-15 kDa

d. Single domain antibodies

A single domain Ab (sdAb) is an Ab fragment consisting of a single monomeric variable domain which retains the original Ag-binding characteristics of the full parental Ab (Burton, 1995) (Figure 2). With a MW of only 11-15 kDa, sdAbs are the smallest intact Ag-binding fragments known. There are a number of advantages to using sdAbs instead of the other larger and more complex Ab fragments (e.g., Fab or scFv). First, expression of sdAbs in hosts such as *E. coli* frequently produces high yields of soluble, stable protein with low occurrences of protein aggregation and proteolytic degradation (Arbabi-Ghahroudi et al., 2005). Experiments have also shown that the function and stability of sdAbs are maintained in the presence of harsh environmental conditions such as elevated temperatures, high detergent and reducing agent concentrations, or denaturing conditions (Ladenson et al., 2006; de Marco, 2011). SdAbs also have the unique ability to bind active sites of enzymes that are not typically recognized by conventional Ab fragments (Desmyter et al., 1996; Desmyter et al., 2002). Furthermore, they can be formatted into larger molecules to create drugs with prolonged serum half-lives or other pharmacological activities (Revets et al., 2005).

The first sdAbs were isolated from heavy chain Abs (HCAbs) found in the *Camelidae* family, referred to as V_HH, or the variable domain of Camelidae family HCAbs (Burton, 1995). The Camelidae family represents an atypical form, as Camelids (e.g., camels, llamas and alpacas) are the only known living creatures to possess HCAbs, which inherently lack light chains and consist only of V_H, C_{H2} and C_{H3} domains, in addition to the more familiar IgGs (Hamers-Casterman et al., 1993; Genst et al., 2006). Despite the absence of V_L and V_H combinatorial diversity, single domain V_HH fragments have demonstrated high affinity binding to a wide range of Ag types (Gao et al., 1999; Muyldermans, 2001). HCAbs are believed to

compensate for the loss of V_L domain through extended CDR3 loop regions and a higher rate of somatic hypermutations to form a V_HH paratope that can form a large Ag-binding repertoire as well as penetrating into clefts and cavities of Ag epitopes (Muyldermans, 2001). IgNAR (immunoglobulin new antigen receptor) heavy chain Abs, from which sdAbs called V_{NAR} fragments can be isolated, have been found in cartilaginous fish.

2. Antibody phage display

Ab phage display technology is an *in vitro* technique in which the display of Abs on filamentous bacteriophage particles permits the isolation of mAb clones that exhibit high affinity and stability, from a large repertoire of unique clones (Forrer et al., 1999; Russel et al., 2004). In this approach, the phenotype of a phage surface-displayed Ab is directly linked to the genotype encoding that Ab packaged within the same phage particle, permitting Ab phage libraries to be subjected to a selection process whereby recovered clones are identified by sequencing and amplified for further rounds of selection (Russel et al., 2004; Lowman, 2007). A major advantage to displaying Abs on phage is that repertoires in the size range of 10^{10}-10^{11} can be realized in *E. coli. Also, c*haracterization of isolated clones can be carried out using a variety of means, including surface plasmon resonance (SPR) analysis, sequential magnetic-activated cell sorting (MACS), and fluorescence-activated cell sorting (FACS) (Bessette *et al.,* 2004).

Phage display was reported for the first time in the mid-1980s when peptides fused to the pIII coat protein on the surface of Ff were successfully expressed (Smith, 1985). Only a few years later, phage display was used to obtain Ag-specific scFv clones from a scFv display repertoire (McCafferty et al., 1990). However, there are a few limitations to this type of display, including the potential for unpredictable expression bias that may potentially occur against

selected eukaryotic protein sequences expressed in *E. coli*. The incorporation of any protein fusion into the phage particle relies mostly on the ability of *E. coli* to express that protein in soluble form.

a. Filamentous bacteriophage

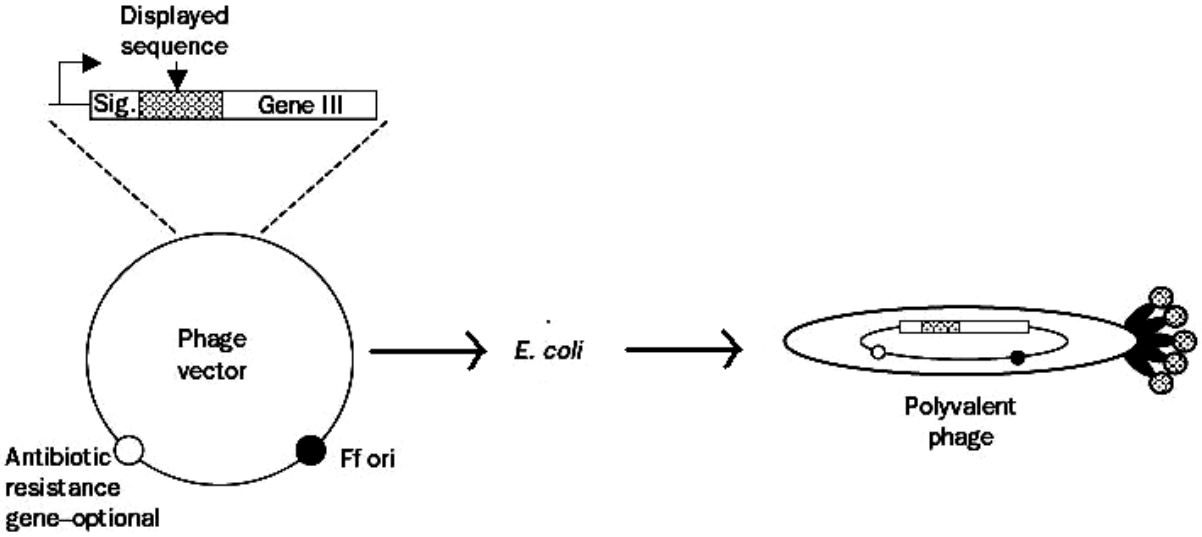
Filamentous bacteriophage (Ff) are non-lytic viral agents that can infect and replicate within Gram-negative bacteria, such as *E. coli*, without destroying the host cell (Russel et al., 2004). Among the advantages to using of Ff as cloning agents are the accommodation of genomic insertions without disruption to their packaging within phage, the possible use of both single stranded and double stranded forms of their DNA (ssDNA and dsDNA, respectively) in cloning and library construction procedures, the fact that the size of Ff is not limited by the DNA they contain, and the retention of original phage infectivity despite modifications made to their coat proteins for subsequent display of proteins (Pande et al., 2010). A beneficial and useful characteristic of Ff is that bacterial host infection does not result in cell lysis, meaning high yields of phage can accumulate within the infected host (Carmen and Jermutus, 2002). The major structural features of Ff are a thin cylindrical shape, a length of approximately 900 nm, a fixed diameter of 6.5 nm, and a circular ssDNA genome encapsulated within a long, hollow, flexible tube (Webster, 1996; Yau et al., 2003).

b. Phage vectors

Phage vectors are used when cloning a particular PCR amplified gene segment of interest during the construction of phage display Ab libraries (Winter et al., 1994). Features shared among different types of phage vectors include an insertion site for the amplified Ab gene between the signal sequence and the gene encoding the coat protein, a Ff origin of replication,

and a selectable antibiotic resistance marker, (Smith and Petrenko, 1997) (Figure 3). Phage vectors result in the production of multivalent Ab phage protein progeny, meaning all copies of the phage coat protein display the Ab of interest (Russel et al., 2004; Smith and Petrenko, 1997). A number of different phage display vectors have been used over the years, with Fd and M13 being the most commonly used (Benhar, 2007). The Fd-tet phage vector is approximately 9.2 kilobases in size, and contains a coding segment for tetracycline (Tet) resistance inserted into the non-coding intergenic (IG) region of the fd genome (a 508 bp segment that does not code for protein and contains the origin of replication) (Zacher et al., 1980; Beck et al., 1978).

Figure 3: The use of phage vector systems in the construction of antibody phage display libraries. The Ab sequence to be displayed is placed in between the secretion signal sequence (Sig.) and the gene encoding a coat protein (in this case, pIII is shown). Phage vectors possess a filamentous phage origin of replication (Ff ori) for synthesis of ssDNA and antibiotic resistance markers allow for selective proliferation of phage vectors in *E. coli*. Phage are produced such that they are multivalent (i.e., all copies of the coat protein display the Ab of interest) (adapted from Russel et al., 2004).


Displayed
sequence
Sig.
Gene III
Phage
vector
Antibiotic
resistance
gene–optional
Ff ori
E. coli
Polyvalent
phage

c. pIII coat protein of filamentous phage

pIII, which is 406 residues in length, represents the most defined and widely utilized Ff coat protein largely due to its surface exposed N-terminal domain that contributes to retention of phage infectivity in spite of large protein insertions (Cwirla et al., 1990; Scott and Smith, 1990; Russel et al., 2004; Smith and Petrenko, 1997; Hoogenboom, 1997). pIII has two N-terminal domains (N1 and N2) and a C-terminal domain (CT) separated by two long, flexible glycine-rich linkers (Russel et al., 2004). These domains mediate phage infection and release of the phage particle following assembly. In addition to aiding in pIII incorporation into the phage particle, the final 132 residues of pIII's CT domain signal termination of assembly and release of phage from the cell (Rakonjac et al., 1999).

3. Development and types of antibody libraries

Library generation almost always involves PCR amplification of Ab variable domains from B cell genes followed by subcloning of the domains into a suitable vector for expression of the Ab fragment (Goldman et al., 2006). The first recombinant Ab library consisted of Ab V_H domains and was derived by PCR from the spleens of immunized mice (Ward et al., 1989). Ab libraries are generally divided into categories based on the source of DNA: naïve, immune, or synthetic.

a. Naïve libraries

Naïve libraries are constructed using the DNA of leukocytes from a host that has no exposure bias toward any particular Ag (Vaughan et al., 1996). Variable genes are typically harvested from the IgM mRNA of peripheral blood leukocytes (e.g., from non-immunized donors) and mRNA is subsequently reverse transcribed to produce cDNA; the variable genes are

then PCR amplified and cloned into a phage vector. Naïve libraries provide a readily available source of Ab fragments binding affinities against a wide variety of target Ags, whereas hyper-immunized repertoires are more likely to contain high affinity Abs specific to the Ag against which the individual was exposed (Marks et al., 1993; Hughes-Jones et al., 1994; Finnern et al., 1995; Watkins and Ouwehand, 2000). Naïve libraries can theoretically yield isolates against an unlimited range of Ags. An example of this library type is the naïve human Fab library constructed by de Haardt et al. (1999), yielding Abs with affinities in the nanomolar range.

b. Immune libraries

Immune Ab libraries are prepared in the same manner as naïve libraries with the only difference being that the host has been repetitively exposed to an immunogen, referred to as hyper-immunization (Burton et al., 1991; Clackson et al., 1991). While the choice of using an immune or naïve library depends upon the ultimate purpose (i.e., the library can either function as a "universal" source of binders, as is the case with naïve repertoires, or it can be designed to yield binders against one specific target Ag, as is the case with immune repertoires), each type has their own advantages and disadvantages (Watkins and Ouwehand, 2000). The immunized Ab library format is based on an Ag-specific response which ameliorates overall chances of obtaining target-specific Abs. However, the possibility of isolating Abs against self-Ags or toxic Ags represents a downside to the screening of immune repertoires. Another aspect to be considered in dealing with immunized libraries is the ethical justification behind the principles of using immunization to isolate immunogen-targeted Abs (Hoogenboom, 1997; Willats, 2002). Immune libraries normally consist of $\leq 10^8$ clones and take advantage of affinity maturation that occurs from immunization of the donor prior to B cell extraction for Ab gene isolation; naive libraries are usually larger since naïve Ab genes are not subjected to *in vivo* affinity maturation.

c. Synthetic libraries

Synthetic library construction incorporates the functional diversity advantages of naïve libraries with progresses continuously being made regarding the knowledge of Ab genes (i.e., the choice of variable gene segments as templates for library construction is guided by factors that will increase the overall performance and quality of the library). Unlike naïve or immune libraries, diversity of fully or semi-synthetic constructed Ab libraries is generated artificially, or *in vitro*. Synthetic oligonucleotides are annealed to a template Ab sequence for introduction of complete or partial sequence degeneracy into the CDR loops. The Ab template contains stable FRs and is selected from a list of human Ab germline consensus sequences based on the favourable biophysical characteristics it exhibits (e.g., high expression, reversible thermal unfolding, and monomericity). Synthetic repertoires often contain Ab genes not normally found *in vivo* (Winter and Milstein, 1991). Screening of synthetic repertoires allows for *in vitro* production of high-affinity Abs that have the potential to be directed toward a virtually limitless variety of target Ags (Winter and Milstein, 1991; Fellouse et al., 2007; Knappik et al., 2000). Synthetic libraries are also devoid of any factors that might limit expression (since the biophysical properties of the Ab template are maintained in the new Abs of the library) or provide a biased immunogenic response against a target Ag (since there was no host immunization). Synthetic repertoires offer the possibility of using human Ig and variable domain DNA sequences with *E. coli* host expression (Griffiths et al., 1994; Cook and Tomlinson, 1995; Andris et al., 1995; Winter, 1998; Knappik et al., 2000; Krebs et al., 2001). Synthetic libraries permit diversities of up to 10^{11} to be obtained, with subsequent isolation of Ab fragments demonstrating affinity values similar to those found in the human secondary immune response (Griffiths et al., 1994). Semi-synthetic libraries are derived from unrearranged V genes of pre-B

cells (i.e., germline cells) or from an Ab framework with genetically randomized CDR3, and are typically constructed *in vitro* by replacing amino acids in CDR1 and/or CDR2 with randomized sequences using PCR techniques (Pini et al., 1998). A further increase in diversity is achieved with fully synthetic libraries, in which a human framework is used in combination with randomized sequences in all three CDRs (Hayashi et al., 1994). Ab fragments selected from fully synthetic libraries have been shown to exhibit affinities in the very low nanomolar range (Knappik et al., 2000). Synthetic libraries are becoming increasingly popular in the isolation of Abs against any conceivable Ag.

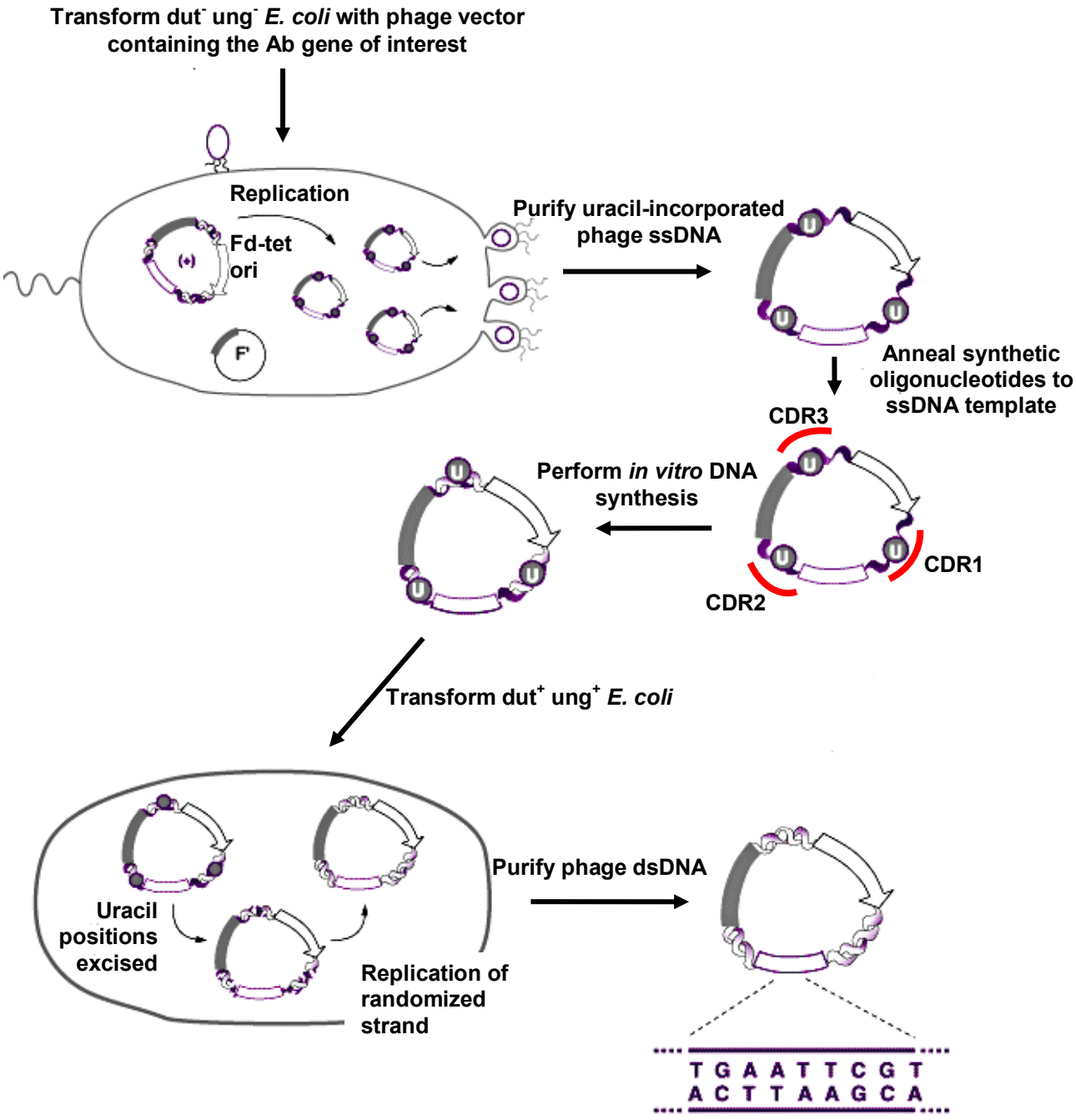
d. Library size

Library size is a crucial aspect in the successful isolation of target-specific Ab fragments: in order to increase the chance of isolating a high affinity Ab fragment from a synthetic library, the potential size of the library must be maximized. Theoretical studies have indicated, expectedly, that the larger the library, the greater the chance of finding clones that bind to any given antigenic epitope, with an increased affinity (Perelson and Oster, 1979). A library size of $\sim 10^{10}$ is likely to yield Abs with favourable affinities, somewhere in the low nanomolar range (Vaughan et al., 1996). Therefore, size and heterogeneity of repertoires play significant roles in the isolation of high affinity Abs. One limiting factor affecting large Ab libraries is the efficiency of plasmid or phage DNA introduction into bacterial hosts, which could theoretically result in a narrowing of the library's diversity down to 10^7-10^8 unique clones; the human natural Ab repertoire approaches the 10^{10}-10^{11} range (Winter et al., 1994).

e. Kunkel's method of mutagenesis

Kunkel's mutagenesis protocol is suitable for Ab libraries with randomized, adjacent positions; the overall scheme is shown in Figure 4. This method involves cloning of the sequence that is to be altered into a phage vector, and infection of a bacterial strain such as *E. coli* strain CJ236, that lacks dUTPase and uracil glycosylase (i.e., a *dut*$^{-}$ *ung*$^{-}$ strain) (Kunkel, 1985; Kunkel et al., 1987). The resultant ssDNA, consisting of uracil bases rather than thymine, can then be isolated and purified according to standard phage precipitation techniques. This makes it possible to anneal synthetic degenerate oligonucleotides *in vitro* to the phage ssDNA, creating a second DNA strand through enzymatic extension that lacks uridine. This dsDNA product can then be transformed into competent *E. coli* cells such as the TG1 strain, which inherently contain dUTPase and uracil glycosylase. These enzymes degrade the resulting uridine-containing strand, and the amplified phage particles each displaying unique Ab sequences can then be collected (Sidhu et al., 2000).

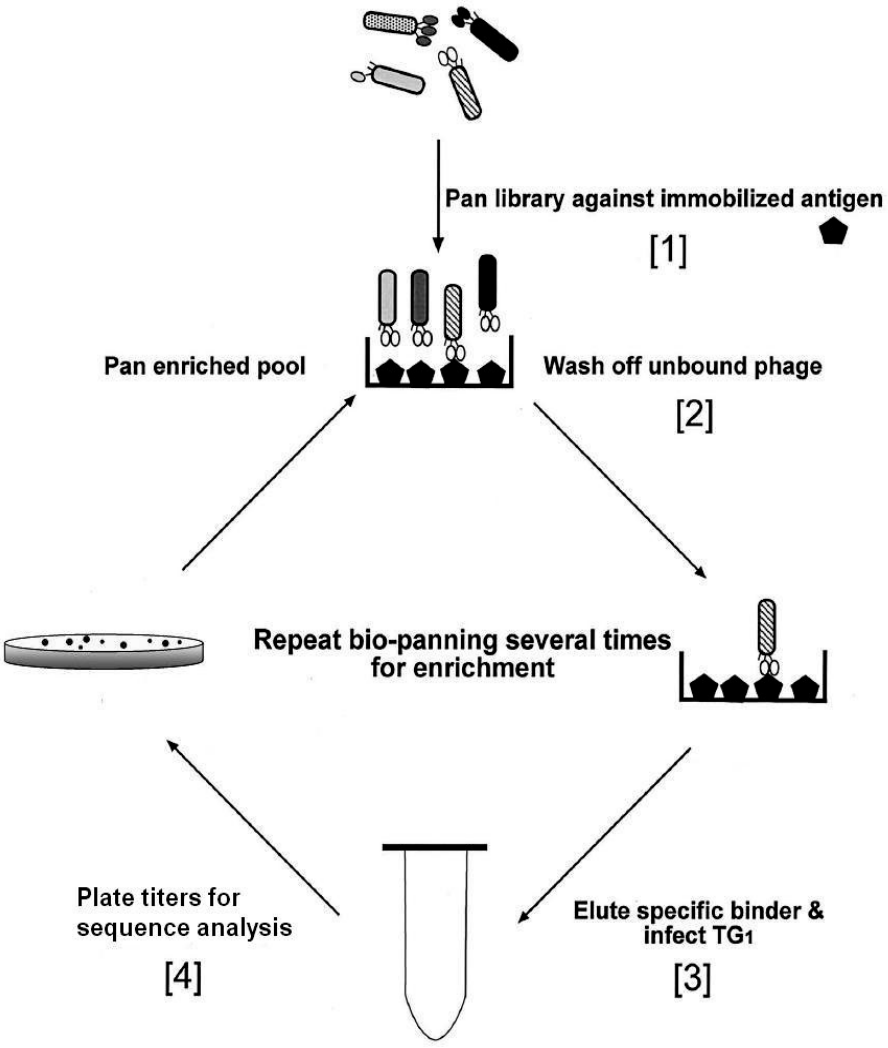
Figure 4: Synthetic randomization using Kunkel's method of mutagenesis. The DNA sequence to be mutated is first cloned into a Ff vector (M13 or Fd-tet). In producing phage ssDNA, a *dut*$^{-}$ *ung*$^{-}$ bacterial host strain (e.g., *E. coli* CJ236) is used such that its lack of dUTPase leads to elevated dUTP levels and the misincorporation of deoxyuridine in place of thymidine. The absence of uracil-N glycosylase also makes the removal of uracil residues impossible. Synthetic degenerate oligonucleotides are then annealed to the uridine-containing ssDNA, with full complementary DNA synthesis mediated by polymerase and ligase. *E. coli* strain TG1 (*dut*$^{+}$ *ung*$^{+}$ contains the enzymes needed to successfully remove uridines from DNA, and is transformed with the phage dsDNA, resulting in degradation of the parental uridine-containing strand. This newly synthesized strand, now containing randomized regions, survives and is replicated to produce intact randomized plasmid (adapted from <http://www.biochem.arizona.edu/classes/bioc471/pages/Lecture6/Lecture6.html>).


Transform dut- ung- E. coli with phage vector containing the Ab gene of interest
Replication
Fd-tet ori
F'
Purify uracil-incorporated phage ssDNA
Anneal synthetic oligonucleotides to ssDNA template
CDR3
CDR1
CDR2
Perform in vitro DNA synthesis
Transform dut+ ung+ E. coli
Uracil positions excised
Replication of randomized strand
Purify phage dsDNA
T G A A T T C G T
A C T T A A G C A

4. Selection of binders from antibody libraries: "panning"

The screening of Abs from a phage display library is carried out *in vitro* and is referred to as "bio-panning" or "panning". The procedure of panning involves the isolation and amplification of phage that display Ag-binding Ab fragments through a series of specific steps (Maynard and Georgiou, 2000; Paschke, 2006). Phage display Ab libraries are generally exposed to the target Ag, which can be immobilized onto plastic surfaces (e.g., ELISA plates) or present in solution, followed by washing to remove those phage carrying non-specific Abs (Kang et al., 1991; Marks et al., 1991; Malmborg et al., 1996). Abs specific to the target Ag are then eluted, or dissociated (e.g., through acidic/basic conditions, or by competition through a finite amount of Ag), before being amplified in host *E. coli* cells (Clackson et al., 1991; Paschke, 2006). Competitive elution improves the specificity of the selected Abs when panning for a blocking Ab or one that competes with a known ligand of a receptor Ag. Clones that prevail are enriched following each panning round and go on to be used in further rounds of selection, effectively narrowing the library to those Abs that bind the target (Yang and Nolan, 2007). High affinity Abs are normally selected within three to four rounds of panning, however more rounds may sometimes be needed to observe successful enrichment (de Bruin et al., 1999). One round of the panning cycle is summarized in Figure 5.

Figure 5: Screening of antibody phage display libraries by panning. Selection of Ag-specific library phage clones is carried out in four basic steps, comprising one round of the panning cycle. Panning begins with [1] exposure and incubation of phage clones with the target Ag (this can be done on a plastic surface, in solution, on cells or *in vivo*) prior to blocking with BSA or milk / casein-PBS, [2] washing steps to exclude phage that bind non-specifically, [3] elution for recovery of target-specific phage as well as amplification of selected phage in *E. coli* (e.g., TG1 strain), and [4] plating of titers for subsequent sequence and binding analysis of individual phage clones. This cycle is repeated 3-5 times depending on the observed enrichment of phage clones (adapted from Azzazy and Highsmith, 2002).


Pan library against immobilized antigen
[1]
Pan enriched pool
Wash off unbound phage
[2]
Repeat bio-panning several times
for enrichment
Plate titers for
sequence analysis
[4]
Elute specific binder &
infect TG1
[3]

Sequencing of titer clones is typically performed following each round of panning for the identification of one or more repeating consensus sequences, which is considered a measure of a successful selection. This way, clones are identified and later characterized for their binding properties. Sequencing can sometimes, however, lead to the identification of irrelevant or non-binding clones are obtained by panning, despite the panning procedure having been correctly carried out using the target Ag of interest. This is because panning not only involves binding to the target Ag, but also to a solid surface (e.g., microwell or magnetic bead) and sometimes an affinity partner (ie: biotin or streptavidin); as result, the isolated clones may instead bind plastic, PBS or streptavidin. There also exists the possibility of a bias towards clones with growth and expression advantages in *E. coli* hosts during amplification steps (Adey et al., 1995). In this case, those clones with high affinity properties but mediocre growth properties would not predominate in the third or fourth panning rounds. However, due to the time investment and laborious steps involved in functional characterization methods for binding, sequencing as a primary tool is used to identify the emergence of a consensus sequence before any functional screening is performed (Yang and Nolan, 2007).

The enrichment factor, another indication of successful selection, is based on the calculated phage yield for each round of panning. Input and output phage numbers are counted as colony-forming units (CFU) on titer plates and phage yield is calculated by dividing the number of output phage (the number of phage following elution and before amplification in *E. coli*) by the number of input phage (the number of phage introduced at the beginning of each panning round; this number is the same for each round). Phage yield should ideally show a considerable increase in later rounds compared with earlier ones. The enrichment factor, which is obtained by dividing the phage yield obtained from the final round of panning by that of the

first round, differs with each panning protocol; however typical values between 10^3- and 10^6-fold are often observed after four rounds of panning (Arbabi-Ghahroudi et al., 2009). Panning progress can also be followed via colony PCR, in which there is generally an increase in the number of clones with Ab insert compared to those with no insert after each successive round of panning (Harrison et al., 1996). Binding properties of the isolated phage clones can be characterized through secondary screening methods, including the enzyme-linked immunosorbent assay (ELISA). However, as previously mentioned, panning may not result in the isolation of high affinity binders.

Screening for Abs in solution eliminates problems associated with conformational changes that are normally found with immobilization of Ag on solid surfaces. This technique also allows for any alterations in Ag concentration to be made, such as decreasing the amount of Ag used for the purpose of isolating higher affinity Abs (Hawkins et al., 1992). Selection of phage-displayed Abs against cell surface-displayed epitopes can be performed with cells grown in suspension, or adherent cultures (Azzazy and Highsmith, 2002). Isolation of specific binders can be enhanced with simultaneous positive and negative cell selection (de Kruif et al., 1995). In this approach, a competitive environment is designed between an Ag-positive target cell line and a cell line not expressing the target of interest, with this negative cell line serving as a sink for the non-specific adherence of irrelevant binders.

5. Protein expression in *E. coli*

Among the many existing hosts for recombinant protein production, the Gram-negative enterobacterium *E. coli* remains one of the most widely used due to its simplicity, safety, high growth rate, ability to reach high densities, well-characterized genome, and the availability of a

large variety of cloning vectors (Sørensen and Mortensen, 2005). Some of the other advantages associated with expression of Ab fragments in *E. coli* include relatively high transformation efficiencies with foreign DNA, well-established genetic manipulation methods, the ability to produce proteins rapidly and in large quantities, ease of genetic manipulation, no lengthy delays between transformation and protein purification, and inexpensive fermentation procedures. However, obtaining abundant yields of correctly folded proteins can often be problematic (Arbabi-Ghahroudi et al., 2005).

The main purpose of Ab protein expression is to obtain an abundant accumulation of soluble product in the bacterial cell. However, there are instances where the metabolic system of the host will not cooperate, leading to cellular stress responses being encountered. Another possibility encountered is the production of misfolded and biologically inactive, insoluble, aggregated proteins, in the form of inclusion bodies; laborious and expensive denaturation-refolding processes are required to recover the active proteins (Villaverde and Carrio, 2003).

a. Protein expression in the periplasm of *E. coli*

A favourable approach to expressing functional Ab fragments occurs within the periplasmic space of *E. coli*, which is considered a highly oxidizing environment. In this method, proteins carrying specific signal sequences (e.g., pelB, phoA and ompA) are directed to the periplasmic region (Pugsley, 1993). Protein expression within the periplasm almost always guarantees proper folding of the protein and is considered analogous to eukaryotic protein synthesis. Unlike the cytoplasmic route, periplasmic expression enables complete *in vivo* assembly of fully functional proteins due to a scarce presence of protease enzymes (Bothmann and Plückthun, 1998; Bothmann and Plückthun, 2000). Ab fragments expressed in the periplasm have been shown to be correctly folded with yields of 0.1-100 mg/L in baffled flasks and 1-2 g/L

in fermentors (Carter et al., 1992). Furthermore, protein purification of Abs from periplasmic extracts is far simpler and less problematic compared to protein purification from cell lysates due to a lower incidence of contamination from bacterial proteins (Arbabi-Ghahroudi et al., 2005).

b. Factors influencing protein expression

Protein production is the cumulative result of complex, sequential steps, including transcription, mRNA processing, translation, post-translational modification, protein folding and assembly, as well as protein export. A reduction in protein yield occurs when any of the aforementioned steps are compromised in any way (Arbabi-Ghahroudi et al., 2005). Among the factors influencing expression are Ab gene structure, stability of mRNA, intrinsic properties of the expression vector, and the presence of inducible promoters (Plückthun, 1994; Skerra, 1994; Daugherty et al., 1999; Arbabi-Ghahroudi et al., 2005). Expression of functional protein in *E. coli* also depends highly upon amino acid sequence. Even though one Ab clone may be favourably expressed in its biologically active and functional form, another clone may not reflect the same properties (Duenas et al., 1995; Nieba et al., 1997; Woo et al., 2002).

c. Protein purification by Immobilized Metal Affinity Chromatography (IMAC)

Metal chelate or immobilized metal affinity chromatography (IMAC) is a powerful technique for isolating Ab fragment proteins under gentle conditions (Lindner et al., 1992; Porath, 1992). In this method, a metal chelating group is first adsorbed onto a chromatographic medium, and a multivalent metal ion (e.g., Cu^{2+}, Ni^{2+}, Zn^{2+}, Co^{2+}) is bound such that a number of coordination sites are left vacant for protein interaction. Typically, a tag of 5-6 histidine residues is added to the C- or N-terminus of the protein to be purified using recombinant techniques. The tag is designed to specifically interact with the chelated metal ions, causing them to be retained

in the chromatographic medium while non-protein components or untagged proteins bind either weakly or not at all. Elution of the protein normally involves an increase in the concentration of a competitive eluting agent, such as imidazole, or a reduction in pH (Schmidbauer and Strobel, 1997).

6. Semaphorins as promoters of neurodegeneration

Neurodegeneration involves the progressive loss of structure or function of neurons, often resulting in the death of neurons. Diseases such as Parkinson's, Alzheimer's, and Huntington's, are a direct result of neurodegenerative processes (Bredesen et al., 2006). In a normal brain, the neuron's dendrites receive signals or impulses which are transmitted through the axon to its target tissue; the axonal growth cone picks up directional cues from the surrounding environment to form synapses with its target tissue. Repellant or "chemorepulsive" agents play a significant role in determining the outcome of stroke or cortical injury; chemorepulsion is defined as the directional movement of a cell away from a stimulus (Vianello et al, 2010). Proteins called Semaphorins function as axonal guidance cues and serve to inhibit axonal outgrowth, effectively preventing new neurons from regenerating and entering the ischemic (i.e., damaged) area (Jiang et al., 2010). Growth cone collapse is described as a rearrangement of the cytoskeleton and detachment of adhesion sites from the extracellular matrix (Mikule et al., 2002).

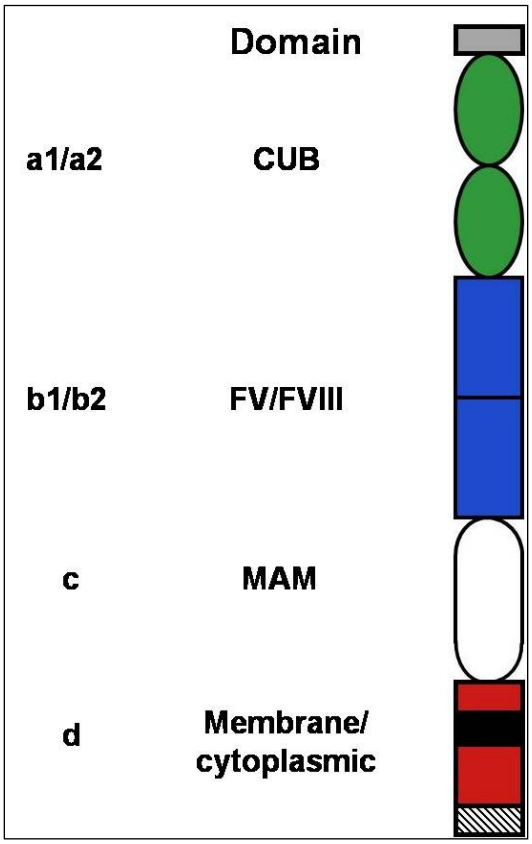
Semaphorins are a conserved family of secreted transmembrane proteins that are involved in axonal repulsion and growth cone collapse, effectively promoting neurodegenerative diseases (Semaphorin Nomenclature Committee, 1999; Pasterkamp and Kolodkin, 2003). Typically 750 residues in length, all members of the Semaphorin family are characterized by a "Sema" domain of approximately 550 amino acids at their amino terminus, and 14-16 conserved

cysteine residues. The plexin and neuropilin (NRP) families are currently the only known receptors that bind the Sema family of proteins (Nakamura et al., 2000).

a. Neuropilins: structure and function

NRPs comprise a family of non-tyrosine kinase receptors for both the secreted class III Semaphorins (i.e., Sema3A) and vascular endothelial growth factor (VEGF) (Soker et al., 1996; Soker et al., 1998; Nakamura and Goshima, 2002; Scarlato et al., 2003). Two homologues with molecular masses of 120-130 kDa, NRP1 and NRP2, have been identified in vertebrates. The conserved primary structure of NRPs consists of three extracellular domains, a single transmembrane domain, and a short cytoplasmic domain (Fujisawa et al., 1997). The first extracellular domain comprises amino-terminal CUB (also referred to as "a1" and "a2") domains named for their homology to complement proteins C1r and C1s, which are responsible for initiating activation of the complement cascade (Ellis, 2006; Gál et al., 2009). The coagulation factor V/VIII (FV/VIII, also referred to as "b1" and "b2") and C-terminal MAM (Meprin, A5, μ-phosphatase, also referred to as "c") domains make up the second and third extracellular domains (Ellis, 2006). A membrane-spanning and short (40-43 residues) cytoplasmic tail ("d") complete the basic structure of NRPs (Nakamura and Goshima, 2002) (Figure 6).

Figure 6: Neuropilin structure. The structure of NRP consists of two adjacent, extracellular CUB (also referred to as "a1/a2") domains, two extracellular FV/FVIII coagulation factor (also referred to as "b1/b2") domains, an extracellular MAM (Meprin, A5, μ-phosphatase, also referred to as "c")) domain, and a transmembrane/cytoplasmic portion (also referred to as "d") (adapted from Nakamura and Goshima, 2002).


Domain
a1/a2
CUB
b1/b2
FV/FVIII
c
MAM
d
Membrane/
cytoplasmic

b. The role of NRP1/Sema3A interaction in neurodegenerative diseases

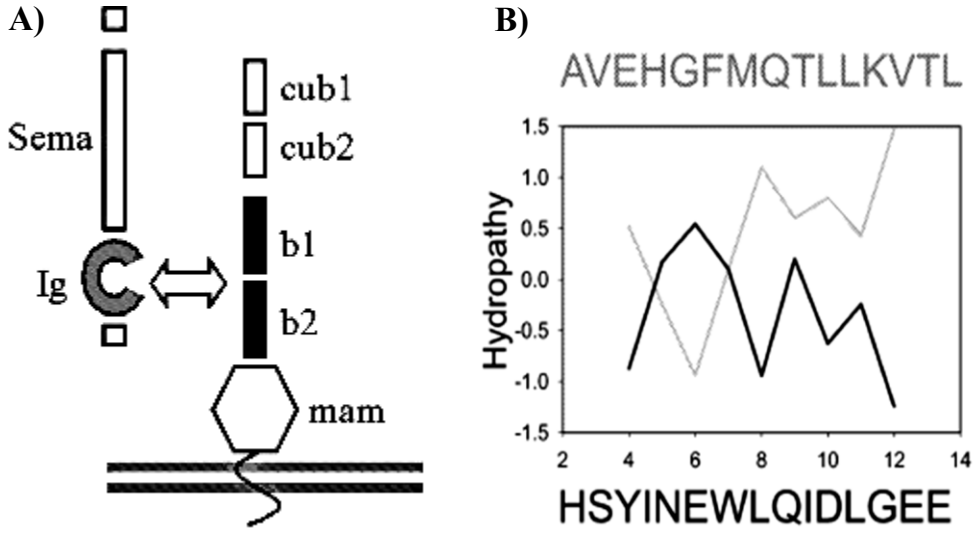
Growth cone collapse, axonal repulsion, and neural cell death have all been shown to be mediated by the interaction of Sema3A and NRP1 in adult ischemic brains (Hou et al., 2008; Jiang et al., 2010). Experiments have shown that NRPs require adapter proteins in order to mediate downstream signaling events (Nakamura et al., 1998). Large transmembrane proteins called plexins, enable Sema3A signaling through the formation of co-receptor complexes with NRP1 (Tamagnone et al., 1999). Via their "Sema" domain, plexins form a ternary complex with NRP1 and Semaphorins, thereby acting as a bridge between the proteins (Renzi et al., 1999; Tamagnone et al., 1999). Class 3 Semaphorins bind NRP1 through both their "Ig" and "Sema" domains (Renzi et al., 1999). The "Ig" domain binds to the "b1/b2" domain of NRP1, and the "Sema" domain binds to the "a1/a2" domain of NRP1 (Giger et al., 1998; Williams et al., 2005). However, the exact binding specificities within the plexin, NRP1, and Semaphorin complex are not known. The role of increased NRP1 and Sema3A expression in the ischemic brain has been established in several neurodegenerative diseases (Shirvan et al., 1999; Gagliardini and Fankhauser, 1999; De Winter et al., 2002; Campbell and Holt, 2003; Ben-Zvi et al., 2008; Hou et al., 2008; Pasterkamp and Giger, 2009). For example, one group observed significant improvement in nerve regeneration of injured spinal cord neurons in adult mice using an agent selectively inhibiting Sema3A (Kaneko et al., 2004). Collectively, data indicates that the Sema3A/NRP1 pathway contributes to the acute cascade of deleterious biochemical events leading to neuronal death.

c. Synthetic Sema3A "Ig" inhibitory peptide

A synthetic peptide (N-Ac-HAVEHGFMQTLLKVTLE) containing the Semaphorin "Ig" domain of Sema3A and specifically targeting NRP1, has been tested both *in vitro* and *in vivo* for

its ability to inhibit Sema3A-induced growth cone collapse (Williams et al., 2005; Jiang et al., 2010). One *in vitro* growth cone collapse assay was performed, whereby injection of the "Ig"-containing synthetic peptide into Sema3A-expressed cultured mouse cortical neurons abolished Sema3A's inhibitory effects on neurite outgrowth and growth cone collapse, with full response inhibition at < 50 μM (Williams et al., 2005; Jiang et al., 2010). In another experiment, the synthetic peptide was administered, along with a control peptide, into the cerebral ventricles of stroke-induced mice. It was determined that the peptide provided significant neuroprotection compared to the brain injected with the control peptide (Jiang et al., 2010). However, studies have thus far only examined the effects of Sema3A/NRP1 following 24 h reperfusions; the long term effects of blocking Sema3A/NRP1 interaction in stroke-induced adult brains remains to be investigated. The AVEHGFMQTLLKVTL motif within the "Ig" domain of Sema3A, which corresponds to the sequence of the synthetic peptide, has been suggested to exhibit its inhibitory effects by interacting with the "b1/b2" domain of NRP1 (Figure 7A). Homology modelling indicates this sequence to be a major exposed loop on the "Ig" domain, and therefore readily available for NRP1 binding (Figure 7B) (Williams et al., 2005).

Figure 7: Interacting domains and hydropathy profile for Sema3A and NRP1. A) Schematic highlighting the interaction that occurs between the Sema "Ig" domain of Sema3A and the "b1/b2" (FV/FVIII coagulation factor) domain of NRP1. B) The computed complementary hydropathy profile between Sema3A "Ig" domain (grey) and the "b1/b2" domain of NRP1 (black) is shown. Sequences from Sema3A (in grey) and NRP1 (in black) are also provided. An anti-parallel interaction is observed (Williams et al., 2005).


A)
Sema
cub1
cub2
b1
b2
Ig
mam
B)
AVEHGFMQTLLKVTL
Hydropathy
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
2
4
6
8
10
12
14
HSYINEWLQIDLGEE

7. Objectives

The objectives of this thesis included constructing a large human V_L phage display library, analyzing it via sequence analysis, subtractively panning the library for isolation of NRP1-specific V_L sdAbs, and characterizing the V_Ls by SEC analysis and flow cytometry.

- Construction of the synthetic phage Ab library utilized a single V_L framework with selected positions in all 3 complementarity-determining regions (CDRs) targeted for *in vitro* randomization using synthetic oligonucleotides

- A synthetic repertoire based on these V_Ls therefore allows for the isolation of novel Abs with high affinity, solubility and thermostability, and provides a promising source of therapeutic proteins

- Specific V_Ls were selected from the repertoire through subtractive panning against a cell line endogenously expressing NRP1 (PC12) as well as a negative cell line that does not express NRP1 (HEK293) with competitive elution carried out using a synthetic peptide derived from the Sema Ig motif of Sema3A

- The synthetic peptide competitively rescued V_Ls that specifically bound the NRP1 epitope displayed on the surface of PC12 cells, through interaction with the b1/b2 domains of NRP1

- Isolated V_Ls were purified by IMAC and analyzed by SEC for monomericity followed by flow cytometry for cell binding

VI. MATERIALS AND METHODS

1. Reagents and solutions

BamHI restriction enzyme and 10x *BamHI* buffer, as well as agarose gel electrophoresis DNA ladders (100 bp and 1 kbp), were purchased from New England BioLabs (Pickering, ON, Canada). *BpiI* (*BbsI*) restriction enzyme and its corresponding buffer (10x Buffer *G*) were obtained from Fermentas (Burlington, ON, Canada). Ampicillin, chloramphenicol, tetracyclin, isopropyl β-D-1-thiogalactopyranoside (IPTG), phenylmethylsulphonyl fluoride (PMSF), isopropanol, bovine serum albumin (BSA), N-(2-hydroxyethyl)-piperazine-N-2-ethanesulfonic acid (HEPES), Tris-HCl, sodium-dodecyl sulfate (SDS), triethylamine (TEA), polyethylene glycol (PEG), and polyoxyethylene sorbitan mono-laurate (Tween-20) were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO, USA). Bacto-tryptone, Bacto-yeast extract, and bacto-agar were purchased from BD Bioscences (Mississauga, ON, Canada). Synthetic oligonucleotides ("*CDR1"*, "*CDR2"*, "*CDR3"*, "*CDR3a"*, and *"CDR3b"*) and PCR primers ("FdtetGIIID" and "Fd-96GIIID") were synthesized by Integrated DNA Technologies (Coralville, IA, USA). PVDF membrane was purchased from Roche Diagnostics (Laval, QC, Canada), mouse anti-His6 IgG-alkaline phosphatase (AP) conjugate was obtained from Abcam (Cambridge, MA, USA) and alkaline phospharase (AP)-goat anti-mouse IgG conjugate was obtained from Cedarlane Laboratories (Burlington, ON, Canada). Protein Plus Standard Ladder as well as AP development substrates were purchased from Bio-Rad Laboratories (Hercules, CA, USA). Anti-6His PE for flow cytometry analysis was purchased from Miltenyi Biotec (Auburn, CA, USA). Taq DNA polymerase, 10x Taq DNA polymerase buffer, and synthetic peptide that

was used for competitive elution of subtractive panning was obtained from GenScript (Piscataway, NJ, USA).

2. Media, bacterial strains, and mammalian cell lines

All media used for this project was sterilized by autoclaving at 15 p.s.i. (pounds per square inch) and 121ºC for 20 mins, unless otherwise noted. 2xYT broth contained 1.6% (w/v) Bacto-tryptone, 1% (w/v) Bacto-yeast extract and 0.5% (w/v) NaCl. Solid media for plates was prepared as for the broth media as described, except that 15 g/L of Bacto-agar was added before autoclaving. Tetracycline (Tet) was prepared as a stock solution at a concentration of 12.5 mg/mL in 50% ethanol, filter sterilized and stored at 4°C in a darkened container; Ampicillin (Amp) and chloramphenicol (Chlor) stock solutions were prepared at 100 mg/mL and 10 mg/mL, respectively, filter sterilized, and stored at -20°C until use. When necessary, Tet, Amp, and Chlor, were individually added to 2xYT medium at final concentrations of 12.5 µg/mL, 100 µg/mL, and 10 µg/mL, respectively; solid media plates were also prepared with these respective concentrations except for Tet, which was at a final concentration of 5 µg/mL. B2xYT media prepared for small-scale (50 mL cultures) protein expression contained 1.6% (w/v) Bacto-tryptone, 1% (w/v) Bacto-yeast extract, 0.5% (w/v) NaCl and 0.4% (v/v) glycerol. Autoclaved separately and later added to the B2xYT media were 1.2% (w/v) K_2HPO_4 and 0.2% (w/v) KH_2PO_4. M9S medium was prepared for large-scale (1 L cultures) protein expression by dissolving 1x M9 salts (0.6% Na_2HPO_4, 0.3% KH_2PO_4, 0.05% NaCl, 0.1% NH_4Cl, w/v) in 1 L ddH_2O and autoclaved. For use in protein expression, the following supplements were added to M9S medium: 0.4% (w/v) filter-sterilized glucose, 5 µg/mL filter-sterilized thiamine-HCl, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$ and 0.4% (w/v) filter-sterilized casamino acids (added to improve the

mutator cell growth rates). SOC medium (0.5% Bacto-yeast extract, 2% Bacto-tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM MgSO4, 20 mM glucose) was prepared and stored at -20ºC until required for bacterial transformations. HEPES (*N*-(2-hydroxyethyl)-piperazine-*N*-2-ethanesulfonic acid) solution (pH 7.4) was prepared at a concentration of 1 M for later use as a major component in the protein dialysis buffer ("Buffer A"). 10x Induction Medium to be added during protein expression consisted of 24 grams of Bacto-tryptone, 48 grams of Bacto-yeast extract, and 12 mL of glycerol per 100 mL of ddH_2O.

E. coli strain TG1 (genotype: *supE hsd 5thi (lac-proAB) F ' [traD36 proAB + lacI9 lacZ M15*) was used for phage Fd-tet vector production, as well as protein expression. *E. coli* strain CJ236 (genotype: F∆*(HindIII)::cat* (Tra^+ Pil^+ Cam^R)/ *ung-1 relA1 dut-1 thi-1 spoT1 mcrA)* was used for phage ssDNA production and isolation for the purpose of synthesizing heteroduplex DNA with three randomized CDR regions. *E. coli* strain CJ236 was grown in 2xYT medium with 10 µg/mL chloramphenicol (Chlor). Pheochromocytoma cells of the rat adrenal medulla (PC12) were cultured in D-MEM (Dulbecco's Modified Eagle Medium) high-glucose (Sigma-Aldrich Chemical Co.), supplemented with 10% horse serum (Invitrogen, Carlsbad, CA, USA), 5% FBS (Fetal bovine serum) (Invitrogen, Carlsbad, CA, USA), and Antibiotic/Antimycotic 100x Solution (Caisson Labs, North Logan, UT, USA). Human embryonic kidney cells (HEK293) were cultured in D-MEM (Dulbecco's Modified Eagle Medium) high-glucose solution, supplemented with 10% FBS and 2 mM L glutamine.

3. Preparation of *E. coli* TG1 electrocompetent cells

Fresh electrocompetent *E. coli* TG1 cells were prepared by picking a single bacterial colony from a freshly streaked M9S medium plate and growing it in 50 mL of 2xYT medium

overnight at 37ºC and 230 rpm. The following day, 3x 10 mL of the overnight culture was transferred into 3x 4 L baffled flasks each containing 2 L 2xYT, and incubated at 37ºC (shaking) until $OD_{600} \approx 0.5$. The flasks were then chilled on ice for 1 hour, with all subsequent steps performed on ice. Cells were harvested by transferring the cultures to 6x 400 mL bottles for centrifugation (1,600 g, 4ºC, 20 min). Cell pellets were washed three times by filling each of the 3 bottles and resuspending the cell pellets with ~330 mL 1 mM ice cold HEPES (pH 7.4), followed by centrifugation (1,600 g, 4 ºC, 20 min). After the third wash, each of the 3 pellets was resuspended in 150 mL 10% (v/v) glycerol (autoclaved), followed by another centrifugation (4,400 g, 4ºC, 15 min). Pellets were then serially resuspended with 10% glycerol, such that 3 mL 10% glycerol was added to the first tube's pellet for resuspension; this resuspended pellet was then transferred to the second tube for resuspension, and so on until all pellets were resuspended and combined into one tube. The final volume of *E. coli* TG1 electrocompetent cells was divided into aliquots for library construction purposes. Aliquots were stored at -80ºC until needed. Prepared electrocompetent cells were tested for their transformation efficiency using commercially available pUC18 vector (~2.7 kbp; GenScript) prior to constructing the synthetic V_L library. Transformation was carried out as follows. First, 0.1 ng (~1 µL) of pUC18 vector was mixed into 50 µL *E. coli* TG1 electrocompetent cells (kept on ice) and transferred to an ice-cold 0.1 cm gap electroporation cuvette (Bio-Rad). Gentle tapping of the cuvette allowed for the cell suspension to be collected toward the bottom, and the cuvette was then wiped down to remove any residual water from the outer surface. The cuvette was then inserted into the slide chamber of a MicroPulser electroporator (Bio-Rad), with the slide pushed into the chamber until the cuvette made firm contact with the chamber electrodes. The capacitor was then charged and a pulse delivered using the "EC1" setting of the MicroPulser, after which the cuvette was removed

for addition of 1 mL SOC with gentle up-and-down pipetting to mix the contents. The ~1 mL transformation material was incubated at 37ºC for 45 min (230 rpm). Dilutions were then made (10^{-2},…,10^{-6}) in 2xYT, with 100 µL of each dilution spread onto agar plates containing Amp (100 µg/mL) for antibiotic selection overnight at 37°C.

4. Construction of the CDR3/3a/3b-randomized synthetic V_L phage display library

a. ssDNA production using *E. coli* CJ236

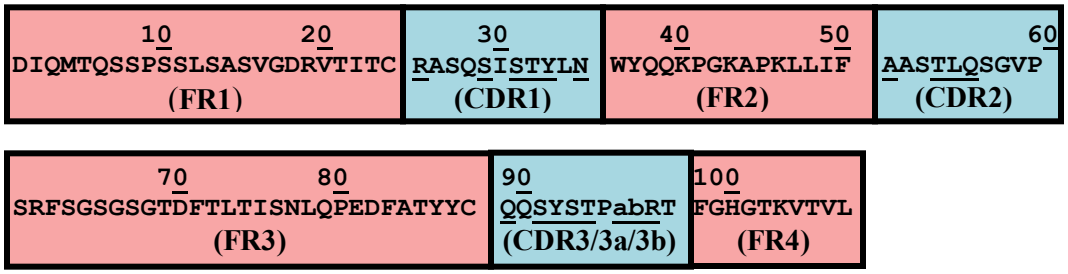
Human V_L-24 gene (which served as a template) was sub-cloned into the Fd-tet phage vector by a senior member of the lab (Fd-tet vector containing V_L-24 insert was named "Fd-tet24" in subsequent sections). First, *E. coli* strain CJ236 (contains gene for chloramphenicol resistance) was grown to log phase ($OD_{600} \approx 0.8$), and streaked on a 2xYT/Chlor (10 µg/mL) agar plate for overnight growth at 37ºC. The next day, one colony was picked and grown in 2xYT/Chlor (10 µg/mL) at 37°C (230 rpm) until log phase. Serial dilutions (10^{-3},…,10^{-12}) of the Fd-tet24 phage were made in sterile 1x PBS and 10 µL from each phage dilution was taken for infection of 200 µL *E. coli* CJ236 log phase cells. Infection titers were incubated at 37ºC for 45 min, then 100 µL from each dilution was plated on 2xYT/Chlor (10 µg/mL)/Tet (5 µg/mL) agar for overnight antiobiotic selection at 37ºC. Four single colonies from the 10^{-10} 2xYT/Chlor (10 µg/mL)/Tet (5 µg/mL) agar plate were then individually grown in 4 x 1 mL cultures containing 2xYT/Chlor (10 µg/mL)/Tet (12.5 µg/mL) with shaking (230 rpm) at 37ºC for 6 hours, after which the 4 x 1 mL cultures were added to 4 x 100 mL baffled flasks each containing 30 mL of 2xYT/Tet (12.5 µg/mL)/uridine (2.5 µg/mL) for overnight growth at 37ºC (230 rpm).

Phage precipitation was carried out with 1/5 volume of 20% PEG 6000/2.5 M NaCl on the overnight cultures to obtain phage with DNA incorporation of uridine rather than thymidine. Briefly, phage precipitation involved centrifugation (6,400 g, 15 min, 4ºC), transfer of supernatant to a fresh tube, addition of 1/5 volume of 20% PEG 6000/2.5 M NaCl to the supernatant, incubation on ice (1 h), a second centrifugation (16,600 g, 30 min, 4ºC), and resuspension of the phage pellet in 0.5 mL 1x PBS. To verify that the majority of phage had incorporated uridine in favour of thymidine, both *E. coli* strains TG1 and CJ236 were grown until log phase, then individually infected with Fd-tet24 phage that was serially diluted (10^{-2},…,10^{-10}) in 1x PBS for phage titers, performed as previously described. 10 µL of each phage dilution was added to 200 µL of either bacterial strain, and each reaction mixture was incubated at 37ºC for 45 min. 100 µL of infected CJ236 dilutions were plated on 2xYT/Chlor (10 µg/mL)/Tet (5 µg/mL) agar, while 100 µL of infected *E. coli* TG1 dilutions were plated on 2xYT/Tet (5 µg/mL) agar. Plates were incubated overnight at 37ºC for antibiotic selection. Phage ssDNA (containing uridine instead of thymidine) was subsequently isolated and purified from *E. coli* CJ236 using the QIAGEN M13 Purification™ kit as per the manufacturer's instructions (QIAGEN, Mississauga, Ontario).

b. *In vitro* mutagenesis (CDR3/3a/3b randomization)

In vitro primer extension and ligation to obtain CDR3/3a/3b-randomized dsDNA was first performed using 3 different synthetic mutagenic oligonucleotides. The human V_L template protein sequence is illustrated in Figure 8, along with positions throughout the CDRs that were selected for randomization.

Figure 8: Human V_L-24 template amino acid sequence with positions randomized during *in vitro* mutagenesis. The human V_L template amino acid sequence with a length of 107 residues is depicted, and divided into framework regions (1-4) and CDRs (1-3), as indicated. Randomized positions are underlined, and residues are numbered according to the nomenclature described by Kabat et al. (1991). Synthetic oligonucleotides were first annealed to the DNA region corresponding to CDR3 during *in vitro* mutagenesis for randomization of selected positions to generate three individual lengths of CDR3 DNA (321, 324, and 327 nucleotides, respectively). Following creation of the CDR3/3a/3b V_L library, a second *in vitro* mutagenesis step resulted in the randomization of selected positions within CDR1 and CDR2, for creation of the synthetic V_L phage display library.


10 20
DIQMTQSSPSSLSASVGDRVTITC
(FR1)
30
RASQSISTYLN
(CDR1)
40 50
WYQQKPGKAPKLLIF
(FR2)
60
AASTLQSGVP
(CDR2)
70 80
SRFSGSGSGTDFTLTISNLQPEDFATYYC
(FR3)
90
QQSYSTPabRT
(CDR3/3a/3b)
100
FGHGTKVTVL
(FR4)

The DNA sequences for the synthetic oligonucleotides used in randomization of the 3 CDRs are shown in Table 1, whereby "N" represents a 25% mix each of adenine, thymine, guanine, and cytosine nucleotides; "K" represents a 50% mix each of thymine and guanine nucleotides, and "R" represents a 50% mix each of adenine and guanine nucleotides.

Table 1. Synthetic oligonucleotide DNA sequences for *in vitro* mutagenesis annealing.

"oligo1"	5' AGA GTC ACC ATC ACT TGC NNK GCA AGT CAG RGC ATT NNK NNK NNK TTA NNK TGG TAT CAG CAG AAA CCA 3'
"oligo2"	5' CCT AAA CTC CTG ATC TTT NNK GCA TCC NNK CKN NNK AGT GGG GTC CCA TCA AGG 3'
"oligo3"	5' TTT GCA ACT TAC TAC TGT NNK CAG NNK NNK NNK NNK CCT NNK ACG TTC GGC CAC GGG ACC 3'
"oligo3a"	5' TTT GCA ACT TAC TAC TGT NNK CAG NNK NNK NNK NNK CCT NNK NNK ACG TTC GGC CAC GGG ACC 3'
"oligo3b"	5' TTT GCA ACT TAC TAC TGT NNK CAG NNK NNK NNK NNK CCT NNK NNK NNK ACG TTC GGC CAC GGG ACC 3'

DNA Phosphorylation

The DNA phosphorylation reactions (one for each of the five oligonucleotides) were set up in total volumes of 20 µL with the following components: 1 µg (~0.6 µL) "oligo3", "oligo3a", or "oligo3b", 2 µL of 10 mM ATP, 20 units of T4 polynucleotide kinase (New England BioLabs), and 2 µL of 10x polynucleotide kinase buffer (New England BioLabs). The reaction mixtures were incubated at 37ºC for 1 h. Phosphorylated "oligo1" and "oligo2" were stored at -20°C until used for annealing.

DNA Annealing

The DNA annealing reactions (one reaction for each of the "oligo3", "oligo3a" and "oligo3b" oligonucleotides) were set up in total volumes of 15 µL with the following components: 1 µg of Fd-tet24 phage ssDNA (~4.4 µL), 20 ng (~0.33 µL) of phosphorylated "oligo3", "oligo3a" or "oligo3b", 1 µL each of 1 mM ATP and 6 mM DTT, and 1.5 µL of buffer (500 mM Tris-HCl, 150 mM $MgCl_2$, pH 8.0). The reaction mixtures were incubated at 90ºC (5 min), then at 50ºC (10 min), and finally at room temperature (15 min).

DNA Synthesis

The DNA synthesis reaction was set up in a total volume of 20 µL with the following components: 15 µL of the previously annealed DNA, 0.7 µL of 10 mM ATP, 0.7 µL of 25 mM dNTP, 1.0 µL of 100 mM DTT, 0.5 µL (2.5 units) of T4 DNA ligase (Invitrogen), and 0.7 µL (7 units) of T7 DNA polymerase (New England BioLabs). The reaction mixture was incubated at room temperature for 3 h.

c. Transformation of phage heteroduplex DNA into *E. coli* TG1

Each of the synthesized CDR3/3a/3b heteroduplex phage DNA was purified separately using the QIAquick PCR Purification™ kit (QIAGEN), eluted in 2x 35 µL ddH_2O, and CDR3/3a/3b phage DNA was quantified (~2.6 µg of each) using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

Selection against the uracil-template DNA strand and completion of the CDR3/3a/3b V_L library was achieved by electroporetically transforming the phage heteroduplex DNA into *E. coli* TG1. CDR3/3a/3b heteroduplex DNA was combined (~7.5 µg total) and concentrated down

from ~200 µL (~37 ng/µL) to ~50 µL (~150 ng/µL) using the Savant SpeedVac® DNA 110 Concentrator (Thermo Scientific). 2x 25 µL aliquots of the CDR3/3a/3b heteroduplex DNA was then used to electroporetically transform 2x 350 µL electrocompetent *E. coli* TG1 cells with 2x 3 mL SOC using 0.2 cm gap electroporation cuvettes (Bio-Rad) and a MicroPulser electroporator (Bio-Rad) at the "EC2" setting for creation of the CDR3/3a/3b-randomized V_L phage display library. The ~7 mL transformation material was first combined for equal distribution purposes, then incubated separately (2x ~3.5 mL) at 37ºC for 1 hour (230 rpm). Titers (10^{-2},…,10^{-6}) for the calculation of the library size were carried out by serial dilution in 2xYT. The remaining ~7 mL volume of transformation material was transferred to 90 mL of 2xYT/Tet (12.5 µg/mL) for overnight growth at 37ºC (230 rpm). The following day, the OD_{600} reading for the ~100 mL overnight culture was taken, and the culture was spun down (1,600 g, 4ºC, 15 min), with the pellet resuspended in 10 mL sterile 2xYT/15% glycerol. 50 µL, 100 µL, and 1 mL aliquots were then made and stored at -80ºC.

5. Construction of the CDR1/2/3/3a/3b-randomized synthetic V_L phage display library

a. Isolation of ssDNA from constructed CDR3/3a/3b synthetic V_L phage display library

Phage ssDNA was isolated from the constructed CDR3/3a/3b V_L library for CDR1 and CDR2 randomization. 100 µL of the CDR3/3a/3b library was defrosted from -80ºC and added to 200 mL of 2xYT/Tet (12.5 µg/mL), followed by overnight growth at 37ºC (230 rpm). Phage were precipitated from the overnight culture as previously described. Phage titers (10^{-6},…,10^{-12}) were also prepared as previously described.

1 mL of precipitated phage (corresponding to ~3 x 10^9 clones) containing randomized CDR3, CDR3a, and CDR3b, was used to infect 100 mL of *E. coli* CJ236 log phase culture, with incubation at 37ºC (30 min, 200 rpm). This entire culture (~100 mL) was then added to 300 mL of 2xYT/Tet (12.5 µg/mL)/uridine (2.5 µg/mL) and grown overnight at 37ºC (200 rpm). Phage were precipitated once again using 1/5 volume PEG/2.5 M NaCl with the resulting phage pellet resuspended in 4 mL 1x PBS, and infection titers (10^{-2},…,10^{-10}) with both *E. coli* strains TG1 and CJ236 were again made for comparison of uracil incorporation into phage DNA. Phage ssDNA (now consisting of uracil rather than thymine nucleotides) was purified using the QIAGEN M13 Purification™ Kit (QIAGEN). *In vitro* mutagenesis was carried out for a second time on the newly synthesized phage ssDNA (CDR3/3a/3b) using the CDR1 and CDR2 synthetic oligonucleotides (recipes described in **4.b.**):

b. *In vitro* mutagenesis (CDR1/2 randomization)

DNA Annealing

Since different concentrations of Fd-tet24 phage ssDNA were isolated from the CDR3/3a/3b V_L library, separate annealing reaction mixtures were made. Individual DNA annealing reactions were set up in total volumes of 150 µL with the following components: 10 µg of Fd-tet24 phage ssDNA, 200 ng (~3.3 µL) of phosphorylated "oligo1", 200 ng (~3.3 µL) of phosphorylated "oligo2", 5 µL each of 1 mM ATP and 6 mM DTT, and 15 µL of buffer (500 mM Tris-HCl, 150 mM $MgCl_2$, pH 8.0). The reaction mixtures were incubated at 90ºC (5 min), then at 50ºC (10 min), and finally at room temperature (15 min).

DNA Synthesis

Separate DNA synthesis reactions were established in total volumes of 200 µL with the following components: 150 µL of the previously annealed DNA, 7 µL of 10 mM ATP, 7 µL of 25 mM dNTP, 10 µL of 100 mM DTT, 25 units of T4 DNA ligase (Invitrogen), and 70 units of T7 DNA polymerase (New England BioLabs). Reaction mixtures were incubated at room temperature for 3 h.

c. Transformation of phage heteroduplex DNA into *E. coli* TG1

Each of the CDR1- and CDR2-annealed heteroduplex phage DNA was PCR purified using the QIAquick PCR Purification kit and eluted in 3x 35 µL ddH_2O. Total phage DNA was quantified, after combining CDR1 and CDR2 heteroduplex phage DNA, using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific). This heteroduplex DNA was then transformed into electrocompetent *E. coli* TG1 cells with SOC medium for creation of the complete CDR1/2/3/3a/3b-randomized V_L phage display library.

Fifty-three transformations (corresponding to a total volume of ~88 mL) were carried out as previously described, in 53 mL of SOC medium using 0.2 cm gap electroporation cuvettes (Bio-Rad) and a MicroPulser electroporator (Bio-Rad) at the "EC2" setting. Each transformation consisted of ~2 µg phage heteroduplex DNA electroporated into 350 µL of *E. coli* TG1 cells. The ~88 mL transformation material was then distributed into 20 mL aliquots, followed by incubation at 37ºC (1 hour, 300 rpm). 1 mL from each of the aliquots was taken and combined for titration purposes; ~500 µL was used for making serial dilutions and titers were plated for calculation of the library size were carried out by serial dilution. The remaining ~88 mL volume

of transformation material was transferred to 2x 2 L of 2xYT/Tet (12.5 μg/mL) in 4 L baffled flasks for overnight incubation at 37ºC (230 rpm).

The next day, 1 mL of the ~4.1 L culture was measured for its OD_{600} reading and spun down (4,400 g, 4ºC, 20 min). The supernatant (containing V_L library phage), was precipitated for phage isolation. Resulting phage pellets were resuspended in 10 mL 1x PBS, divided into 1 mL aliquots with 5 μL reserved for titrations, and stored at -80ºC. Bacterial cell pellets that resulted from spinning down the 4.8 L overnight culture (i.e., library cell stock), were serially resuspended in 30 mL 2xYT/15% glycerol and divided into 500 μL aliquots for storage at -80ºC.

6. Identification and sequencing of V_L library clones

Randomly selected clones from the 2xYT/Tet (5 μg/mL) agar library titer plates for colony PCR using universal reverse primers "Fd-96GIIID" (5'-CCCTCATAGTTAGCGTAACGGTAACGATCT-3') and "FdtetGIIID" (5'-GTGAAAAAATTATTATTCGCAATTCCT-3'). Each clone to be amplified was picked from the titer plate and dissolved in 50 μL ddH_2O ("dissolved template") for use in the PCR reaction. PCR reaction mixtures were set up in total volumes of 20 μl for each clone amplified, with the following components: 1 μL of "dissolved template", 10 pmol of "Fd-96GIIID" primer, 10 pmol of "FdtetGIIID" primer, 2 μL of 10x reaction buffer, 0.4 μL of 10 mM dNTP and 1.5 units of *Taq* DNA polymerase (GenScript). The PCR protocol consisted of an initial step at 94ºC for 3 min, followed by 30 cycles of 94ºC for 1 min, 55ºC for 30 s, 72ºC for 30 s and a final extension step at 72ºC for 7 min. Amplified PCR products were electrophoresed on a 1% agarose gel and bands were observed at ~500 bp, corresponding to presence of V_L insert. PCR products were

purified using the QIAquick PCR Purification Kit (QIAGEN) as per the manufacturer's instructions.

Purified DNA clones were sequenced using the dideoxy method (Sanger et al., 1977), and translated into their corresponding protein sequences using the EditSeq software program (DNA Star; Madison, WI, USA). Sequences were aligned by ClustalW for analysis using the MegAlign software program (DNA Star) and numbered according to the Kabat numbering system (Kabat et al., 1991). Actual (observed) frequencies were tabulated for each randomized position in the V_L insert and graphed along with expected frequencies (Table S1 of the "Supplemental" section) for comparison. Expected frequencies were calculated according to the oligonucleotides (Table1) used for each position randomized. The following sample calculation for alanine (A) illustrates how the expected frequencies were obtained (i.e., for position 24 that was randomized with the "NNK" trinucleotide):

Sample calculation for expected frequency of alanine at position 24 ("NNK"):

N = 25% adenine, 25% thymine, 25% cytosine, and 25% guanine

K = 50% thymine and 50% guanine

Alanine = GCU, GCA, GCC, GCG

Since "NNK" allows for a total of 32 possible codons and, in the case of alanine, only GCU and GCG can be encoded, the probability (P_A) that alanine is obtained using the "NNK" trinucleotide is:

$$P_A = (2)/(32)$$

$$= 0.0625$$

Therefore, there is a 6.25% probability that alanine is obtained at position 24 using NNK as a degenerate DNA codon.

Phage clones from each round of subtractive panning were amplified by PCR using the same set of forward and reverse universal primers ("Fd-96GIIID" and "FdtetGIIID", respectively) and the same PCR protocol steps as mentioned above. Amplified products were electrophoresed on a 1% agarose gel for verification of V_L insert size (bands at ~450 bp), and enrichments were assessed by sequencing analysis such that any prevailing clones would appear several times throughout the sequence alignment.

7. Cloning of V_L inserts

Random colonies from the synthetic V_L library 2xYT/Tet (5 μg/mL) agar titer plates were selected for DNA amplification of their V_L inserts. These inserts were then cloned into the ~3 kbp expression vector pSJF2H (Arbabi-Ghahroudi et al., 2009) for expression analysis. All randomly picked colonies were resuspended in 100 μL ddH_2O ("Fd-tet24 colony mixture") and 1 μL of this "Fd-tet24 colony mixture") was used as template for a PCR reaction for the addition of restriction enzyme sites *Bbs*I and *Bam*HI to the 5' and 3' ends of the V_L DNA fragments via gene-specific sense primer "HVL42 *Bbs*I B" (5'-TATGAAGACACCAGGCCGACATCCAGATGACCCAGTCT-3') and anti-sense primer "HVL89 *Bam*HI F" (5'-TTGTTCCGGATCCTAGGACCGTCACCTTGGTCCC-3').

Two identical PCR reaction mixtures were set up in total volumes of 50 μl with the following components: 1 μL of "Fd-tet24 colony mixture", 10 pmol of "HVL89 *Bam*HI F" primer, 10 pmol of "HVL42 *Bbs*I B" primer, 5 μL of 10x reaction buffer, 1 μL of 10 mM dNTP

and 20 units of *Taq* DNA polymerase (GenScript). The PCR protocol consisted of an initial step at 94ºC for 3 min, followed by 30 cycles of 94ºC for 1 min, 55ºC for 30 s, 72ºC for 30 s and a final extension step at 72ºC for 7 min. The amplified PCR products were electrophoresized in a 1% agarose gel and bands were observed at ~300 bp, corresponding to V_L insert size. PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN).

Both PCR products were combined into one tube (~100 µL) and purified using the QIAquick PCR Purification kit (QIAGEN) as per the manufacturer's instructions. 72 µL of ddH_2O was used for DNA elution, which resulted in a yield of 720 ng pure V_L DNA. Both the PCR product and pSJF2H vector were then digested with *Bam*HI (New England BioLabs). Briefly, V_L DNA and pSJF2H vector digestion reaction mixtures were set up in total volumes of 100 µL and 50 µL, respectively. The V_L DNA digestion reaction mixture consisted of 70 µL V_L DNA, 10 µL of 10x *Bam*HI buffer, 10 µL of 10x BSA, and 20 units of *Bam*HI restriction endonuclease. The pSJF2H vector digestion reaction mixture contained 1 µg (~20 µL) of undigested pSJF2H vector, 5 µL of 10x *BamHI* buffer, 5 µL of 10x BSA, and 40 units of *Bam*HI restriction endonuclease. Both digestion reactions were incubated in a 37°C water bath (2 hours). Digested V_L DNA was then purified using the QIAquick PCR Purification Kit (QIAGEN), with DNA elution carried out in 72 µL ddH_2O. 4 µL of the digested pSJF2H vector was loaded onto a 1% agarose gel to verify size against uncut vector, using a 1 kbp ladder (100 V, 20 min). After successful digestion was confirmed by observation of a single band at ~3 kbp, the remaining 46 µL of vector was loaded onto another 1% agarose gel (100 V, 20 min). Bands were cut and purified using the QIAquick Gel Extraction kit (QIAGEN) as per the manufacturer's instructions, with DNA elution in 72 µL ddH_2O. Both purified digestion products were then set up in secondary digestion reactions with *Bpi*I (*Bbs*I), with total volumes

of 80 µL. 70 µL of purified DNA was used as template, along with 8 µL of 10x buffer *G* and 20 units of *Bpi*I (*Bbs*I) restriction endonuclease (Fermentas). A control digestion was also included, in which undigested pSJF2H vector served as the DNA template.

*Bpi*I (*Bbs*I) digestion reactions were incubated in a 37°C water bath (2 hours). 1 µL of alkaline phosphatase (Sigma-Aldrich) was added to the 80 µL pSJF2H vector digestion reaction, followed by incubation for an additional 30 min in a 37ºC water bath, for dephosphorylation of phosphate groups at the 5' ends to prevent formation of self-ligation products. 5 µL of both digested and undigested pSJF2H vector were electrophoresed on a 1% agarose gel for band comparison (100 V, 20 min, 1kbp ladder). *BamHI*/*BpiI* (*BbsI*)-digested insert "mixture" product and *BamHI*/*BpiI* (*BbsI*)-digested pSJF2H vector were purified using the QIAquick PCR Purification Kit (QIAGEN), with the V_L DNA eluted in 51 µL ddH_2O (total yield of ~500 ng), and pSJF2H vector eluted in 62 µL ddH_2O (total yield of ~300 ng).

V_L insert DNA was then ligated into pSJF2H vector, resulting in the addition of C-terminal c-Myc and His_6 tags to V_L inserts. Briefly, a 15:1 molar ratio (V_L DNA):(pSJF2H vector) ligation reaction consisted of ~150 ng of V_L insert DNA, 100 ng of pSJF2H vector, 3 µL of 5x DNA ligase buffer, and 2.5 units of DNA ligase (Invitrogen). A control ligation with water as a replacement for V_L DNA was also included. Ligation reactions were incubated in a 37°C water bath (1 hour).

Two individual transformations were performed in SOC and 50 µL aliquots of electrocompetent *E. coli* TG1 cells; the first consisted of 1 µL of the ligated material (with V_L insert), and the second involved 1 µL of the control ligation material (no V_L insert). Both

transformations were incubated at 37ºC (45 min, 230 rpm). 50 µL, 100 µL, and 200 µL portions were plated individually on 2xYT/Amp (100 µg/mL) agar for overnight growth at 32ºC.

The following day, colony-PCR was performed on 30 randomly picked colonies from the 2xYT/Amp (100 µg/mL) transformation plates, using universal primers "M13FP" (5'-TGTAAAACGACGGCCAGT-3') and "M13RP" (5'-CAGGAAACAGCTATGACC-3'). Each of the 30 amplified PCR products (~5 µL each) were loaded onto a 1% agarose gel for verification of V_L insert presence.

8. Expression of His_6-tagged V_L proteins

Small-scale expression of V_L library clones, using ~10 µL of glycerol stocks as template, was carried out by inoculating 2 mL of B2YT/Amp (100 µg/mL) with ~5 µL from the glycerol stock for overnight growth (27ºC, 230 rpm). The following day, the culture was centrifuged (4,400 g, 4ºC, 10 min.) and the resulting pellet re-suspended in 2mL of fresh B2YT/Amp (100 µg/mL). 0.5 mL of this "pre-culture" was added to 50 mL of B2YT/Amp (100 µg/mL) in a 250 mL baffled flask for incubation at 37ºC (200rpm) until $OD_{600} \approx 0.3$. IPTG was then added at a final concentration of 0.1mM followed by an additional incubation at 37ºC (200rpm) for 24 hours.

Large-scale expression of V_L library clones, using ~10 µL of glycerol stocks as template, was performed by initially growing the clone in 5 mL of 2xYT/Amp (100 µg/mL) overnight at 37 ºC on a rotary shaker (200 rpm). The following day, ~2 mL of this overnight culture was transferred to 100 mL of M9S medium + Amp (100 µg/mL) with added nutrients, for 24 hours on a rotary shaker (200 rpm) at 28ºC. 30 mL of this culture was then transferred to 1 L

M9S/Amp (100 μg/mL) plus added nutrients (previously mentioned) for 24 hour growth on a rotary shaker at 28ºC (200 rpm). Following induction with 0.1 mM IPTG and 10x Induction Medium, the culture was grown for an additional 48 hours with shaking (200 rpm) at 28ºC.

9. Extraction of V_L proteins

a. Total Cell Lysis

Isolation of the V_L protein that was randomly selected from the constructed library for expression analysis was performed according to the "total cell lysis" protocol. First, cells were harvested by centrifugation (4,400 g, 4ºC, 20 min) and placed on ice for all subsequent steps. Cell pellet was resuspended in 100 mL ice-cold lysis buffer (50 mM Tris-HCl pH 8.0, 25 mM NaCl, 2 mM EDTA) and stored overnight at -20ºC. The following day, the frozen suspension was removed from -20ºC. PMSF and DTT were then added at final concentrations of 1 mM and 2 mM, respectively, while the frozen suspension thawed at room temperature with occasional shaking. Cell lysis was accomplished by the addition of 5 mL of freshly prepared lysozyme (100 μg/mL final concentration from a 3 mg/mL aqueous solution) to the thawed suspension, followed by incubation at room temperature for 45 min with occasional shaking until an increase in viscosity was observed. 300 μL of DNase I (Sigma-Aldrich Chemical Co.) (15 U/μL stock in 1 M $MgCl_2$) was then added to the lysate, followed by incubation at room temperature for 30 min or until the suspension became "watery". At this point, 30 μL was reserved for Western blotting as the "total lysate". Soluble and insoluble fractions of the lysate were then separated by centrifugation (23,200 g, 4ºC, 30 min). Centrifugation was repeated until the supernatant was clear, and 30 μL was reserved for Western blot analysis. The supernatant was then dialyzed overnight at 4ºC against 6 L of buffer "A" solution (10 mM HEPES, 500 mM NaCl, pH 7.4) for

IMAC buffer exchange using cellulose dialysis tubing (Sigma-Aldrich). The supernatant was filtered the following day using a Millipore 0.22 μm bottle top filter unit with a PVDF Membrane. The lysate pellet was resuspended in 80 mL ice-cold lysis buffer and kept at -20ºC, with 30 μL reserved for the Western blot.

b. Periplasmic Extraction

Proteins of the V_L subtractive panning isolates were extracted from the bacterial periplasm using TES buffer (0.2 M Tris-HCl pH 8.0, 0.5 mM EDTA, 20% sucrose). Cells were harvested from the 1 L culture by centrifugation (4,400 g, 4ºC, 20 min), with resulting pellets resuspended in 30 mL TES (30 mL of TES buffer used for every 1 L of culture) and 300 μL of 100 mM PMSF, followed by incubation on ice for 1 hour. 45 mL of TES buffer diluted in ddH_2O (v/v) was then added to each resuspended pellet, and samples were left for an additional incubation on ice for 1 hour. Proteins were collected into the supernatant by high-speed centrifugation (23,200 g, 4ºC, 30 min), then dialyzed overnight for IMAC buffer exchange against 6 L of buffer "A" solution, at 4ºC. Proteins were then filtered using a Millipore 0.22 μm bottle top filter unit with a PVDF Membrane.

10. Western blot

10 μL of the protein supernatant was mixed with 10 μL of 2x SDS-loading dye (0.005% Bromophenol Blue, 4% SDS, 100 mM Tris-Cl pH 6.8, 20% glycerol, 200 mM DTT added immediately before use) and boiled at 100°C (5 min). Similarly, culture pellet was resuspended in ddH_2O and 10 μL were used to mix with 10 μL 2x SDS-loading dye followed by boiling (100°C, 5 min). Samples were loaded onto 12.5% SDS-PAGE gels and electrophoresed (200 V, 30 min) followed by transfer to a PVDF membrane (Roche Diagnostics, Laval, QC, Canada) for

1 hour (100 V) using Transfer buffer (25 mM Tris-HCl buffer pH 8.0, 192 mM glycine, 4% methanol). Blocking of the membrane was carried out with 30 mL of 3% casein-TBST (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 0.05% Tween-20) for 1 hour. Mouse anti-His_6 IgG (Abcam) diluted 1:2,500 in TBST (4 µL of anti-6His IgG in 10 mL of TBST) was added to the membrane for 1 hour followed by 3 consecutive washes with 10 mL of TBST. AP-labeled goat anti-mouse IgG conjugate (Cedarlane Laboratories, Burlington, ON, Canada) diluted at a ratio of 1:5,000 in TBST (2 µL of goat anti-mouse IgG conjugate in 10 mL of TBST), was used to probe the membrane for 1 hour followed by a final set of 3 consecutive washes with 10 mL of TBST. Membranes were then subjected to 25 mL of AP substrate (Bio-Rad) (25 µL of substrate A + 25 µL substrate B diluted in 25 mL TBST) for 7 min, washing in dH_2O and drying by air.

11. Purification by Immobilized Metal Affinity Chromatography

V_L protein was purified by immobilized metal affinity chromatography (IMAC) using a HiTrap Chelating 5 mL column (GE Healthcare) charged with Ni^{2+} ions from a 5 mg/mL $NiCl_2$ solution. The column was first attached to a single-channel Peristaltic P1 Pump (Pharmacia/GMI, Albertville, Minnesota, USA) with the flow rate set at ~1 mL/min. Approximately 25 mL of degassed ddH_2O was passed through the column, followed by ~25 mL of a 5 mg/mL $NiCl_2.6H_2O$ solution; the column was then washed again with ~25 mL of degassed ddH_2O. Next, ~20 mL of degassed, filter-sterilized buffer "A" was passed through the HiTrap Chelating 5 mL column followed by the protein sample to be purified (which was dialyzed against buffer "A"). The protein sample flow-through (i.e., components that did not bind to the Ni^{2+}-coated column) was also collected and saved in the event that protein was eluted without column binding. Freshly prepared buffer "A" and buffer "B" solution (10 mM HEPES, 500 mM

NaCl, 500 mM Imidazole, pH 7.4) were degassed, and the HiTrap Chelating 5 mL column was then attached to the AKTA fast protein liquid chromatography (FPLC™) system (GE Healthcare) for protein purification, according to the manufacturer's instructions. Ni^{2+}-bound protein was eluted in 1 mL fractions with a 10-500 mM imidazole gradient, using buffer "B" solution. The fractions corresponding to the "eluted" peaks on the chromatogram were collected for protein purity analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). OD_{280} values for each V_L were measured for calculation of protein concentration using molar extinction coefficients that were computed with the "ProtParam" tool of ExPASy software (Swiss Institute of Bioinformatics, Geneva, Switzerland).

12. Size Exclusion Chromatography

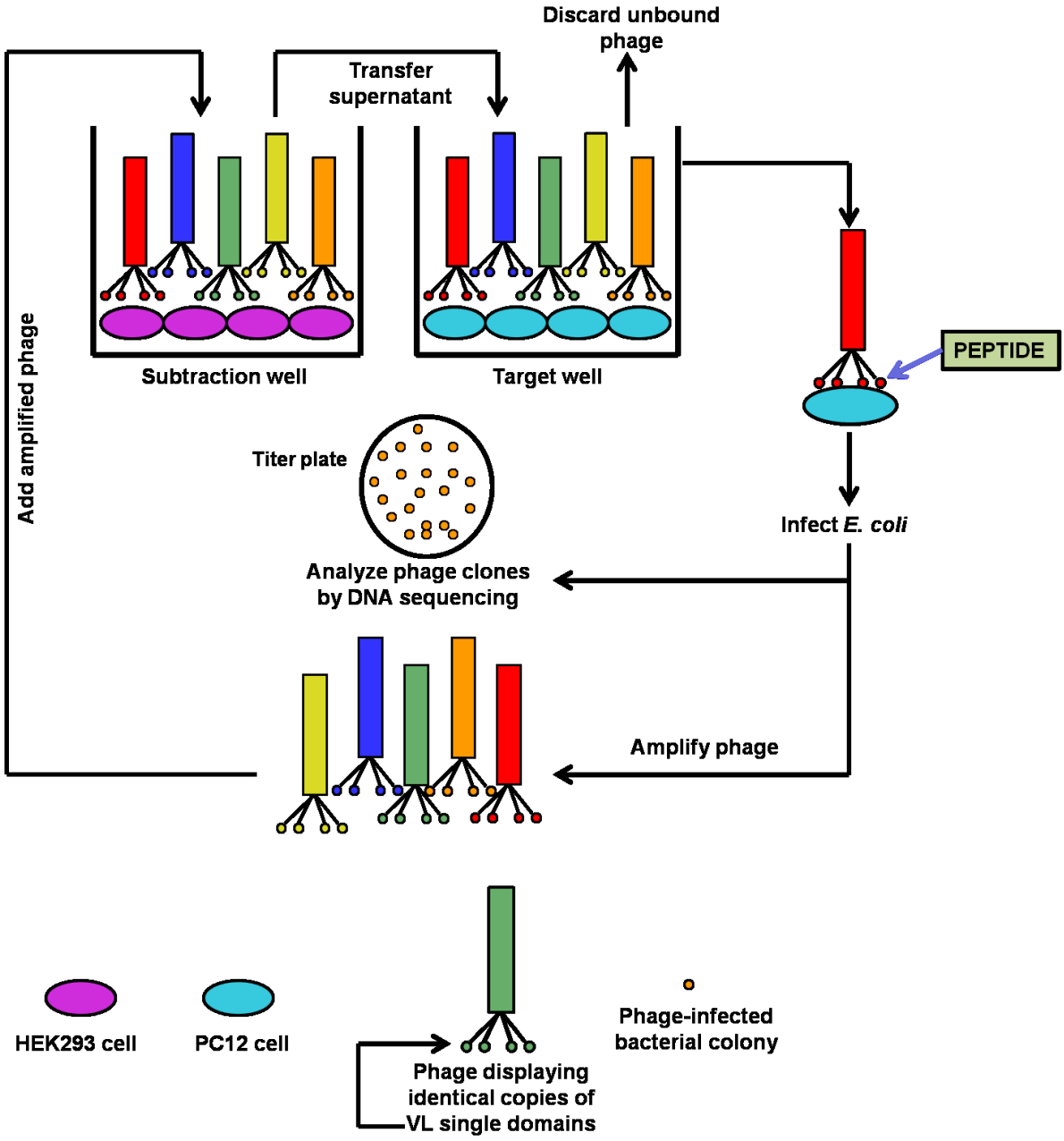
Size exclusion chromatography (SEC) was performed on all purified V_Ls using a Superdex^{TM} 75 column (GE Healthcare) for verification of their state of aggregation under the control of an AKTA^{TM} FPLC (GE Healthcare), as per the manufacturer's instructions. Following attachment of the column to the AKTA^{TM} FPLC, the column was washed with degassed ddH_2O at a flow rate of 0.5 mL/min for 90 min. A baseline was then established by passing degassed, filter-sterilized "SEC buffer" solution (10 mM HEPES, 150 mM NaCl, 0.005% (v/v) P-20 detergent, pH 7.4) through the column at 0.5 mL/min for 90 min or at 0.1 mL/min overnight. Purified protein sample was then injected into the column using a syringe, and protein elution was carried out using"SEC buffer" solution. Sample fractions from the size exclusion column were collected and individually dialyzed against 4 L of 1x PBS overnight at 4ºC four consecutive times for use in flow cytometric analysis.

Protein properties including molecular weight, molar extinction coefficient, and isoelectric point (pI) were calculated using the "Compute pI/MW" and "ProtParam" tools of ExPASy.

13. Subtractive panning of the constructed synthetic V_L phage display library

The steps involved in the subtractive panning of the synthetic V_L library are schematically represented in Figure 9. Two milligrams of competitively binding NRP1 inhibitory peptide (HAVEHGFMQTLLKVTLE) containing part of the sequence for Sema3A (GenScript) peptide was dissolved in 900 µL of ddH_2O and distributed into 45 µL aliquots. When needed for competitive elution, 5 µL of 10x PBS was added to one aliquot of 45 µL ddH_2O-dissolved peptide (final concentration of 1 mM).

Figure 9: Scheme for subtractive panning of the synthetic V_L phage display library. The library was subtractively panned against two mammalian cell lines. Pheochromocytoma (PC12) cells of the rat adrenal medulla, a tumor cell line that endogenously expresses NRP1, served as the target cell line while Human embryonic kidney- 293 (HEK293) cells served as the subtractor cell line which did not express NRP1. Blocked phages were pooled with HEK293 cells, spun down, and the depleted supernatant transferred to PC12 cells. Competitive elution (the recovery of bound phage to the target site) was carried out by adding a synthetic Sema3A "Ig" peptide (at a concentration of 1mM) to the cell pellet. Competitively eluted phage were used to infect *E. coli* TG1 log phase cells with an aliquot used to make titers plated on agar media for clone sequencing purposes, while the remaining infected *E. coli* was incubated overnight amplification (adapted from Arbabi-Ghahroudi et al., 2009).


Discard unbound phage
Transfer supernatant
Add amplified phage
Subtraction well
Target well
PEPTIDE
Titer plate
Infect E. coli
Analyze phage clones by DNA sequencing
Amplify phage
HEK293 cell
PC12 cell
Phage displaying identical copies of VL single domains
Phage-infected bacterial colony

The constructed synthetic V_L phage display library was subjected to four rounds of subtractive panning against two mammalian cell lines. Pheochromocytoma (PC12) cells of the rat adrenal medulla, a tumor cell line that expresses NRP1 endogenously (Lee et al., 2002), were used as the positive cell line while Human embryonic kidney- 293 (HEK293) cells, which do not express NRP1, served as a negative cell line. Both adherent cell lines were grown and generously provided by the Experimental Stroke Research Group (National Research Council, Ottawa, ON). 2 x 10^{13} phages (~100 µL) were first incubated in 2 mL of fresh tissue culture medium (D-MEM high-glucose supplemented with 10% FBS and 2 mM L glutamine) with 2.5% BSA (1 hour). 4 x 10^6 (~100 µL) HEK293 cells were spun down (705 g, 4ºC, 1 min) and resuspended in 2 mL fresh tissue culture medium.

Blocked phages were then mixed with HEK293 cells, such that the cell pellet was resuspended in blocked phage, followed by incubation at room temperature for 1 hour. 4 x 10^6 (~100 µL) PC12 cells were centrifuged (705 g, 4ºC, 1 min) and the cell pellet was resuspended in 2 mL fresh tissue culture medium (D-MEM high-glucose supplemented with 10% horse serum, 5% FBS, and Antibiotic/Antimycotic 100x Solution). This step was repeated a second time. After the 1 hour incubation of phage with HEK293 cells, this mixture was centrifuged (705 g, 4ºC, 1 min) and the HEK293-depleted phage supernatant added to the PC12 cell pellet for resuspension; this was followed by an incubation period of 1 hour at room temperature with occasional gentle agitation. After centrifugation (705 g, 4ºC, 1 min) to pellet PC12 cell-bound phage, the supernatant was discarded and the cell pellet was washed twice with 0.05% PBST (1x PBS containing 0.05% (v/v) Tween-20) and twice more with 1x PBS. Wash steps were carried out by gently resuspending the pellet in 1 mL 0.05% PBST, transferring the suspension to a fresh tube, and centrifuging (705 g, 4ºC, 1 min). After a second 0.05% PBST wash, the pellet was

gently resuspended in 1 mL 1x PBS, transferred to a fresh tube, and centrifuged (705 g, 4ºC, 1 min).

Bound phage were competitively eluted by adding 50 µL of a synthetic Sema3A "Ig"-containing peptide (at a concentration of 1 mM) to the phage-bound PC12 cell pellet with resuspension and an incubation of 1.5 hours at room temperature. After centrifugation (4,412 g, 4ºC, 2 min), the supernatant (eluted phage) was used to infect 30 mL of *E. coli* TG1 log phase cells and incubated in a water bath at 37°C (30 min). A 5 µL aliquot was reserved for phage titrations and the remaining 30 mL infection culture was centrifuged (1,600 g, 4°C, 10 min), and the resulting pellet resuspended in 1 mL of 2xYT. The entire 1 mL of phage-infected *E. coli* TG1 was spread onto 4 large round dishes containing 2xYT/Tet (5 µg/mL) agar for overnight amplification at 37ºC.

Phage titers were determined by infection of *E. coli* TG1; input and output phage from each round of panning were determined by comparing the number of CFUs on 2xYT/Tet (5 µg/mL) agar titer plates that were made before (output) and after (input) overnight amplification in *E. coli* TG1. The following day, cells were collected by scraping the large dishes using 2x 5 mL of 2xYT per plate and a glass spreader. Cells were then combined and mixed by vortexing, followed by a centrifugation step (1,600 g, 4°C, 10 min). Cell pellets were resuspended in 2xYT/15% glycerol and stored at -80°C as part of the cell stock. The phage-containing supernatant was pelleted at high-speed (13,800 g, 4°C, 10 min,) and passed through a Millipore 0.45 µm GP Express™ Plus Membrane filtration system. Phage were then precipitated with 1/5 volume PEG/2.5 M NaCl and kept on ice for 1 hour, after which centrifugation (4,400 g, 4°C, 30 min) resulted in a pellet that was subsequently resuspended in 5 mL of 1x PBS. 1 mL of

PEG/2.5 M NaCl was then added to the 5 mL resuspended phage for further precipitation, with incubation on ice (20 min) then centrifugation (3,000 g, 4°C, 30 min) and a final pellet resuspension in 1 mL of 1x PBS. Titers (10^{-6},…,10^{-12}) of the precipitated phage were made as previously described, and CFUs counted on each plate enabled calculation of the number of phage and determination of what volume would correspond to 1 x 10^{12} phage for use in the subsequent round of subtractive panning (adapted from Lee et al., 2007).

14. Flow Cytometry

Flow cytometry was carried out to test protein binding of four isolated V_L clones to both PC12 and HEK293 cells, using an anti-6His PE (phycoerythrin)-conjugated secondary Ab (Miltenyi Biotec). Controls included each cell line incubated 1) alone, 2) with anti-6His PE, and 3) with a V_L clone not isolated from subtractive panning. 100 µL of cells at a concentration of 5 x 10^6 per mL were washed once with 15 mL of PBS-1% BSA and spun down (705 g, 4ºC, 2 min) with the resulting pellet resuspended in 50 µL of PBS-2% BSA. Cells were then incubated either alone, with 10 µL of anti-6His PE Ab, with 5 µg of each of the four V_L clones and 10 µL of anti-6His PE Ab, or with 5 µg of the V_L control and 10 µL of anti-6His PE Ab. Each of the reaction mixtures were incubated on ice for 1 hour. Following a PBS wash, consisting of centrifugation (705 g, 4°C, 2 min) and resuspension of the cell pellet in 3 mL 1x PBS, each of the reaction mixtures (except for the "cells alone" reactions) was incubated with 10 µL of anti-6His PE for 1 hour on ice. After another wash with 3 mL 1x PBS, cell reaction mixtures were fixed with 300 µL of PBS-1% formaldehyde. Data was acquired out using a BD FACSAriaII cell sorter (BD Biosciences) and the BD FACS Diva software (BD Biosciences).

VI. RESULTS

1. Efficiency of transformation for freshly prepared electrocompetent *E. coli* TG1 cells

A final volume of ~20 mL *E. coli* TG1 electrocompetent cells was obtained (please refer to **3. Preparation of *E. coli* TG1 electrocompetent cells** of the "Materials & Methods") and divided into aliquots for library construction purposes. Transformation efficiency of the prepared electrocompetent cells was verified prior to constructing the synthetic V_L library, using a variety of approaches.

First, transformation efficiencies of the freshly prepared electrocompetent cells were compared to commercially available electrocompetent cells (VWR International, Mississauga, ON, Canada) using 0.1 ng of pUC18 vector (~2.7 kbp; GenScript) as DNA. For the prepared electrocompetent cells, ~300 CFUs were counted on the 10^{-2} plate, and ~26 CFUs on the 10^{-3} plate. Efficiency of transformation was calculated using the 10^{-3} plate:

(26 CFUs)(10^3 dilution factor)

$= 3 \times 10^4$ CFUs in plated volume without dilution

(3×10^4 CFUs)/(100 µL volume used to plate)

$= 3 \times 10^2$ CFUs/µL plated volume

(3×10^2)(1,000 µL volume used to plate)

$= 3 \times 10^5$ CFUs/mL transformation reaction

(3×10^5 CFUs/mL)/(0.1 ng pUC18 vector)

$= 3 \times 10^6$ CFUs/ng pUC18 vector

(3×10^6 CFUs/ng)(1,000 ng/1 µg)

$= 3 \times 10^9$ CFUs/µg DNA

Therefore, the transformation efficiency for the freshly prepared *E. coli* TG1 electrocompetent cells was calculated to be ~3 x 10^9 CFUs/µg DNA.

For the commercially available electrocompetent cells, ~60 CFUs were found on the 10^{-2} plate, and ~7 CFUs on the 10^{-3} plate. Efficiency of transformation was calculated to be ~7 x 10^8 CFUs/µg DNA (~4x lower than the transformation efficiency obtained for the prepared electrocompetent cells).

Next, transformation efficiency of both freshly prepared and commercially available electrocompetent cells was tested using ~250 ng of heteroduplex phage DNA obtained from mutagenesis and randomization. For the freshly prepared electrocompetent cells, it was found that the 10^{-3} plate contained ~1,500 CFUs, the 10^{-4} plate consisted of ~180 CFUs, and 5 CFUs were observed on the 10^{-6} plate. Therefore, using the 10^{-4} plate, a transformation efficiency of ~8 x 10^7 CFUs/µg heteroduplex phage DNA was calculated. For the commercial electrocompetent cells, ~700 CFUs were counted on the 10^{-3} plate, while ~80 CFUs were found on the 10^{-4} plate. Therefore, using the 10^{-4} plate, a transformation efficiency of ~3 x 10^7 CFUs/µg heteroduplex phage DNA was calculated (approximately 2x lower than the efficiency obtained for the freshly prepared electrocompetent cells).

Finally, two different final concentrations of Tet (12.5 µg/mL and 5 µg/mL on the agar plates, separately) were used with heteroduplex phage DNA to assess any significant difference in transformation efficiency for the freshly prepared electrocompetent cells. ~300 ng of heteroduplex DNA was electroporated into 50 µL aliquots of TG1 electrocompetent cells, with 100 µL of diluted transformation material (10^{-2},…,10^{-6}) spread onto 2xYT/Tet (12.5 µg/mL and 5 µg/mL, separately) agar plates for overnight selection (37°C). For the plates containing a Tet

concentration of 12.5 μg/mL, ~100 CFUs were counted on the 10^{-4} plate and 11 CFUs were observed on the 10^{-5} plate. These numbers corresponded to a transformation efficiency of ~3 x 10^7 CFUs/μg heteroduplex DNA. For the plates containing a Tet concentration of 5 μg/mL, ~300 CFUs were counted on the 10^{-4} plate, 30 CFUs were found to be on the 10^{-5} plate, and 4 CFUs were observed on the 10^{-6} plate. Therefore, the transformation efficiency was calculated to be ~1 x 10^8 CFUs/μg heteroduplex DNA using agar plates containing a Tet concentration of 5 μg/mL. This transformation efficiency was ~3x higher than that obtained for the agar plates containing a Tet concentration of 12.5 μg/mL.

2. Synthetic V_L phage display library

A synthetic V_L phage display library was created in *E. coli* TG1 using a human V_L gene ("V_L-24") as template. V_L-24, containing a V_L scaffold (321 base pairs; 107 amino acids) and sub-cloned into the Fd-tet phage vector, was obtained as a gift. A human naïve library was previously screened against protein L (an IgG light chain-binding protein expressed by anaerobic bacterial strains such as *Peptostreptococcus magnus* (Akerström and Björck, 1989; Nilson et al., 1992)) through which several protein L binding V_Ls were isolated; among the isolates was V_L-24. Protein L binds the V_L domain through the V_L framework regions as opposed to the CDRs, allowing for the selection of V_Ls displaying a high degree of structural integrity influencing V_L stability and solubility. In comparison to the remaining V_L isolates, V_L-24 demonstrated a number of favourable biophysical properties, including non-aggregation, resistance to gastro-intestinal proteases (e.g., trypsin, chymotrypsin, and pepsin), a high Tm (melting temperature), high protein expression, reasonable Thermal Refolding Efficiency, and a relatively high affinity for protein L. V_L-24 also deviated by only 2 amino acids from the original V germline sequence; germline cells are cells in the reproductive organs of multicellular organisms that divide by

meiosis to produce gametes (O'Connor, 2008). Therefore, V_L-24 was used as a scaffold for the generation of a diverse synthetic V_L library with the intention of obtaining high-affinity V_L clones displaying considerable stability and solubility.

A synthetic V_L phage display library containing randomized CDR3/3a/3b (three different lengths of CDR3; 9, 10, and 11 residues long, respectively) regions was first constructed using *in vitro* mutagenesis with degenerate oligonucleotides. Phage ssDNA with randomized CDR3/3a/3b was then isolated and purified for production of phage heteroduplex DNA with randomized CDR1/2/3/3a/3b via a second *in vitro* mutagenesis procedure. CDR1/2/3/3a/3b heteroduplex DNA was electroporated into *E. coli* TG1 electrocompetent cells for generation of the CDR1/2/3/3a/3b-randomized synthetic V_L phage display library.

A large proportion of the synthetic oligonucleotides (please refer to Table 1 for respective DNA sequences) used for *in vitro* mutagenesis were based on the "NNK" degenerate trinucleotide, where "N" represents an equal proportion of the four nucleotides adenine, thymine, guanine, and cytosine; and "K" represents thymine and guanine. For position 28 (CDR1), the nucleotide "RGC" was used to obtain either a serine (S) or glycine (G), and for position 54 (CDR2), "CKN" was used in obtaining either leucine (L) or arginine (R). The use of "NNK" allows for the use of all 20 amino acids while minimizing the incorporation of stop codons (i.e., to a probability of 1 out of 32). Previous studies examining the randomization of CDR positions within V_L domains revealed the specific positions throughout the CDRs that are critical in Ab-Ag interaction. The CDR randomization design described here was largely based on previous sequence alignment results of the V_L isolates (including germline and non-germline, against Protein L) that revealed a high degree of sequence variability in the 3 CDRs. The developed

randomization strategy took into account previous studies which indicated what specific positions and amino acids within the CDRs were critical for protein stability and solubility. Since CDR3 length also plays an important role in stability and solubility, 3 different lengths of CDR3 were incorporated into the creation of the synthetic V_L library. Diversification of highly conserved residues at specific positions (i.e., 28 and 54) was limited, as these amino acids are believed to contribute additional structural stability. Oligonucleotides were annealed to DNA regions corresponding to the 3 CDRs for randomization of the following positions (28, 30, 31, 32, 34, 50, 53, 54, 55, 89, 91, 93, 94, 95[a/b], 96) of the V_L24 template sequence (for recipes, please refer to **4.a. Construction of the CDR3/3a/3b-randomized synthetic V_L phage display library** of the "Materials & Methods").

The mutagenesis method used in creating a synthetic library consisting of V_Ls with randomized CDRs is based on the use of a DNA template containing uracil residues in place of thymine; the uracil-containing DNA is produced within an *E. coli dut*$^-$ *ung*$^-$ strain (e.g., CJ236). In the combined *dut*$^-$ *ung*$^-$ mutant, deoxyuridine is incorporated into DNA in place of thymine and is not removed. Phosphorylated synthetic oligonucleotides (i.e., containing the degenerate nucleotides "N", "K", and "R") are annealed *in vitro* to the targeted positions of the wild-type sequence (i.e., V_L-24) using DNA polymerase, ligase and a dNTP mixture. Uracil is then removed from the template strand by uracil *N*-glycosylase and glycosylase treatment is achieved by introducing the DNA products of the *in vitro* incorporation reaction into *ung*$^+$ *E. coli* (e.g., TG1) cells, whose enzymes degrade the resulting uracil-containing strand, leaving the complementary strand which contains the randomized positions.

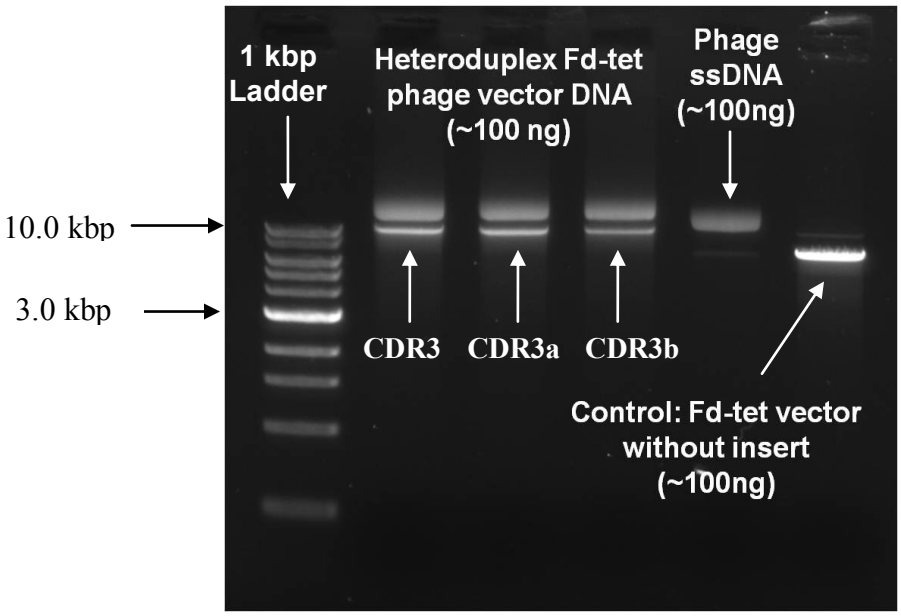
a. Construction of the CDR3/3a/3b-randomized synthetic V_L phage display library

Infection titers for *E. coli* strains TG1 and CJ236 were compared to assess efficiency of uracil incorporation into phage ssDNA. 2 colonies were counted on the 10^{-3} TG1 dilution plate, while the 10^{-8} CJ236 dilution plate displayed 8 colonies. CJ236 phage infection was therefore ~40,000-fold more efficient than TG1 infection, signifying successful and efficient incorporation of uracil into phage ssDNA. Phage ssDNA (containing uracil instead of thymine nucleotides) was subsequently isolated and purified from the *E. coli* strain CJ236. A total yield of ~23 µg ssDNA was obtained.

The first *in vitro* mutagenesis procedure was carried out with "oligo3", "oligo3a" and "oligo3b"; each of the synthesized CDR3/3a/3b heteroduplex phage DNA was purified separately, eluted in 2x 35 µL ddH_2O, and CDR3/3a/3b-randomized phage DNA was quantified (~2.6 µg of each) using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

Formation of heteroduplex DNA was confirmed by loading ~100 ng of CDR3/3a/3b DNA (Figure 10, lanes 2-4) on a 0.7% agarose gel, with phage ssDNA and Fd-tet (phage vector without V_L insert) as controls. Heteroduplex DNA (lanes 2-4) was characterized by the appearance of two bands in contrast to the single band observed for phage ssDNA (lane 5). The lower band represents supercoiled DNA and the upper band indicates linear DNA (lanes 2-4) (Gabibov and Makarevitch, 1995). Size of the Fd-tet vector (linearized) without V_L insert was also verified, with band appearance at ~9.2 kbp (lane 6).

Figure 10: 0.7% agarose gel profile of heteroduplex DNA formation following *in vitro* mutagenesis with synthetic randomized oligonucleotides "oligo3", "oligo3a", and "oligo3b". From left to right: Lane 1 = 1 kbp ladder (New England BioLabs), Lane 2 = 100 ng of *CDR3* heteroduplex DNA, Lane 3 = 100 ng of *CDR3a* heteroduplex DNA, Lane 4 = 100 ng of *CDR3b* heteroduplex DNA, Lane 5 = 100 ng of phage ssDNA (before *in vitro* mutagenesis), Lane 6 = 100 ng of Fd-tet phage vector without V_L insert (control). Photograph was taken with AlphaImager 3400 (Alpha Innotech Corporation, San Leandro, CA).


1 kbp
Ladder
Heteroduplex Fd-tet
phage vector DNA
(~100 ng)
Phage
ssDNA
(~100ng)
10.0 kbp
3.0 kbp
CDR3
CDR3a
CDR3b
Control: Fd-tet vector
without insert
(~100ng)

The constructed CDR3/3a/3b V_L library was found to consist of ~5 x 10^7 clones based on 2xYT/Tet (5 μg/mL) titer plates; 74 CFUs were counted on the 10^{-4} plate, and the total transformation cell reaction volume was ~6.7 mL. Therefore, using the 10^{-4} titer plate:

(74 clones)(10^4 dilution factor)

= 7 x 10^5 total clones in plated volume without dilution

(7 x 10^5)/(100 μL volume used to plate)

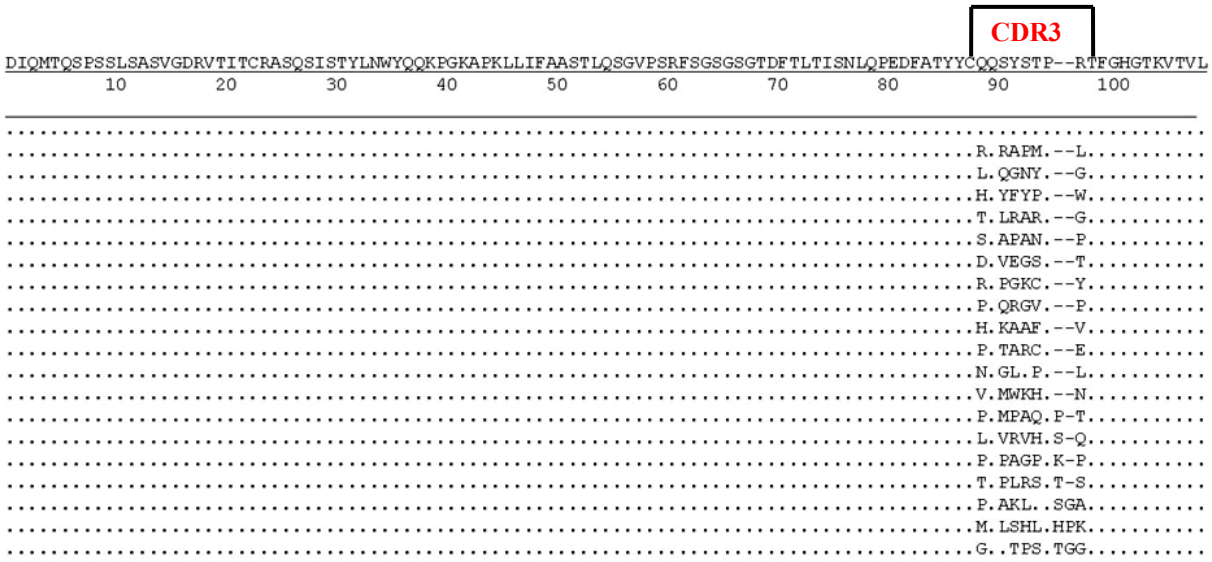
= 7 x 10^3 clones/μL plated volume

(7 x 10^3 clones/μL)(6,700 μL total volume)

= ~5 x 10^7 clones in the 6.7 mL total reaction

Twenty clones randomly selected from the library's titer plates were electrophoresed on a 1% agarose gel to verify presence of insert at ~450 bp using the universal forward and reverse primers "FdtetGIIID" and "Fd-96GIIID" (please refer to "Materials and Methods" section for respective primer sequences) and sequenced using the EditSeq program (DNAStar) followed by ClustalW alignment (MegAlign program, DNAStar). Of the 20 clones, 1 was found to be "wild-type" (i.e., sharing 100% amino acid identity with the V_L-24 template, in terms of both CDR3 and framework regions). Twelve clones were found to be of typical CDR3 length (i.e., 9 amino acids), 4 clones contained one amino acid insertion (i.e., 10 amino acids in CDR3), and 3 clones had two amino acid insertions (i.e., 11 amino acids in CDR3). Aligned amino acid sequences of the 20 clones are displayed in Figure 11, along with the V_L-24 template sequence.

Figure 11: Aligned protein sequences of 20 randomly picked clones from the CDR3/3a/3b-randomized synthetic V_L phage display library. Dots in the sequence entries indicate amino acid identity with the human V_L-24 template sequence on which mutagenesis was performed; dashes are included for sequence alignment, and CDR3 is boxed. The Kabat numbering system is used (Kabat et al., 1991).


CDR3
DIQMTQSPSSLSASVGDRVTITCRASQSISTYLNWYQQKPGKAPKLLIFAASTLQSGVPSRFSGSGSGTDFTLTISNLQPEDFATYYCQQSYSTP--RTFGHGTKVTVL
10
20
30
40
50
60
70
80
90
100

b. Construction of the CDR1/2/3/3a/3b-randomized synthetic V_L phage display library

CDR3/3a/3b-randomized phage ssDNA was isolated from the newly constructed V_L library and following phage precipitation, phage titers plated on 2xYT/Tet (5 μg/mL) revealed ~287 CFUs on the 10^{-6} plate and 3 CFUs on the 10^{-8} plate. Therefore, a concentration of 3 x 10^6 phage/μL was calculated using the 10^{-6} plate:

(~300 CFUs)(10^6 dilution factor)

= 3 x 10^8 total phage in plated volume

(3 x 10^8 total phage in plated volume)/(100 μL volume used to plate)

= 3 x 10^6 phage/μL

CDR3/3a/3b-randomized isolated phages were used to infect CJ236 once again with added uridine, in order to carry out mutagenesis with "oligo1" and "oligo2" for randomization of CDRs 1 and 2. Infection titers were compared between *E. coli* strains TG1 and CJ236 to assess efficiency of uracil incorporation into phage ssDNA. 141 CFUs were found on the CJ236 10^{-6} titer plate, while 30 CFUs were counted on the 10^{-3} TG1 plate. This translated to a ~5,000-fold increase in uracil incorporation into phage ssDNA. Phage ssDNA (consisting of uracil instead of thymine nucleotides) was then purified, with a total yield of ~65 μg ssDNA that contained randomized CDR3/3a/3b regions.

In vitro mutagenesis was carried out using the newly synthesized phage ssDNA (CDR3/3a/3b-randomized) and the CDR1 and CDR2 synthetic oligonucleotides ("oligo1" and "oligo2", as described in **5.a. Construction of the CDR1/2/3/3a/3b-randomized synthetic V_L phage display library** of the "Materials & Methods"). Each of the CDR1- and 2-annealed heteroduplex phage DNA was PCR purified, and eluted in 3x 35 μL ddH_2O. ~100 μg of total

phage DNA was quantified, when combined, using the NanoDrop (Thermo Scientific). Heteroduplex DNA was transformed into electrocompetent *E. coli* TG1 cells for creation of the final CDR1/2/3/3a/3b-randomized V_L phage display library. Fifty-three transformations (in ~88 mL) were carried out as described (**5.c. Transformation of phage heteroduplex DNA into *E. coli* TG1** of the "Materials and Methods")**.** Titers for the calculation of the library size were carried out by serial dilution, and the remaining transformation material (~88 mL) was transferred to 2x 2 L of 2xYT/Tet (12.5 µg/mL) for overnight incubation at 37ºC (250 rpm). The 2xYT/Tet (12.5 µg/mL) 4.1 L *E. coli* TG1 culture was measured for its OD_{600} reading (~2.2) following overnight growth. Bacterial cell pellets that resulted from spinning down the 4.1 L overnight culture (i.e., library cell stock), were resuspended in 30 mL 2xYT/15% glycerol and divided into 500 µL aliquots for storage at -80ºC. Cell concentration was calculated to be 5 x 10^{10} cells per 500 µL aliquot using the recorded OD_{600} reading of ~2.2, where an OD_{600} reading of 1.0 ≈ 3 x 10^8 *E. coli* cells/mL (Mülhardt, 2007).

Following phage precipitation, phage titer plates were grown overnight for calculation of the total number of phage per mL of 1x PBS (~10^{14} phage/mL). ~12 CFUs were counted on the 10^{-12} plate, therefore:

(12 CFUs)(10^{12} dilution factor)

= 1 x 10^{13} total phage in plated volume

(1 x 10^{13} total phage in plated volume)/(100 µL volume used to plate)

= 1 x 10^{11} phage/µL plated volume

(1 x 10^{11} phage/µL)(1000 µL/1 mL)

= 1 x 10^{14} phage/mL 1x PBS

Titer plates made using the library transformation material indicated were found to be inconsistent, as discrepancies were observed regarding the number of CFUs counted on one dilution plate versus those of another. The 10^{-3} plate revealed ~400 CFUs, while both the 10^{-4} and 10^{-5} plates showed similar colony growth numbers (~150 CFUs and ~100 CFUs, respectively). It was also noted that the 10^{-5} plate displayed a few irregularities in terms of colony dispersion and clumping (i.e., growing of individual colonies in very close proximity to one another). The 10^{-6} and 10^{-7} plates also shared similar colony numbers between them, with 13 CFUs and 14 CFUs respectively counted; 2 CFUs were found on the 10^{-8} plate.

Due to the variation in titer plate results, the library size was determined from the transformation efficiency previously obtained for the freshly prepared *E. coli* TG1 electrocompetent cells using 300 ng of pure phage heteroduplex DNA (1×10^8 clones/μg heteroduplex DNA; please refer to **1. Efficiency of transformation for prepared electrocompetent *E. coli* TG1 cells** of the "Results" section). ~100 μg of phage heteroduplex DNA was used to create the V_L repertoire; therefore, it was estimated that the library contained a diversity of 10^{10} clones.

3. Sequencing and statistical analysis of the constructed synthetic V_L phage display library

The quality of an Ab phage display library is characterized by its completeness or diversity (i.e., the proportion of all possible amino acid sequences that have been incorporated into the library). To assess its quality/diversity, the synthetic V_L library was characterized by sequence and statistical analysis. One hundred clones were randomly selected from the constructed library's titer plates (colonies of transformed bacteria after their electroporation with mutagenized phage DNA), and electrophoresed on a 1% agarose gel to verify presence of insert

at ~450 bp using the universal forward and reverse primers "FdtetGIIID" and "Fd-96GIIID" (please refer to "Materials and Methods" section for respective primer sequences). Following sequencing of each sample, the DNA sequences were translated to protein sequences using the EditSeq program (DNAStar) and aligned by ClustalW method with MegAlign (DNAStar). Of the 100 clones, 3 were found to be "wild-type" sequences (i.e., shared 100% amino acid identity with the V_L-24 template). In addition, 10 other clones were "wild-type" only in CDR3 and 2 were "wild-type" for only CDR2. With respect to CDR3, 42 clones had one amino acid insertion (i.e., termed "CDR3a" signifying CDR3 lengths of 10 amino acids) and 16 were found to have 2 insertions (i.e., termed "CDR3b" signifying CDR3 lengths of 11 amino acids). Aligned sequences of the 100 randomly picked clones are displayed in Figure 12, along with the V_L-24 template sequence.

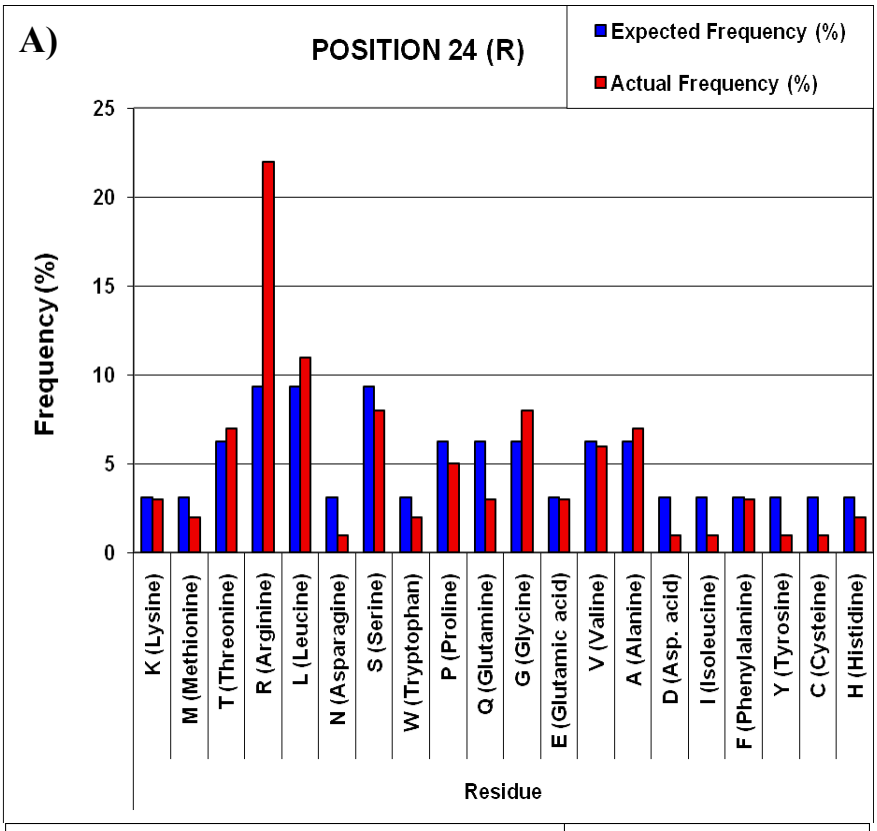
Figure 12: Aligned protein sequences of 100 randomly picked clones from the fully randomized synthetic V_L phage display library. The "WT" human V_L-24 template sequence is shown and dashes are included in CDR3 for sequence alignment. Dots in the sequence entries indicate amino acid identity with the human V_L-24 template sequence on which mutagenesis was performed. The Kabat numbering system is used (Kabat et al., 1991).

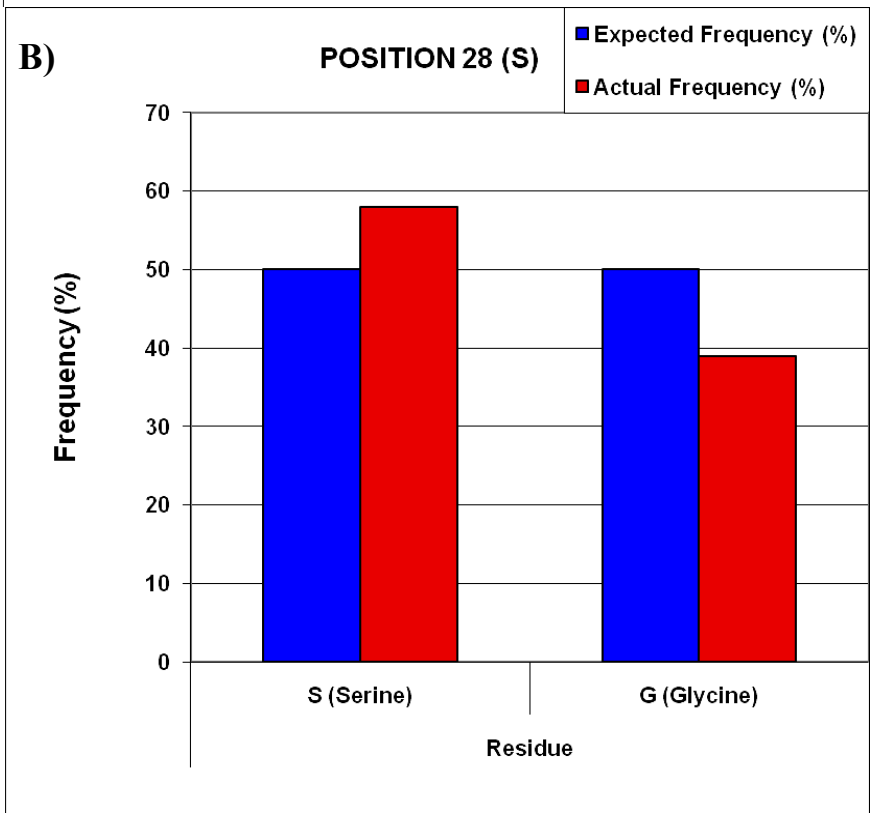
Clone #	CDR1	CDR2	CDR3
	24	50	89
WT	RASQSISTYLN	AASTLQS	QQSYSTP--RT
1(WT)	.....	.....	.....--..
2(WT)	.....	.....	.....--..
3(WT)	.....	.....	.....--..
4	....G.EDQ.S	L..WRS.	G.RS.I.--Q.
5	Y.....TSG.Q	G..L.Y.	..QLVS.--M.
6	L...G.VPD.L	Q..VRW.	..FMRG.--S.
7	....G.PAG..	L..LRG.	A.AT.G.V-W.
8	P.....TVR.L	L...R..	.....--G.
9	S.....RES.I	E..P.S.	R.PADP.H-V.
10	E...G.FRS.D	...G.S.	.....--Y.
11	V...G.LDS.P	P.....	P.ASKF.--L.
12	A......RS.M	S...RL.	D.TTGV.--G.
13	L...G.VMP.A	D..IRF.	L.PPMP.T-S.
14	S.....APR..	R..RRH.	H.PAPH.I-..
15	......PQ..S	T..M.I.	V.AARP.RPA.
16	......T...V	L..P.A.	V.LPYF.--T.
17	....G.F...L	F..NRG.	S.GMHR.KQL.
18	T.....DAT.V	V..DRK.	M.ATLG.--S.
19	W.....I.K.R	G..GRR.	G.GRV..SGL.
20	G.....FIW.T	...A...	..ARPL.--L.
21	......PG..V	P..S.C.	A.NLRP.EGH.
22	L...G.GIR..	S..R.R.	G.LNTF.--T.
23	T......RF..	M..PRP.	A.MAYL.--E.
24	S...G.RNL.W	Y..C.G.	S.RSFA.--S.
25	T......R..G	R..W.A.	A.TR...P-..
26	....G.YFM.T	V..YRT.	L.G.RM.--D.
27	L.....YF..P	T..Y.S.	.......R-S.
28	G.....AGT.I	...G.R.	G.IDQH.F-G.
29	E.....KAV.L	G..P.K.	A.RSWL.P-S.
30	A.....RP..R	.....	G.TKIS.YKM.
31	.......RD.L	S..PRS.	P.VHW..A-A.
32	G.....QPN.I	T..KRM.	R.RRPS.--D.
33	....G.KME.D	I..PRV.	R.RP...L-G.
34	T.....MSI.K	P..PRP.	P.LSVL.R-..
35	....G.V.A..	M..RRA.	.....--..
36	......LP...	...G...	I.KLHN.--..
37	L.....PRP.I	P..RRT.	S......--..
38	H.....QKP.R	L..S.P.	S.VPTP.--L.
39	T...G.D.T..	R..PRI.	L.QPQP.A-G.
40	G.....W.L.P	...PRV.	T..TPN.KGM.
41	F...G.PDP.C	G..V.R.	T.ATRN.KGM.
42	W...G.PW..S	V..LRA.	.....--..
43	L.....LNR.K	P..HRK.	A.FPKY.--M.
44	K......L...	S..G...	R.GS.S.S-P.
45	E.....RSI.G	V..L.P.	P.LIA..--S.
46	P...G.TG..I	R..A.L.	S.TRRQ.--V.
47	A.....GKI..	N....A.	L.MGMN.--V.
48	S...G.PPF.F	S...RC.	.....--..
49	K.....WVH.K	S....D.	E.GVAG.--R.
50	K.....VAV.T	N..I.G.	E.HKII.--V.
51	Q......LW..	E..R.L.	P.PPAP.A-L.
52	....G..LD.L	V..G.N.	A.RGPG.G-T.

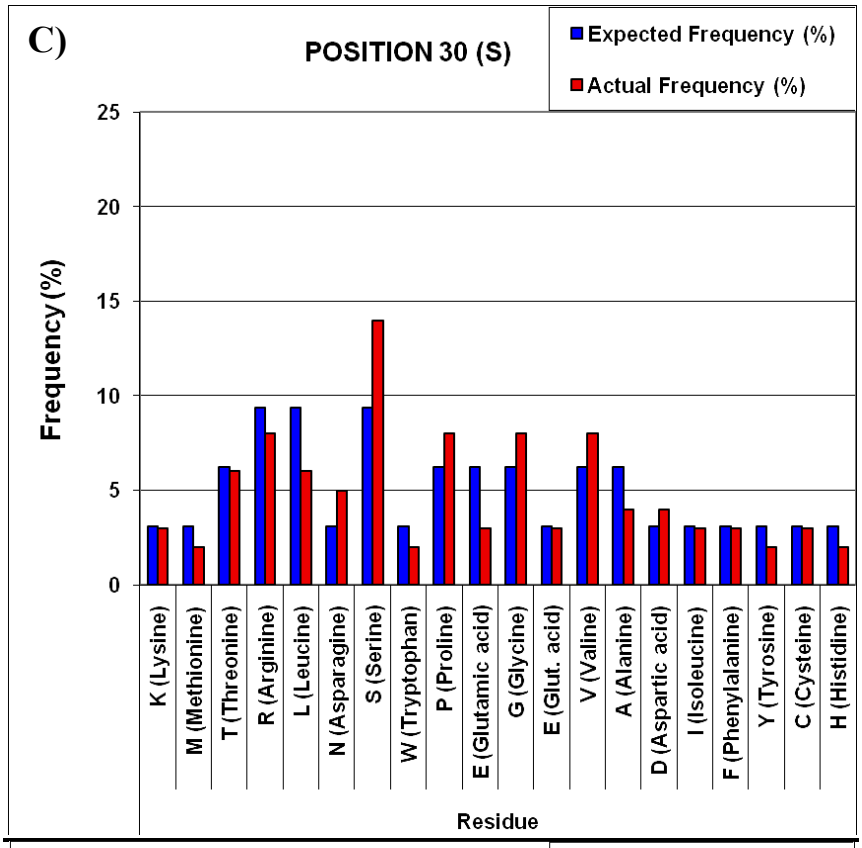
53	....G.GNL.L	T..L.T.	T.AEIE.--..
54	G.....RSE.P	S...RE.	L.QEAL.LQ..
55	G...G.NRK.C	S..P.R.	F.TGWH.--..
56	V...G.GV...	L....L.	C.GRCL.--..
57	P...G.ELM.S	P..PRS.	.......--..
58	P...G.TCR.K	N..F.L.	G.PSRV.PN..
59	S.....CQA.S	P..SRL.	A.GREP.REP.
60	G...G..RQ.W	...V...	..VHAI.KY..
61	V.....GSR.V	Y..D.T.	.......--..
62	V.....RSG.M	F..RRS.	S.AGGG.PP..
63	A...G.LVS.H	V....I.	K......--..
64	L...G.N.D..	P...P.A	.......--..
65	V...G.RLE.G	L..PRA.	.......--..
66	....G.NHS..	M..L...	C.PPAP.--..
67	F.....HIP.R	P..Q.S.	N......--..
68	......MYA.L	.....Y.	A..R.A.G-P.
69	V......AT.Q	T...RC.	D......--G.
70	S.....NSL.E	V....R.	S.LQLR.--..
71	W......LG.R	S......	R..RVG.QLA.
72	S.....DVL.S	T..HRL.	A.HAPP.S-Q.
73	S.....TCS.C	I..SR..	.......--D.
74	L......HF..	...L...	R.PPVP.R-E.
75	Q...G.KEN.E	P..NRV.	V.MHVY.--S.
76	G......WR..	P..WRT.	L...RQ.LGD.
77	A...G.VA...	T..PRL.	W......--..
78	L.....LPD.M	T..HRP.	Y.PGGH.--F.
79	......I.V.L	N..PRW.	R.HRAS.A-K.
80	A...G.DRP.L	L..I.F.	T.AAPA.--P.
81	H.....GHA.K	W..SRP.	.......--..
82	....G.CFR.I	P..H.R.	Y.WSRS.--L.
83	P...G.CVF.G	.....R.	.......--..
84	I.....VYL.K	....	H.THVV.--V.
85	D...G.H.S.R	R..PRP.	R.VS.G.--P.
86	.......RA.R	P..Q.P.	.......--..
87	L...G.N.D.W	P..VRD.	..NQFL.L-..
88	M...G.GKV.M	C..K.N.	L.ARL..T-E.
89	M.....IG..G	L..FRM.	M......--..
90	L...G.TQN.F	E..ERC.	A.PSC..P-..
91	......GWG.S	...G...	A.VPPH.P-..
92	Q.....VFL.Y	P..QRP.	F......--T.
93	T.....P.T..	R..ARL.	T.RAYH.G-V.
94	....G..W..A	K..YRG.	K.LPWY.S-A.
95	T.....AAQ.R	P..A.A.	.......--K.
96	A.....LSR.H	....R..	A.RA.A.--P.
97	F.....VAS.V	Y..RRL.	A.VPPA.F-..
98	N.....ENE.V	S..PRL.	L..PRR.GV..
99	C.....RVK.E	G..R.R.	V.PAGE.--P.
100	....G.ASA.F	R..F.N.	V.KVVS.AAV.

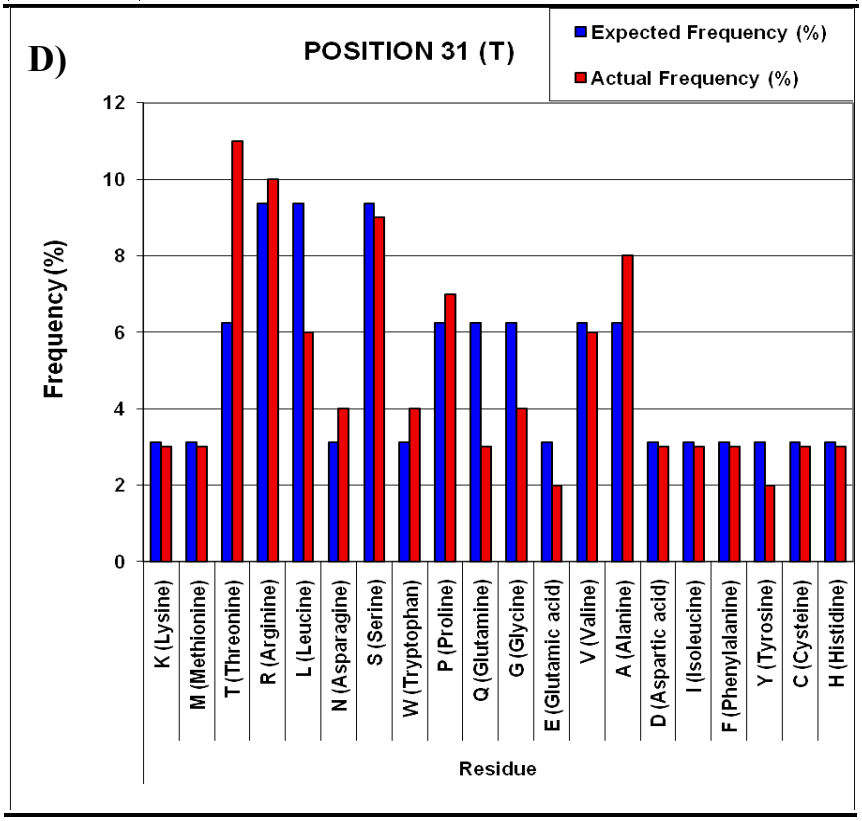
Actual (observed) frequencies were calculated for each randomized position (24, 28, 30, 32, 34, 50, 53, 54, 55, 89, 91, 92, 93, 94, 95a/b, 96) of the 97 randomly selected library V_Ls (the 3 WT clones were excluded from the analysis) and graphed alongside expected frequencies for comparison. Expected frequencies were calculated according to each randomized position (i.e., position 24 = "NNK"; position 28 = "RGC", etc.). Please refer to **6. Identification and sequencing of V_L library clones** of the "Materials & Methods" section for a sample calculation. For expected and actual (observed) frequency values, please refer to Table S1 ("Supplemental" section). Results are shown in Figure 13, graphs A-R.

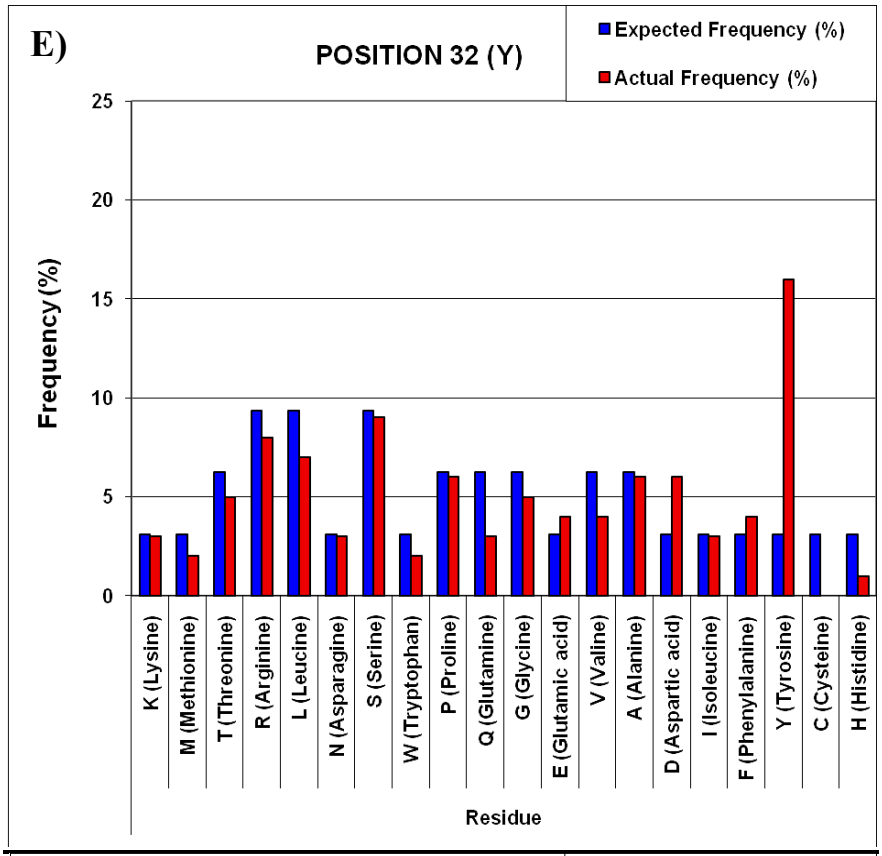
Figure 13: Comparison of the expected and actual amino acid frequencies for 100 randomly selected synthetic V_L library clones. 100 clones were selected at random from the constructed synthetic V_L library and sequenced. The expected (in blue) and actual (observed, in red) amino acid frequencies were then computed and graphed for each randomized position (A-R for positions 24, 28, 30, 32, 34, 50, 53, 54, 55, 89, 91, 92, 93, 94, 95a, 95b, and 96, respectively). Wild-type residues at each position are indicated in brackets for each graph.

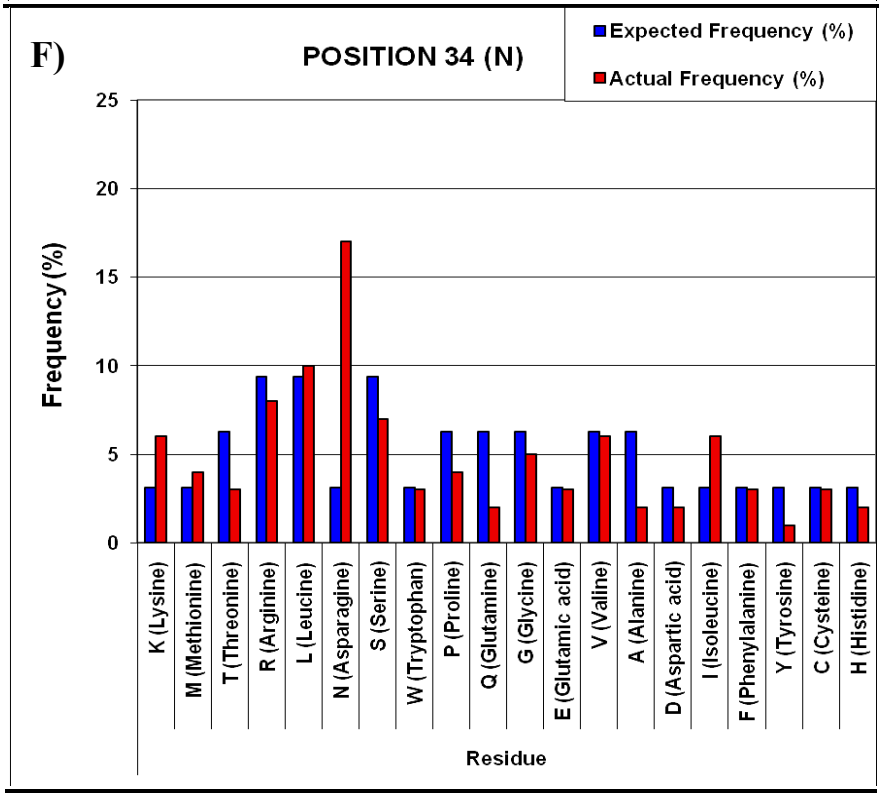

A)
POSITION 24 (R)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Asp. acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

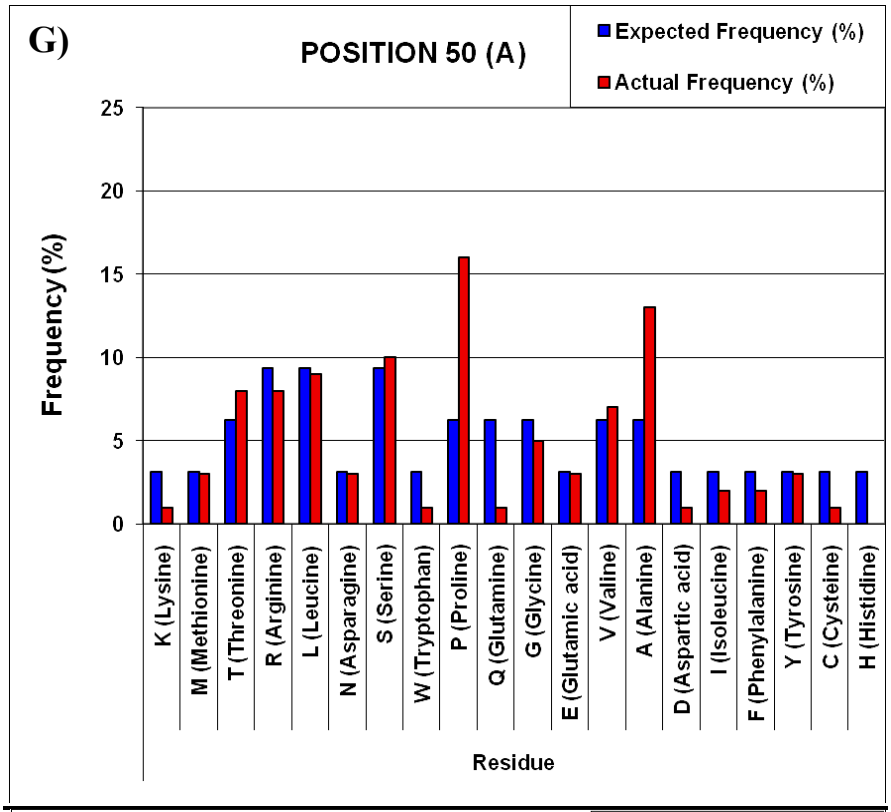

B)
POSITION 28 (S)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
70
60
50
40
30
20
10
0
S (Serine)
G (Glycine)
Residue

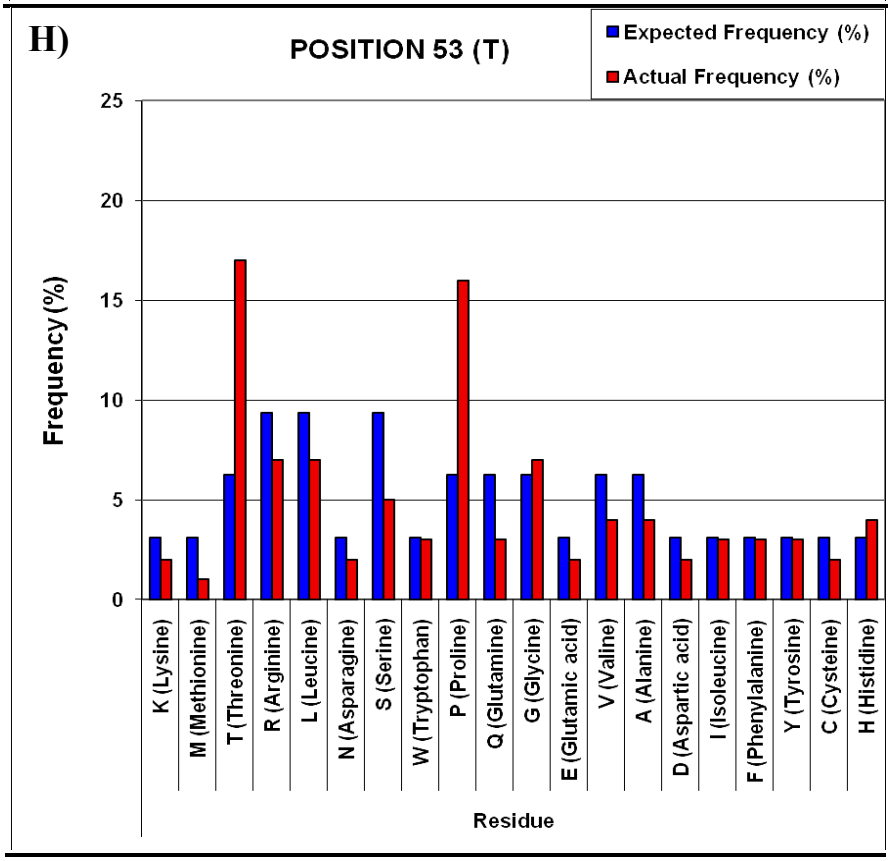

C)
POSITION 30 (S)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
E (Glutamic acid)
G (Glycine)
E (Glut. acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

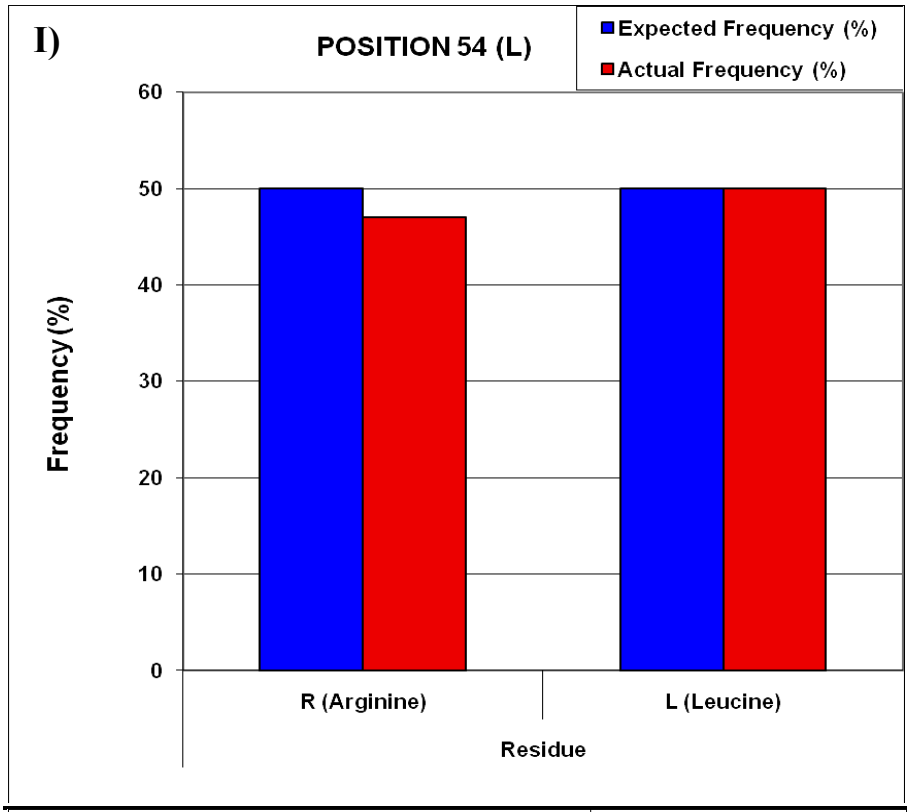

D)
POSITION 31 (T)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
12
10
8
6
4
2
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

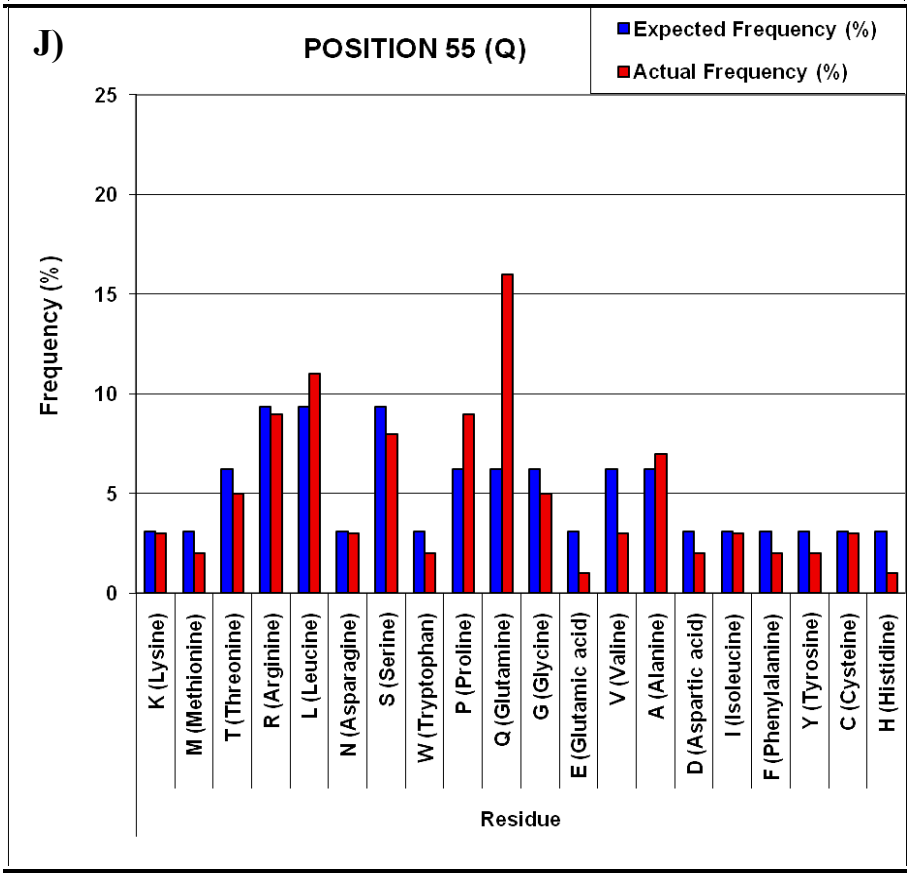

E)
POSITION 32 (Y)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

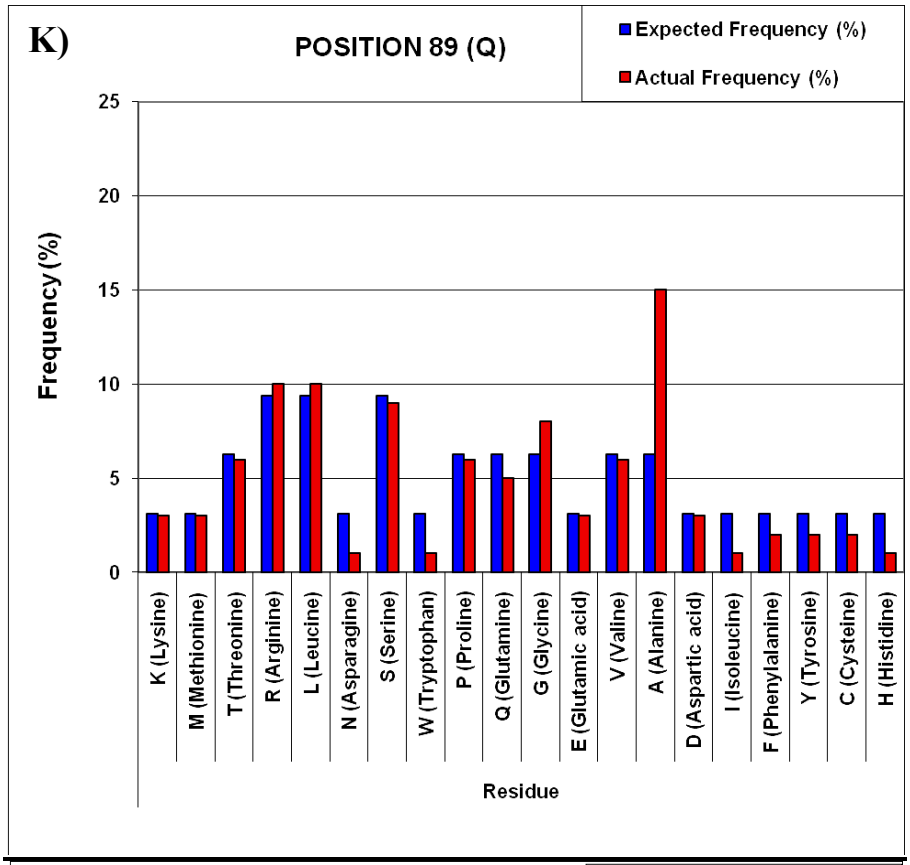

F)
POSITION 34 (N)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

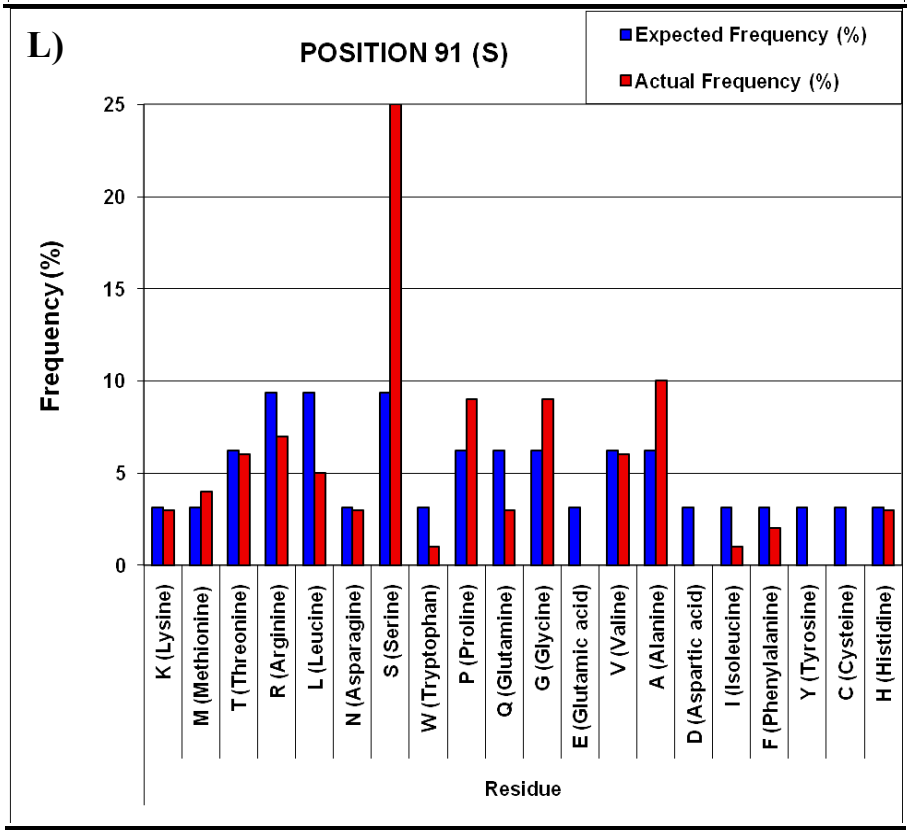

G)
POSITION 50 (A)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

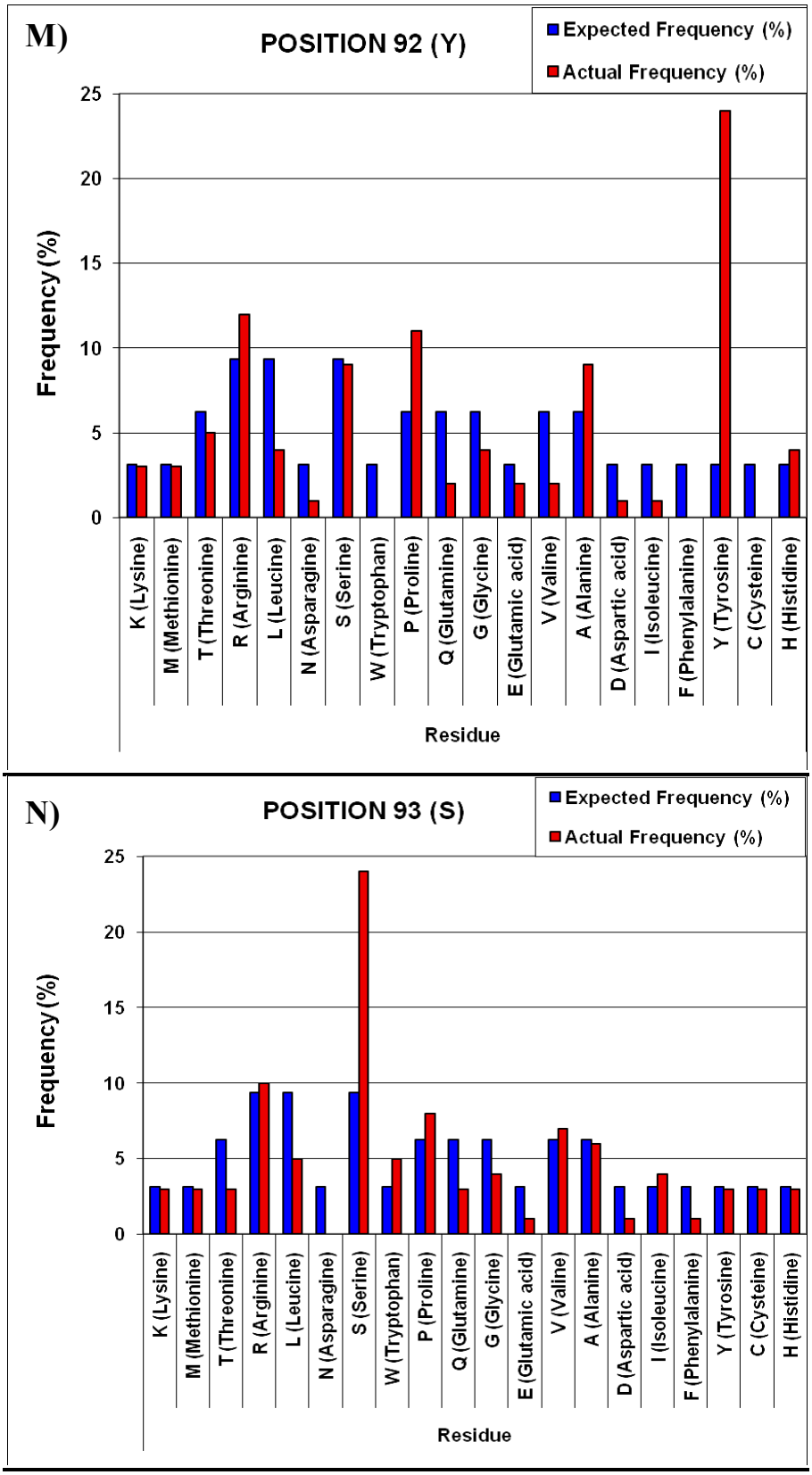

H)
POSITION 53 (T)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

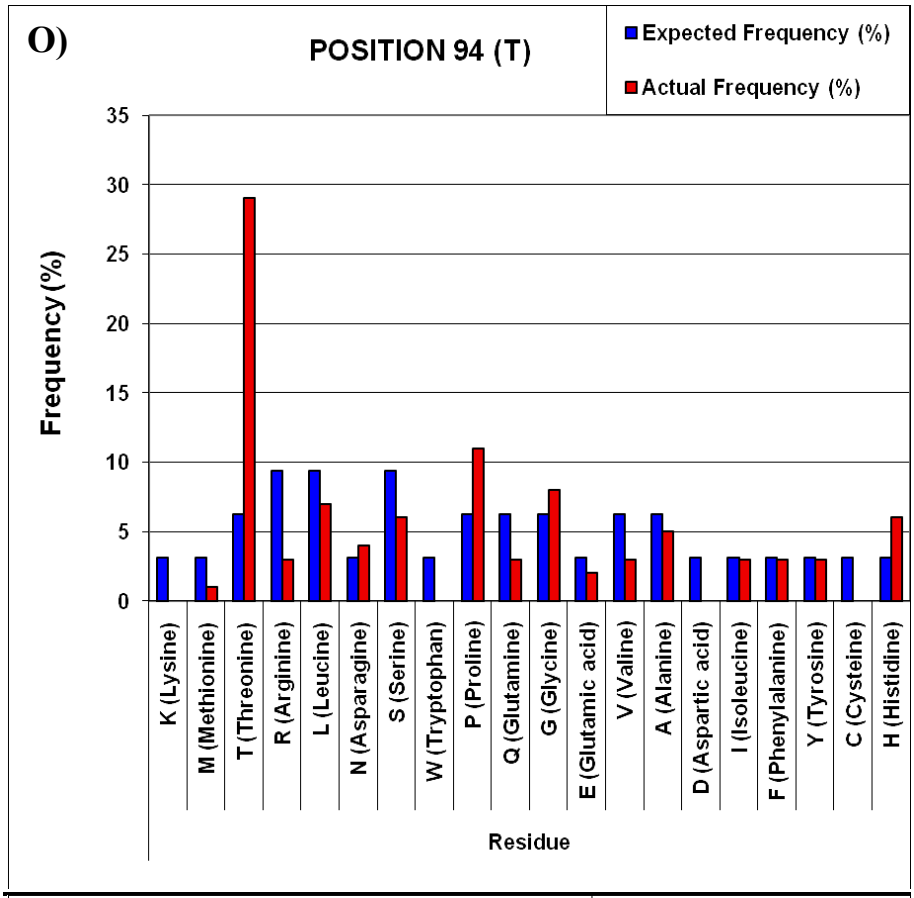

I)
POSITION 54 (L)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
60
50
40
30
20
10
0
R (Arginine)
L (Leucine)
Residue

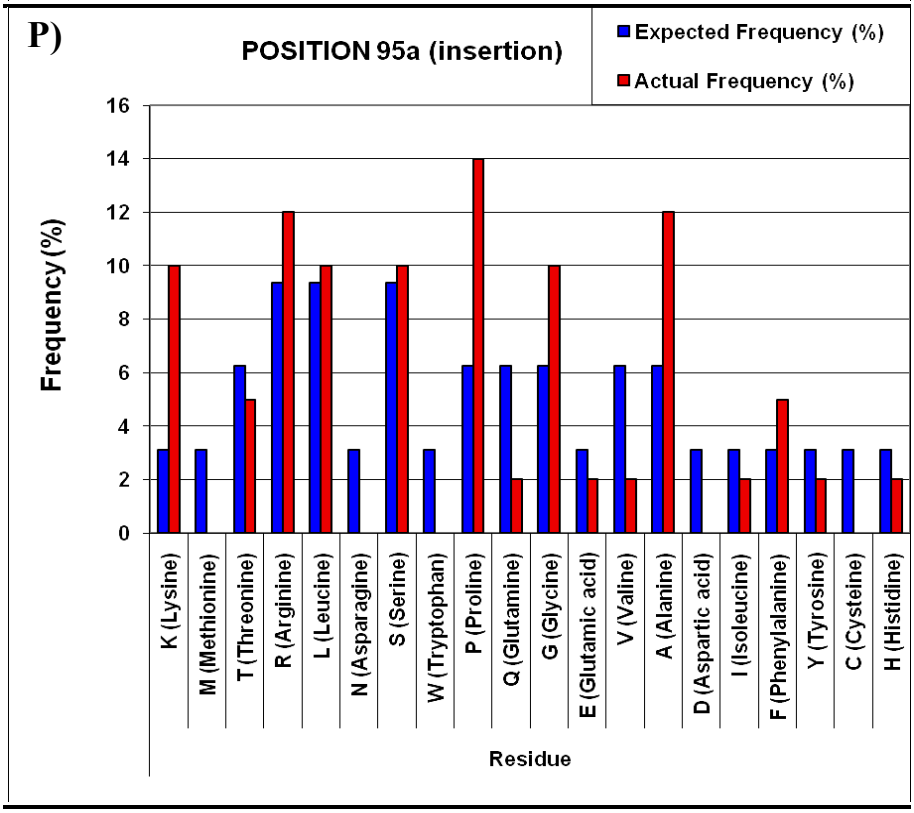

J)
POSITION 55 (Q)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

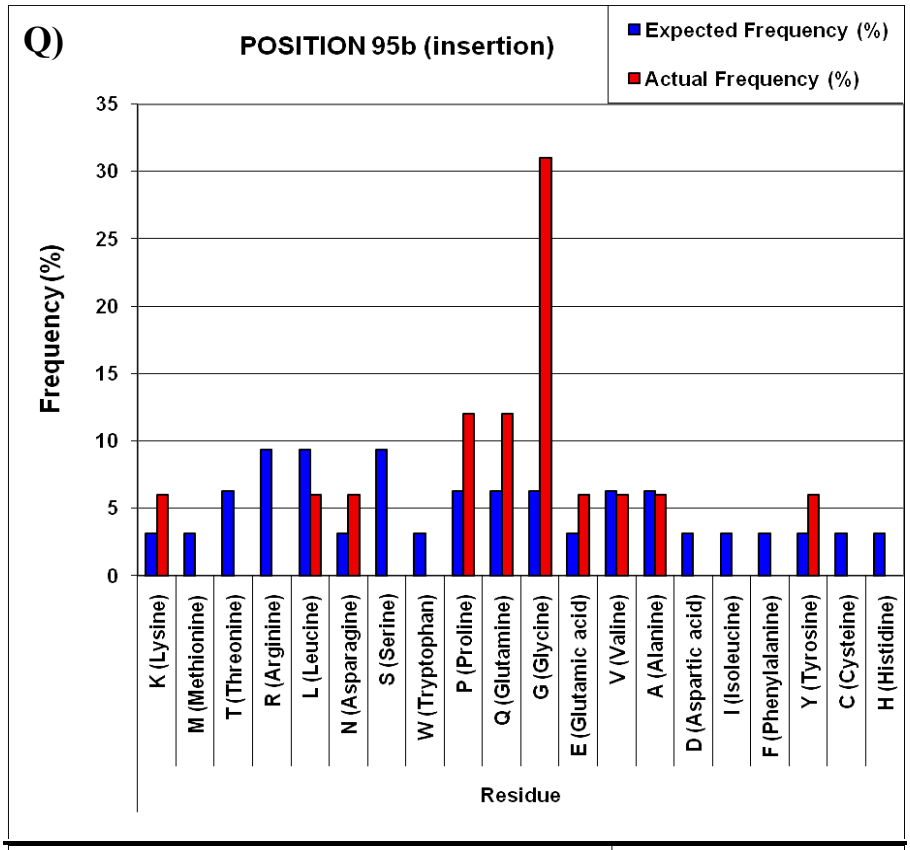

K)
POSITION 89 (Q)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

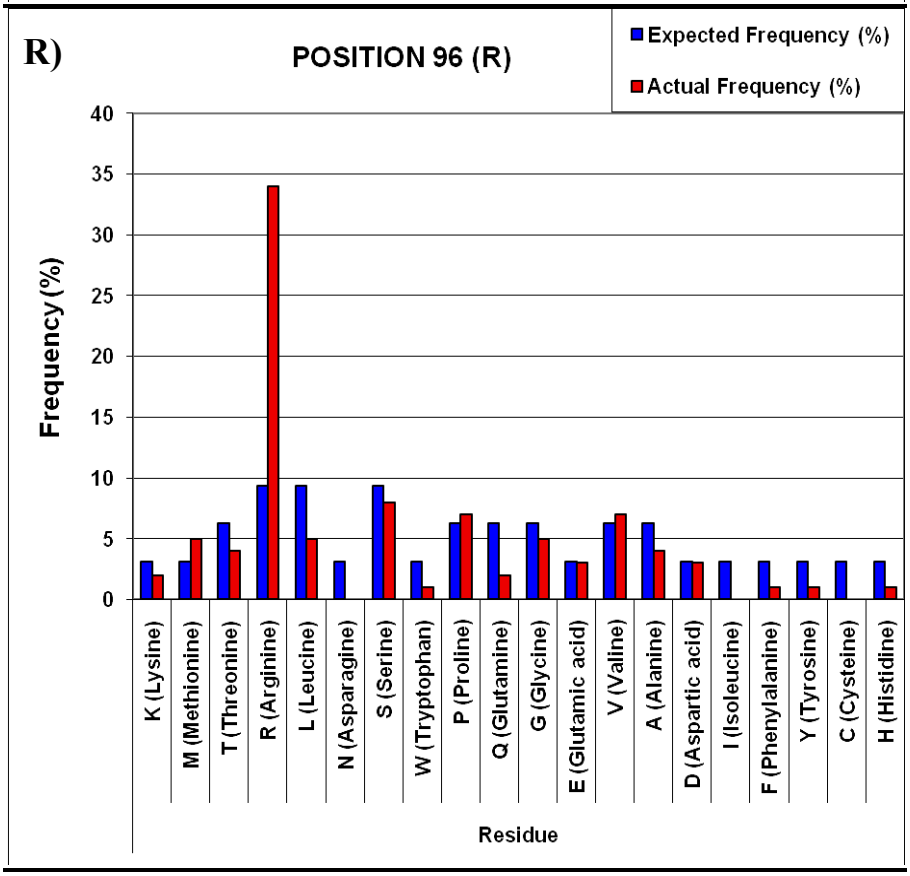

L)
POSITION 91 (S)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue


M)
POSITION 92 (Y)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue
N)
POSITION 93 (S)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue


O)
POSITION 94 (T)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
35
30
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue


P)
POSITION 95a (insertion)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
16
14
12
10
8
6
4
2
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue


Q)
POSITION 95b (insertion)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
35
30
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue


R)
POSITION 96 (R)
Expected Frequency (%)
Actual Frequency (%)
Frequency (%)
40
35
30
25
20
15
10
5
0
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
Residue

While sequence analysis confirmed the degree of heterogeneity of the constructed V_L library, its results also suggested a number of cases in which there were clear differences between the observed and expected frequencies for a particular amino acid at a specific randomized position.

The diversity of constructed Ab phage libraries is often influenced by biological censoring during the phage amplification process, effectively lowering the likelihood of selecting high affinity clones from the library. Biological censoring is a phenomenon that commonly occurs in phage libraries when their diversity becomes limited by a disproportionate representation of different phage clones during phage reproduction such that residues compromising phage stability or replication are suppressed within a library, while those such as "wild-type"/template residues that may be advantageous to phage, are favoured and propagated within the library (Kuzmicheva et al., 2009). Statistical analysis of the 100 V_L sequences was carried out using the Frequency Disparity Index (FDI) equation for each of the residues at each randomized position in the V_L insert. The FDI formula is an adapted equation (Kuzmicheva et al., 2009) derived from Rosner (2006), and provides an indication of how much the expected (denoted as "E") frequency deviates from the observed (denoted as "O") frequency for the amino acid at a given randomized position within the V_L insert. "N" represents the total number of V_L clones used in sequence analysis (in this case, N = 100).

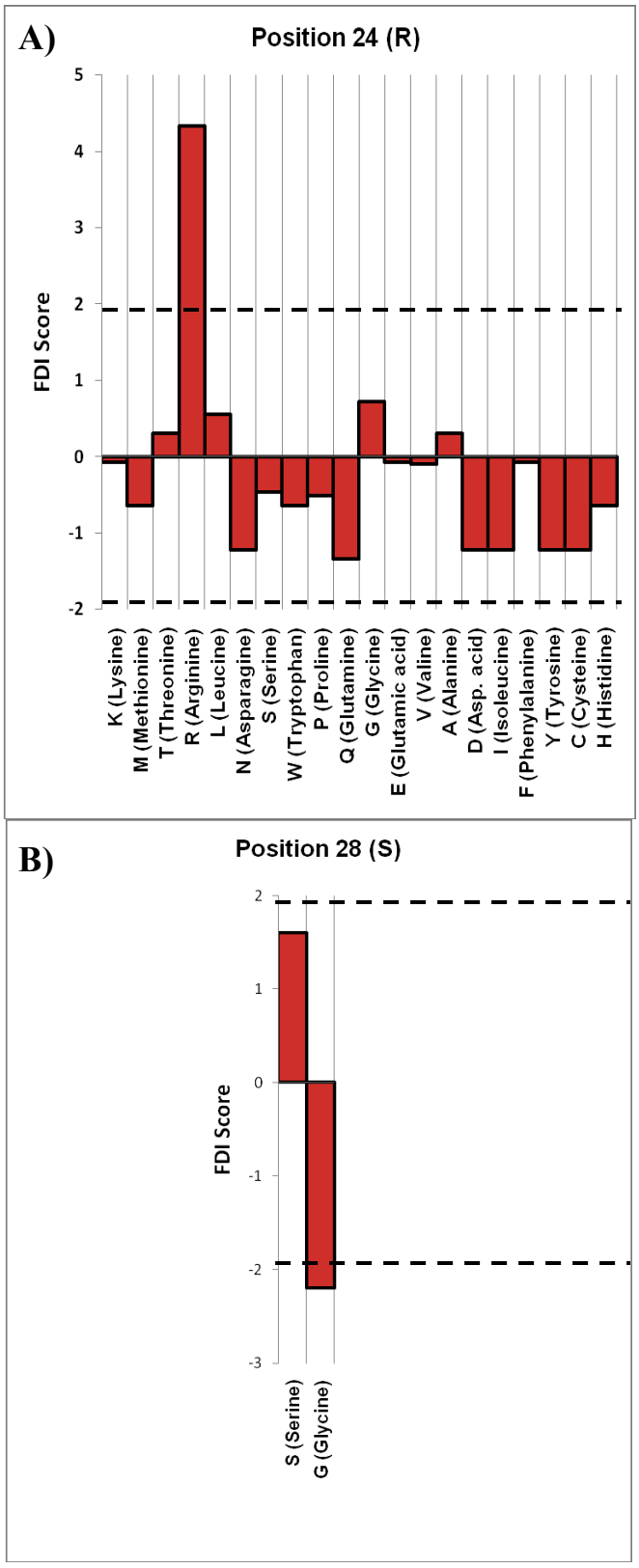
$$\mathbf{FDI} = \frac{O - E}{\sqrt{(E(1 - E)/N)}}$$

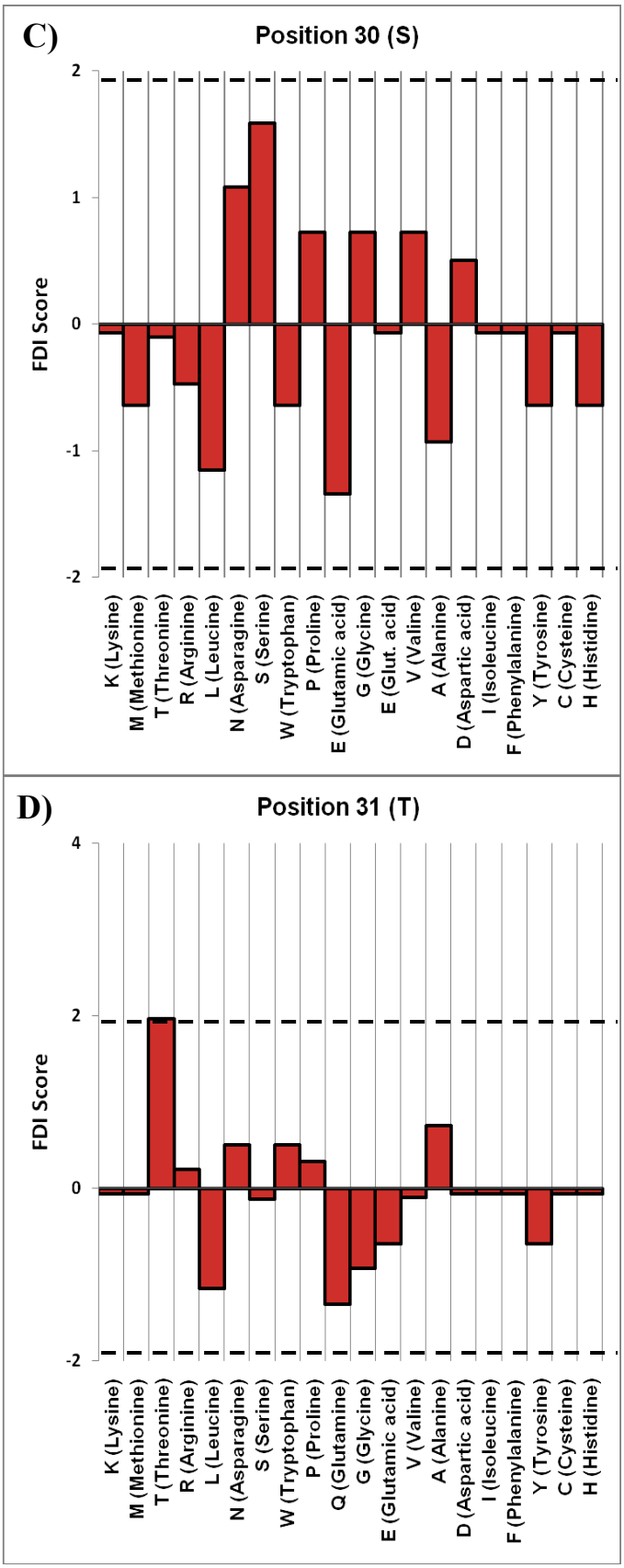
The FDI values calculated may be positive, negative, or zero, depending on whether the particular residue was reported at a higher frequency than expected (in which case the FDI value would be positive), at a lower frequency than expected (negative FDI value), or at the expected

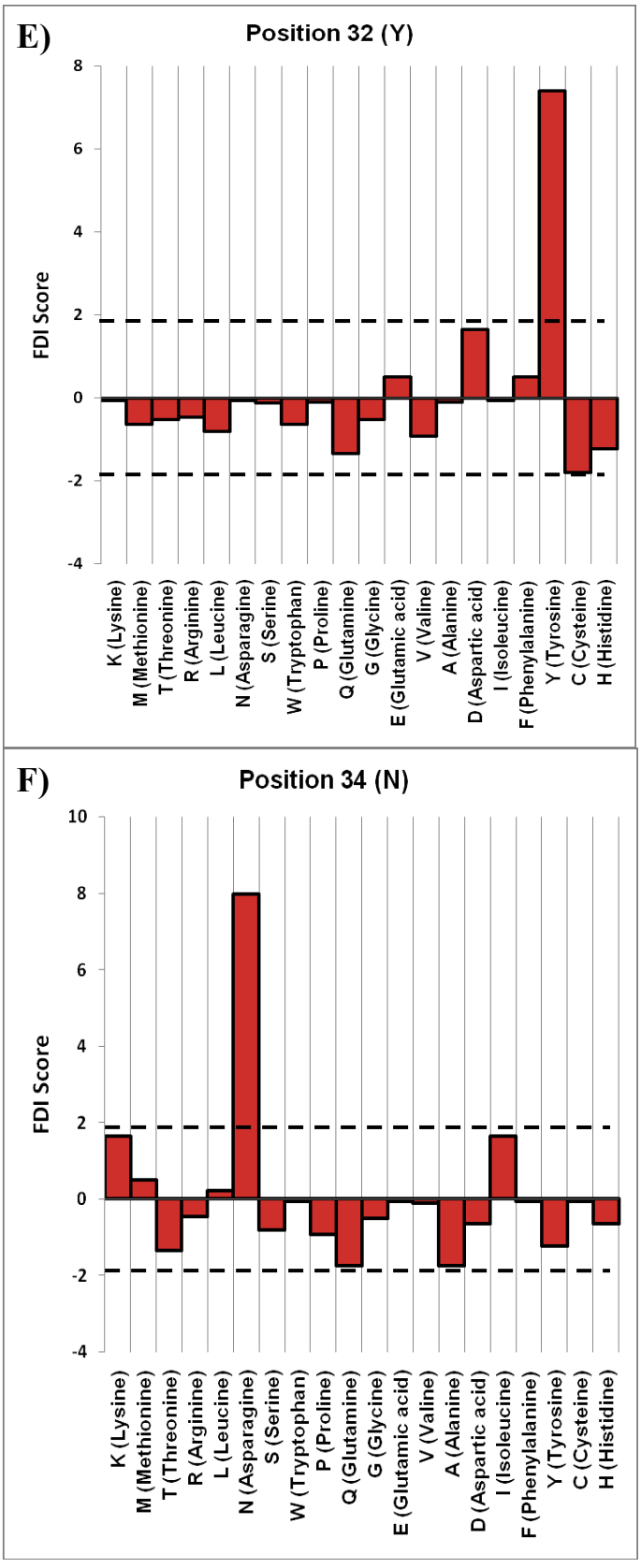
level (FDI of zero). Over- or under-representation of different amino acids corresponding to the randomized positions of the V_L insert is best visualized by graphical means (Figure 14, A-R). Since a relatively large number of clones (i.e., 100) was analyzed, the FDI values should conform to a standard distribution in which two-tailed P-values derived from a Z-table can be used to demonstrate the statistical significance of any differences between the expected and actual (observed) frequencies of an amino acid at a given position. In statistical significance testing a P-value is the probability of observing a sample statistic as extreme as the theoretical value (Schervish, 1996). Since the test statistic is a Z-score, a normal (symmetrical) distribution is used to assess the probability associated with that Z-score. The null hypothesis is traditionally rejected (i.e., the observed value is statistically significant) when the P-value is either less than/equal to or greater than/equal to the significance level (i.e., threshold value) represented by the Greek letter α, which is often 0.05 (Gosling, 1995). The 0.05 level of significance corresponds to a 95% confidence interval with Z-scores of +/-1.96 that represent cutoff points (found in a standard Z-table) (Gosling, 1995).

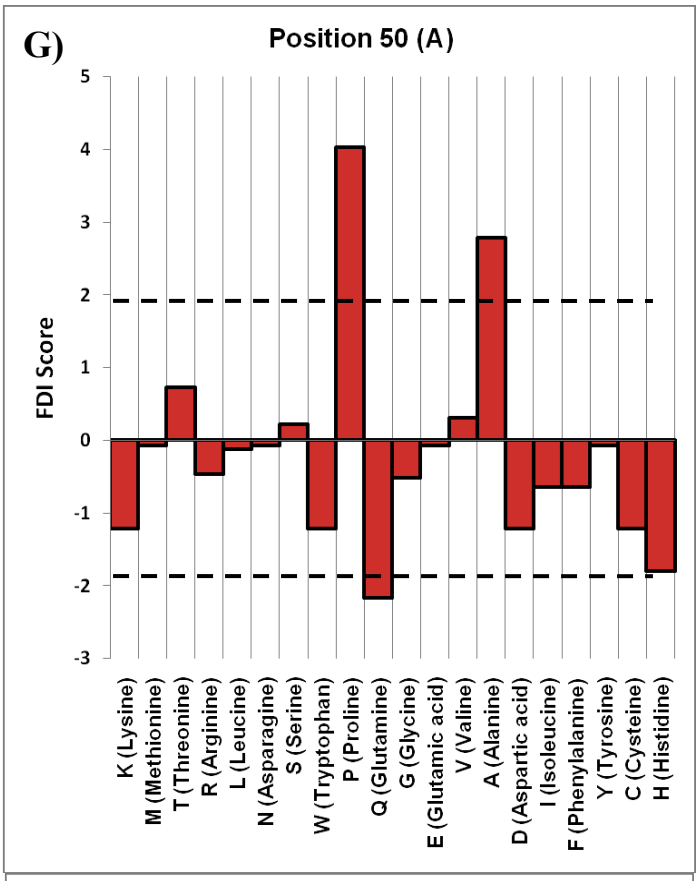
It was found that the observed frequency of amino acids at different randomized positions did not strictly match the theoretical frequency values. At a α-level of 0.05, FDI scores greater than or equal to 1.96 represented significant deviation of the observed frequency from the expected value (with over- or under-representation of the residue depending on the sign of the FDI score). It should be noted that any significant over- or under-representations observed are exclusive events for each residue at each randomized V_L position; each residue reported at each individual position is considered an isolated occurrence, meaning no correlations can be made between adjacent positions or to other residues. Please refer to Table S1 ("Supplemental" section) for calculated FDI scores. Results are shown in Figure 14, graphs A-R.

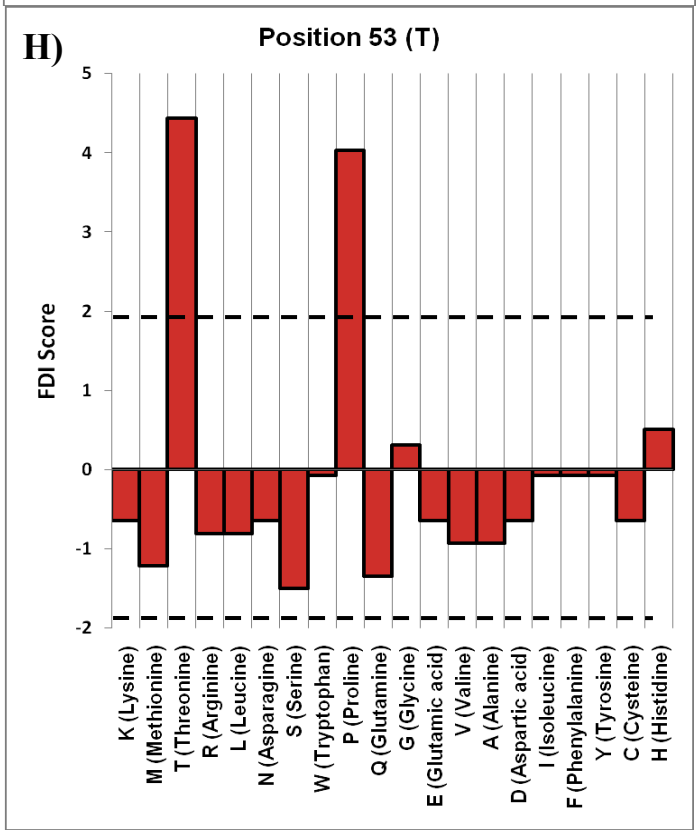
Figure 14: Statistical analysis of amino acids at randomized positions in the 100 randomly selected V_L library clones. The expected (E) frequency of amino acids and the actual (observed) frequency calculated were used in an adaptation of the one-sample tests for proportion (Rosner, 2006). The results of this analysis were expressed as a frequency disparity index (FDI) which is a measure of how much the actual frequency deviates from the expected. Amino acids for each randomized position (A-R for positions 24, 28, 30, 32, 34, 50, 53, 54, 55, 89, 91, 92, 93, 94, 95a, 95b, and 96, respectively) of the V_L library clones are displayed along the x-axis whereas the y-axis represents the FDI score. FDIs for each amino acid are presented by clusters of columns, and dashed horizontal lines represent statistical significance of a frequency disparity at the α-level of 0.05. Wild-type residues at each position are indicated in brackets for each graph.

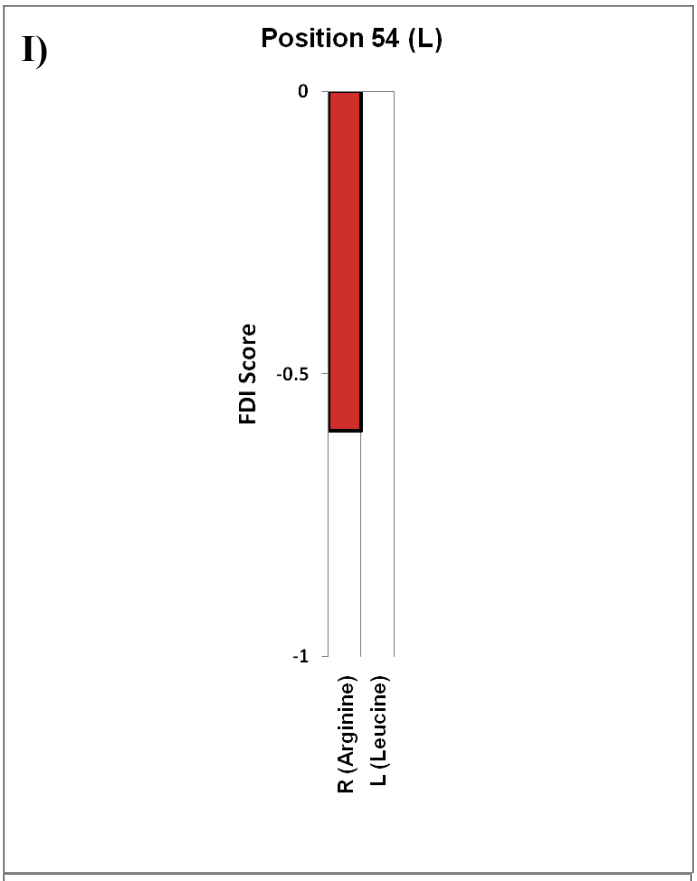

A)
Position 24 (R)
FDI Score
5
4
3
2
1
0
-1
-2
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Asp. acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
B)
Position 28 (S)
FDI Score
2
1
0
-1
-2
-3
S (Serine)
G (Glycine)

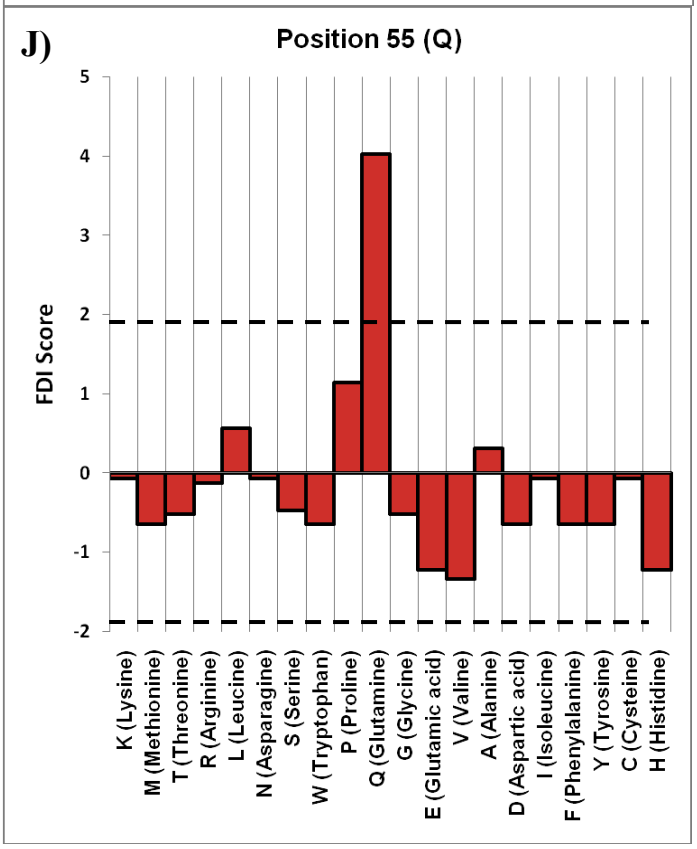

C)
Position 30 (S)
FDI Score
2
1
0
-1
-2
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
E (Glutamic acid)
G (Glycine)
E (Glut. acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
D)
Position 31 (T)
FDI Score
4
2
0
-2
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

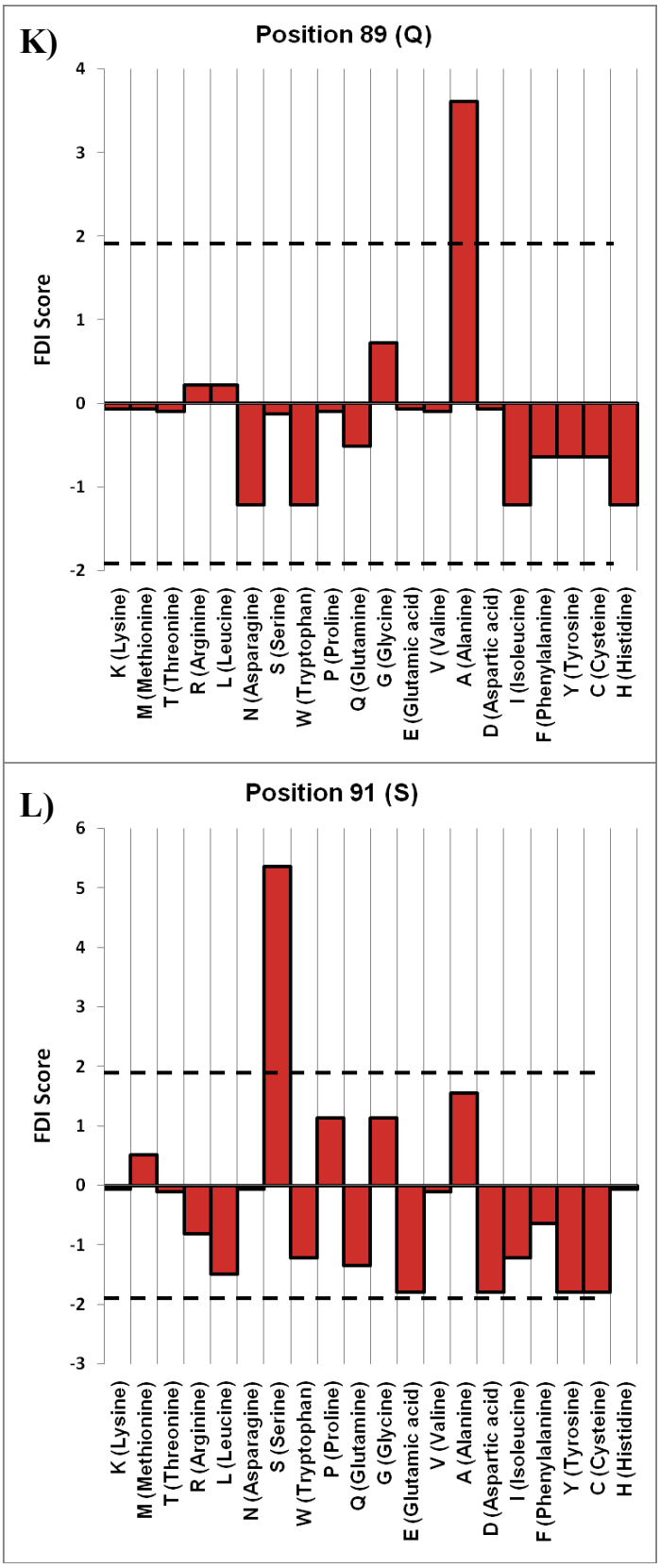

E)
Position 32 (Y)
FDI Score
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
F)
Position 34 (N)
FDI Score
10
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

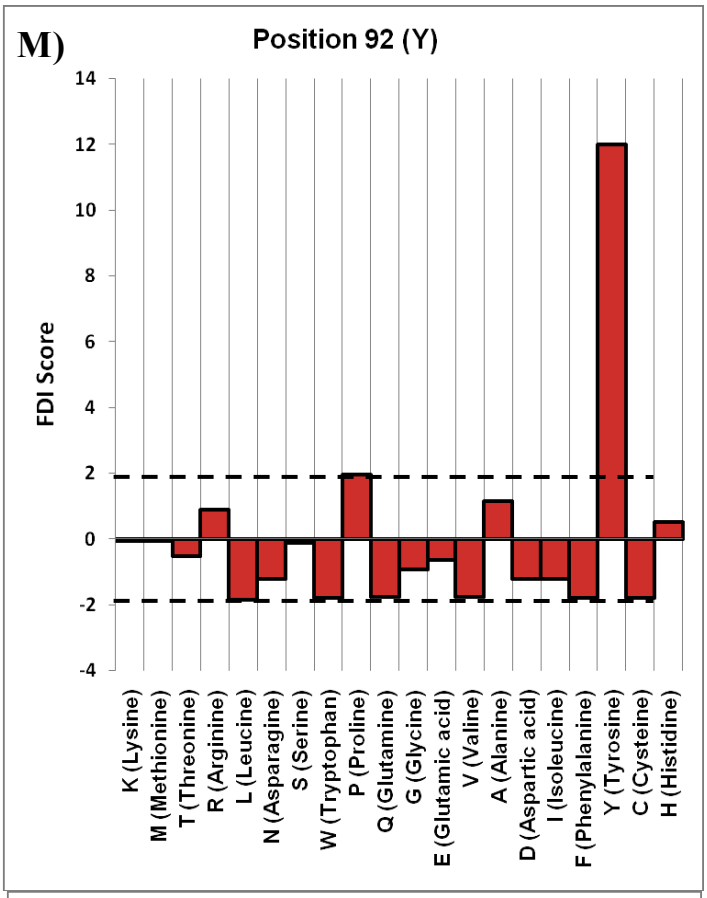

G)
Position 50 (A)
FDI Score
5
4
3
2
1
0
-1
-2
-3
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

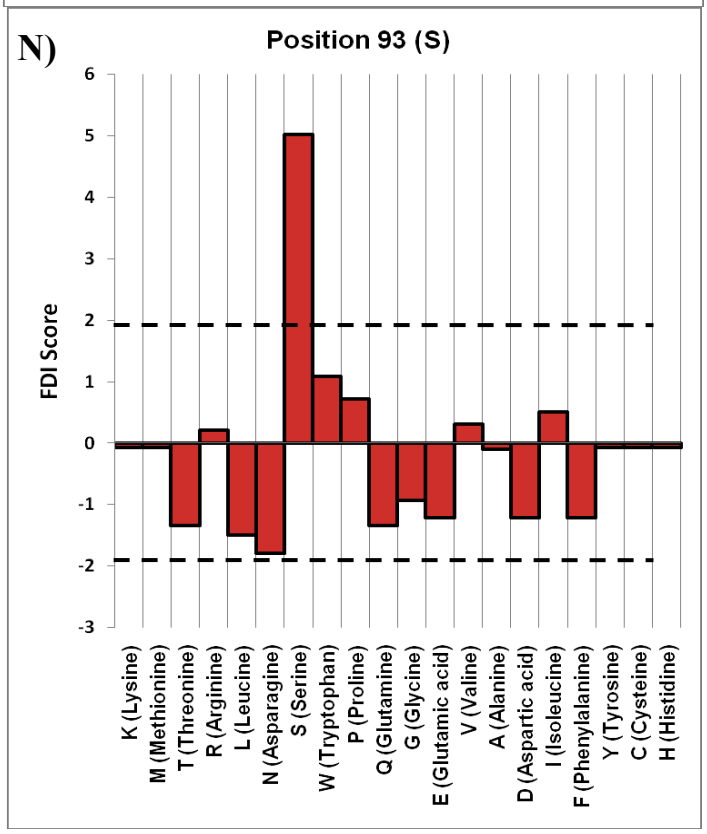

H)
Position 53 (T)
FDI Score
5
4
3
2
1
0
-1
-2
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

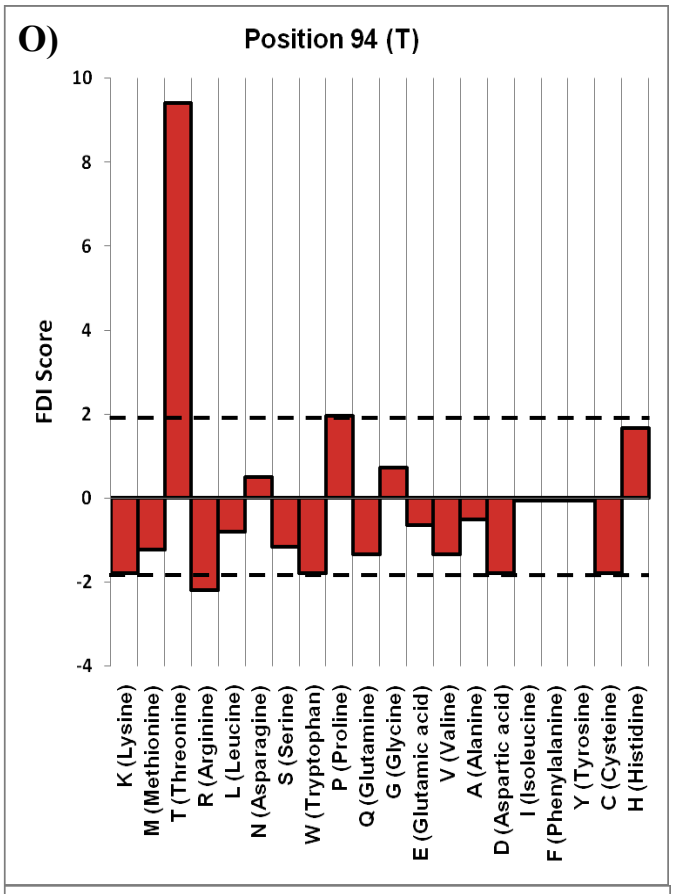

I)
Position 54 (L)
0
-0.5
-1
FDI Score
R (Arginine)
L (Leucine)

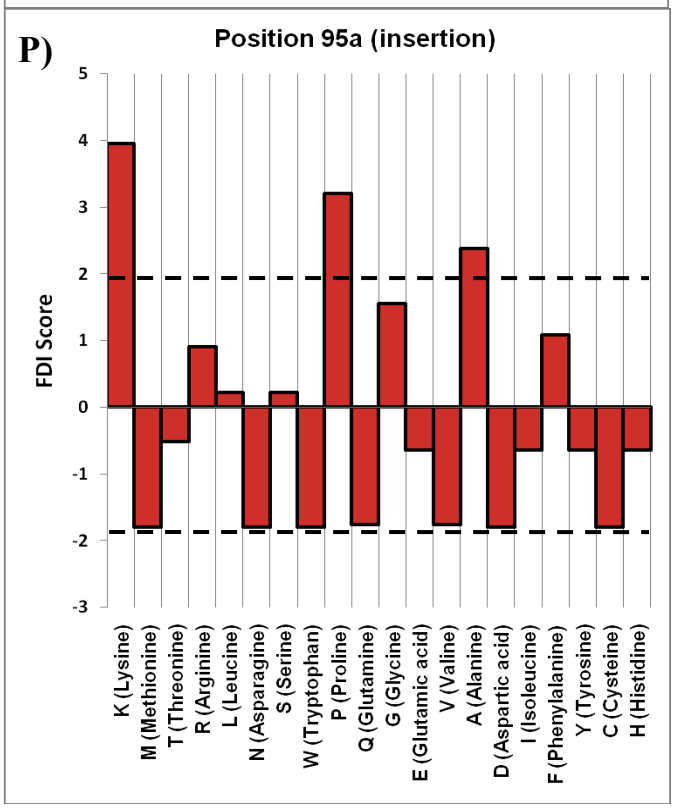

J)
Position 55 (Q)
5
4
3
2
1
0
-1
-2
FDI Score
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

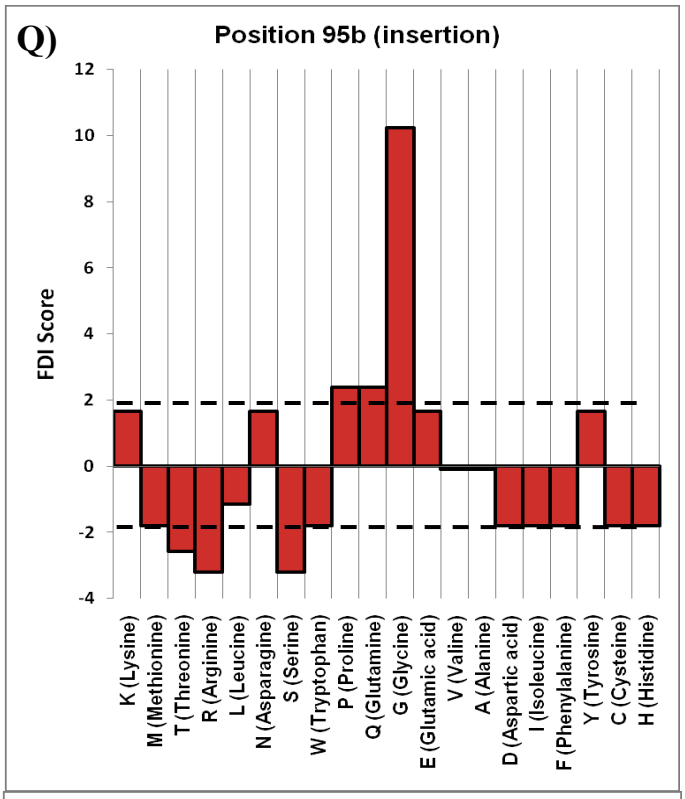

K)
Position 89 (Q)
FDI Score
4
3
2
1
0
-1
-2
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)
L)
Position 91 (S)
FDI Score
6
5
4
3
2
1
0
-1
-2
-3
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

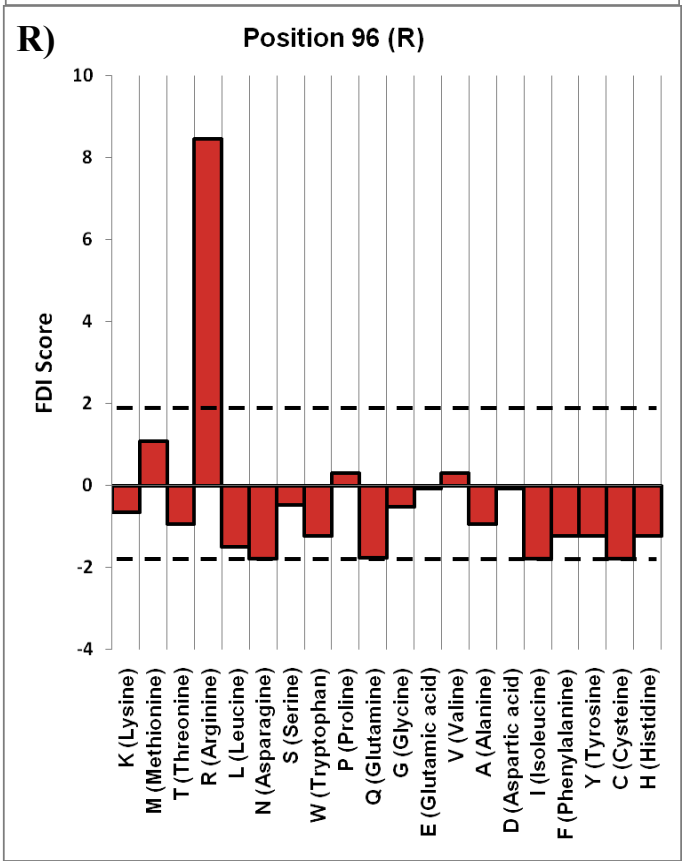

M)
Position 92 (Y)
FDI Score
14
12
10
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)


N)
Position 93 (S)
FDI Score
6
5
4
3
2
1
0
-1
-2
-3
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)


O)
Position 94 (T)
FDI Score
10
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)


P)
Position 95a (insertion)
FDI Score
5
4
3
2
1
0
-1
-2
-3
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)


Q)
Position 95b (insertion)
FDI Score
12
10
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)


R)
Position 96 (R)
FDI Score
10
8
6
4
2
0
-2
-4
K (Lysine)
M (Methionine)
T (Threonine)
R (Arginine)
L (Leucine)
N (Asparagine)
S (Serine)
W (Tryptophan)
P (Proline)
Q (Glutamine)
G (Glycine)
E (Glutamic acid)
V (Valine)
A (Alanine)
D (Aspartic acid)
I (Isoleucine)
F (Phenylalanine)
Y (Tyrosine)
C (Cysteine)
H (Histidine)

A of residues displayed FDI scores greater than 1.96, thus indicating strong enrichment and significant over-representation at the selected positions (Figure 14, graphs A-R). Among these were the predominating "wild-type" amino acids arginine (R) at positions 24 and 96; serine (S) at positions 91 and 93; threonine (T) at positions 31, 53 and 94; tyrosine (Y) at positions 32 and 92; asparagine (N) at position 34; alanine (A) at position 50; and glutamine (Q) at position 55. A considerable preference was observed for proline (P) at two CDR2 positions (50 and 53) as well as several CDR3 positions (92, 94, 95a and 95b). Alanine (A) was also noted to occur frequently at positions 89 and 95a; inserted amino acids occurring at higher frequency than expected included proline (P), lysine (K), and alanine (A) at position 95a, and glycine (G), glutamine (Q), and proline (P) at position 95b.

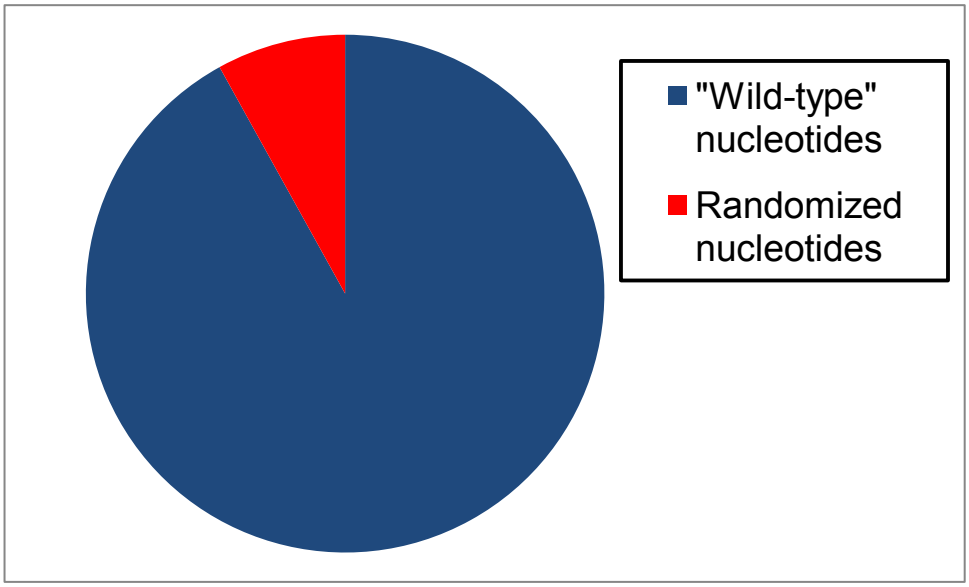
Far fewer cases were observed for different amino acids that displayed FDI scores less than 1.96 (Figure 14, graphs A-R). Residues that appeared at a lower frequency than expected included glycine (G) at position 28; glutamine (Q) at position 50; and serine (S), arginine (R), and threonine (T) at position 95b.

It was determined that the observed frequencies of amino acids at different positions of the randomized regions did not always match the pattern of the expected frequencies. Biological censoring, which occurs during phage amplification of clones, may have played a role in influencing library diversity. However, almost all of the randomized positions displayed a distinct preference for the "wild-type" residue of a specific position; positions 28, 30, and 89 were the only cases that did not follow this trend. This could imply a bias for these residues in the bacterial *E. coli* system or a result of the annealing efficiency, which is affected by a number of factors (addressed under "Discussion"). Among the other major over-representations noted

was the consistent preference toward proline (P) throughout both the CDR2- (positions 50 and 53) and CDR3-randomized regions (positions 92, 94, and 95a/b; Figures 13 and 14, graphs G, H, M, and O-Q). There were only three instances of under-representation and they occurred at positions 28 (glycine (G)), 50 (glutamine (Q)) and 95b (serine (S), arginine (R), and threonine (T)) (Figures 13 and 14, graphs B, G and Q). Comparison of the expected and observed amino acid frequency profiles for the constructed V_L libraries provides some evidence of biological censoring for several positions, as well as certain preferences toward residues that might favour better surviving clones during amplification of the library.

Following sequence analysis on amino acids of the 100 randomly selected V_L library clones, an additional nucleotide analysis was performed on all of the randomized positions throughout the 3 CDRs that had displayed "wild-type" residues (i.e., amino acids identical to those found in the V_L-24 template sequence). Analysis revealed that ~8% of the "wild-type" residues were non "wild-type" in terms of nucleotide composition; in other words, ~92% of these positions had retained their original nucleotide sequences (i.e., had not been randomized) (Figure 15). The same results were observed for each CDR individually, as well. These results may be due to the annealing efficiency of Kunkel's mutagenesis protocol, as well as a number of other factors influencing the efficiency of annealing (further addressed in "Discussion" section).

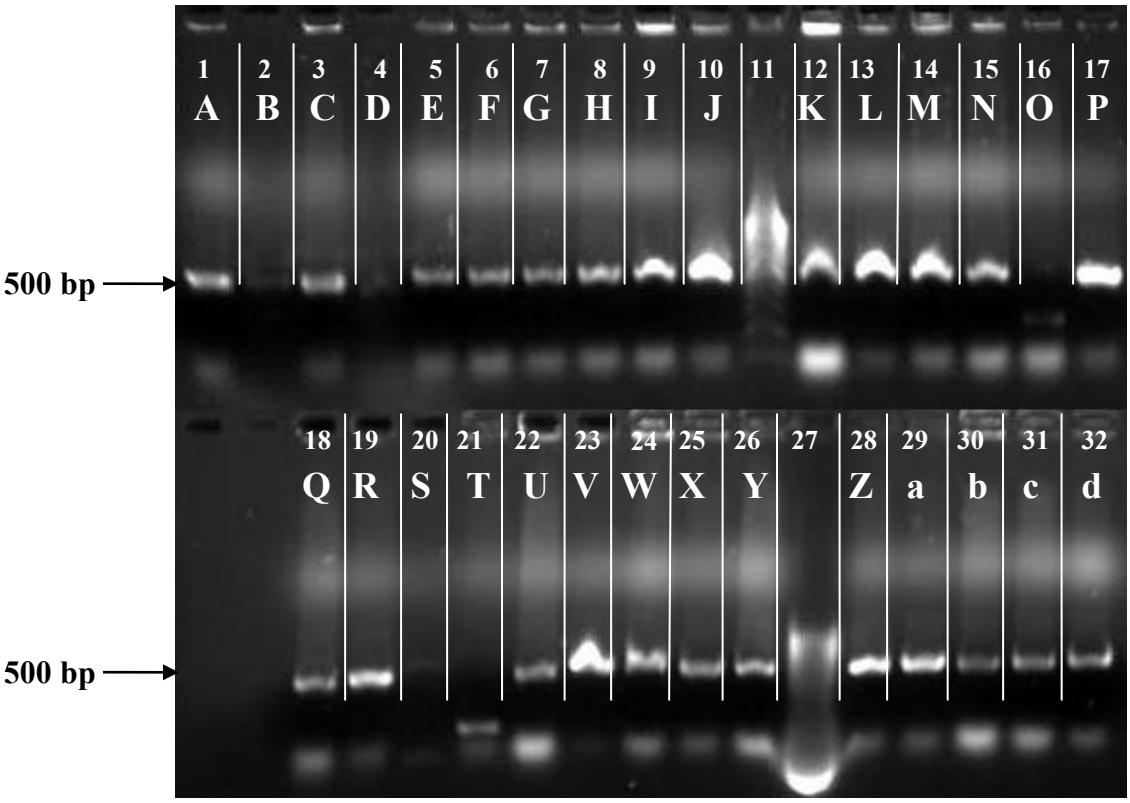
Figure 15: DNA nucleotide analysis of positions that were targeted for randomization and that resulted in amino acids identical to the human V_L-24 template sequence. Nucleotide analysis was performed on all of the randomized positions throughout the CDRs that had displayed residues identical to those found in the V_L-24 template sequence. Results indicated that ~8% of these amino acids consisted of nucleotides that did not match the parent V_L-24 sequence, meaning that ~92% of these positions had retained their original nucleotide sequences (had not been randomized).


"Wild-type" nucleotides
Randomized nucleotides

4. Expression and protein analysis of randomly selected V_L clones from the constructed synthetic library

Thirty of the 100 previously selected library colonies were picked for amplification of their V_L inserts using the Forward and Reverse V_L gene-specific primers "HVL89 BamHI F" and "HVL42 BbsI B" (please refer to "Materials and Methods" section for respective primer sequences). The V_L inserts were then digested with *BamHI* and *BpiI* (*BbsI),* and cloned into the *E. coli* expression vector pSJF2H for expression and protein analysis. Following ligation of the 30 V_L inserts with pSJF2H and transformation into *E. coli* TG1 cells, colony-PCR was performed on 30 randomly picked colonies from the 2xYT/Amp (100 µg/mL) transformation plates, using universal primers "M13FP" and "M13RP" (please refer to **7. Cloning of V_L inserts** of the "Materials & Methods" section for primer sequences). Each of the 30 amplified PCR products (~5 µL each) were loaded onto a 1% agarose gel for verification of insert presence (indicated by appearance of bands at ~500 bp) (Figure 16).

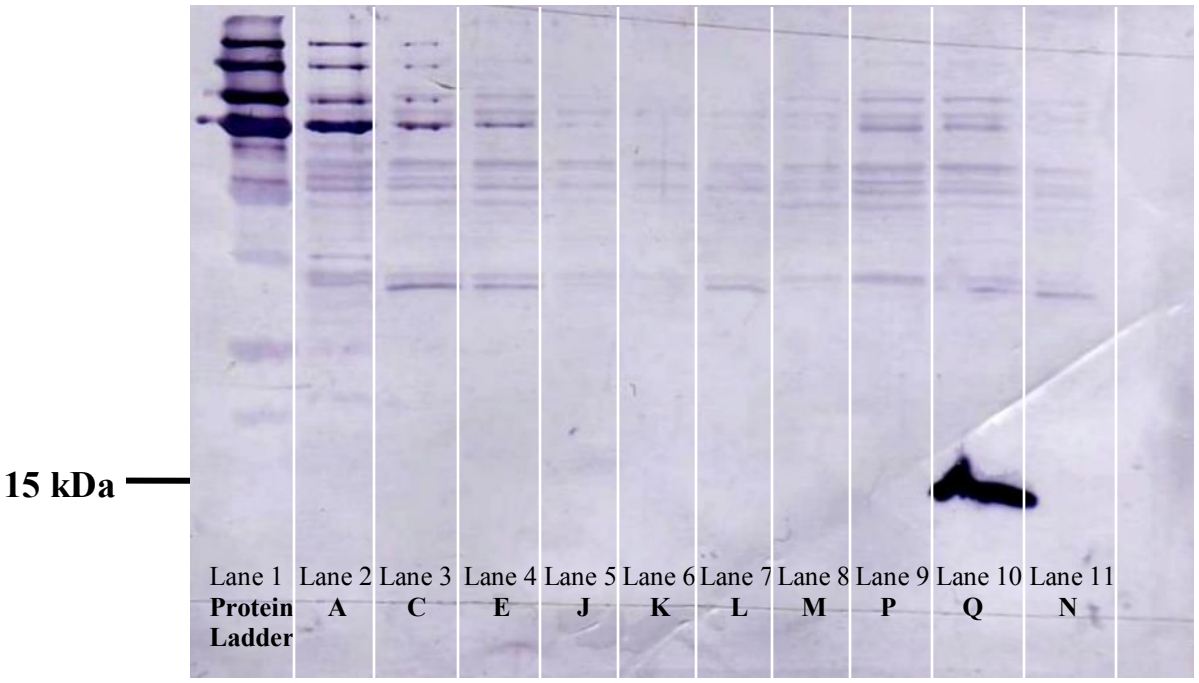
Figure 16: 1% agarose gel profile of colony-PCR performed on 30 randomly selected V_L colonies. The clones have been designated "A-Z, a-d" in order as would be counted numerically from 1-30. From left to right: Lanes 1-10 = V_L clones A-J, Lane 11 = 100 bp ladder (New England BioLabs), Lanes 12-17 = V_L clones K-P, Lanes 18-20-26 = V_L clones Q-Y, Lane 27 = 100 bp ladder (New England BioLabs), Lanes 28-32 = V_L clones Z, a-d. Photograph was taken with AlphaImager 3400 (Alpha Innotech Corporation).


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
A B C D E F G H I J K L M N O P
500 bp
18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
Q R S T U V W X Y Z a b c d
500 bp

V_L clones A, C, E-N, P-R, U-Z, and a-d were sequenced, then translated (EditSeq, DNA Star) and aligned (MegAlign, DNA Star) with the V_L-24 template. Of the 25 sequences, V_L clone Q was found to be a "wild-type" (i.e., shared sequence identity with V_L-24 parent template).

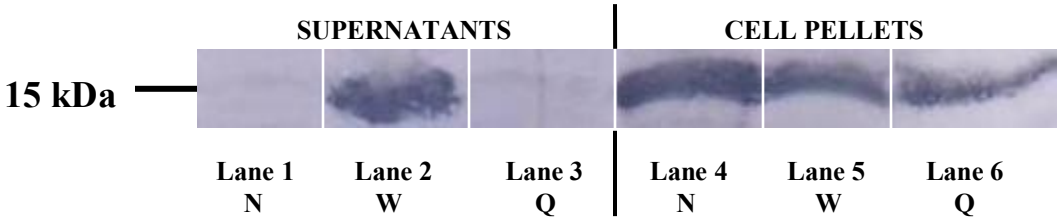
Clones A, C, E, J-N, P-Q were selected for small-scale (50 mL cultures) protein production using B2xYT/Amp (100 μg/mL) growth media (please refer to **8. Expression of 6-His tagged V_L proteins** of the "Materials & Methods" section for protocol steps). Clone Q ("wild-type" sequence) was included in the expression protocol for comparison purposes, and due to the previously known high expression levels of the V_L-24 template. Following extraction of the proteins from the periplasm, Western blot analysis was performed to verify amount of protein production in the protein supernatants (please refer to **10. Western blot** of the "Materials & Methods" section for protocol steps). Only V_L clone Q ("wild-type" V_L-24 sequence) was found to display periplasmic protein, as indicated by the band at ~15 kDa (Figure 17). Additional Western blot analysis (not shown) revealed protein in most of the cell pellets for the same clones (A, C, E, J-N, and P-Q) suggesting failure of protein localization to the periplasm in these cases.

Figure 17: Western blot of selected V_L library clones (A, C, E, J-N, P, Q) grown by the small-scale B2xYT/Amp method. 10 μL of protein supernatant extract samples from the periplasm were loaded onto a 12.5% SDS-PAGE gel and electrophoresed (200V, 30 minutes) followed by transfer to a PVDF membrane. V_L proteins were detected with mouse anti-6His IgG monoclonal Ab (Roche Diagnostics) and alkaline phosphatase (AP) labeled goat anti-mouse IgG conjugate (Cedarlane Laboratories), followed by visualization with an AP conjugate substrate kit (Bio-Rad). Protein expression for only clone Q ("wild-type" V_L sequence) was observed, as indicated by band appearance at ~15 kDa. Lane 1 = Protein Plus Standard Ladder (Bio-Rad), Lane 2 = clone A, Lane 3 = clone C, Lane 4 = clone E, Lane 5 = clone J, Lane 6 = clone K, Lane 7 = clone L, Lane 8 = clone M, Lane 9 = clone P, Lane 10 = clone Q, Lane 11 = clone N.


15 kDa
Lane 1 Protein Ladder
Lane 2 A
Lane 3 C
Lane 4 E
Lane 5 J
Lane 6 K
Lane 7 L
Lane 8 M
Lane 9 P
Lane 10 Q
Lane 11 N

V_L clones N, Q and W were arbitrarily selected for large-scale (1 L cultures) protein expression in M9S/Amp (100 µg/mL) medium (please refer to **8. Expression of 6-His tagged V_L proteins** of the "Materials & Methods" section for protocol steps). Following periplasmic extraction of the proteins, expression for each clone was verified by Western blot using 10 µL of the isolated protein supernatants and 10 µL of the resuspended cell pellets (dissolved in ~25 mL of ddH_2O). All clones produced protein in the cell pellet (Figure 18) while V_L clone W was the only one to display periplasmic expression, as indicated by the band at ~15 kDa (Figure 18).

Figure 18: Western blot of selected V_L library clones (N, W, Q) grown by the large-scale M9S media method. 10 µL of protein supernatant extract samples from the periplasm, as well as corresponding cell pellets, were loaded onto a 12.5% SDS-PAGE gel and electrophoresed (200V, 30 minutes) followed by transfer to a PVDF membrane. V_L proteins were detected with mouse anti-6His IgG monoclonal Ab (Roche Diagnostics) and alkaline phosphatase (AP) labeled goat anti-mouse IgG conjugate (Cedarlane Laboratories), followed by visualization with an AP conjugate substrate kit (Bio-Rad). Protein expression for only clone W was observed, as indicated by band appearance at ~15 kDa. Lane 1 = clone N (protein supernatant), Lane 2 = clone W (protein supernatant), Lane 3 = clone Q (protein supernatant), Lane 4 = clone N (cell pellet), Lane 5 = clone W (cell pellet), Lane 6 = clone Q (cell pellet).


SUPERNATANTS
CELL PELLETS
15 kDa
Lane 1
N
Lane 2
W
Lane 3
Q
Lane 4
N
Lane 5
W
Lane 6
Q

Figures 19 and 20 depict the His_6-tagged V_L clone W, which was grown according to the 1 L M9S method and displayed a significant level of expression. The cell culture was lysed with lysozyme and DNAse for crude protein extraction. The protein-containing supernatant was dialyzed against 10mM HEPES (pH 7.0) and 500mM NaCl, then purified by IMAC. Figure 19 depicts the IMAC profile obtained from purification of V_L clone W. The remaining non-V_L components were eluted first, with a peak forming between 10-15 mL. The His_6-tagged V_L protein was found to elute at the expected volume, as indicated by the (second) protein peak at approximately 23 mL. Protein purity was verified by SDS-PAGE analysis of the fractions corresponding to eluted V_L protein (not shown). Purified V_L protein was then passed through a Superdex 75 column (SEC analysis) for observation of molecular form (i.e., aggregation tendencies). A symmetrical peak at an elution volume of 13 mL was obtained, consistent with a monomeric protein without any aggregation (Figure 20). The chromatograms of non-aggregating sdAbs almost always consist of a single, symmetrical peak (Arbabi-Ghahroudi et al., 2009a) with slight variations in the elution volumes observed; however, monomeric proteins are generally found to be within the 12-15 mL elution volume range according to the Superdex 75 manufacturer (GE Healthcare).

Figure 19: Purification of randomly selected V_L library clone W immobilized metal affinity chromatography (IMAC) profile. V_L protein was loaded onto a Ni^{2+}-charged HiTrap Chelating 5 mL column for subsequent purification and the protein was eluted in 1 mL fractions using a 10-500 mM Imidazole gradient, in buffer containing 10 mM HEPES, 500 mM NaCl, and 500 mM Imidazole (pH 7.4). The bound material was eluted by a stepwise increase of imidazole to 10 mM, and a linear gradient of imidazole from 10 to 500 mM. The dotted line indicates imidazole concentration. The continuous line corresponds to purified protein. First peak (~11 mL) represents non-V_L components and the second peak (~23 mL) is purified V_L protein.

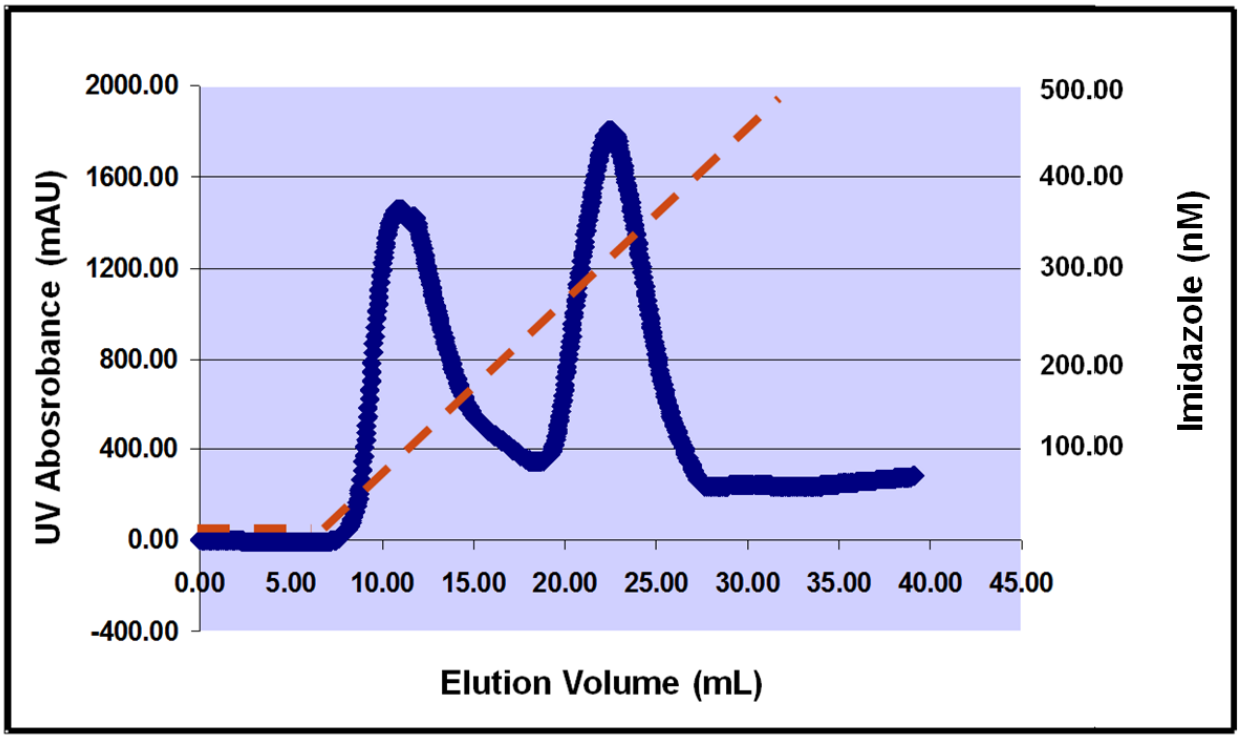
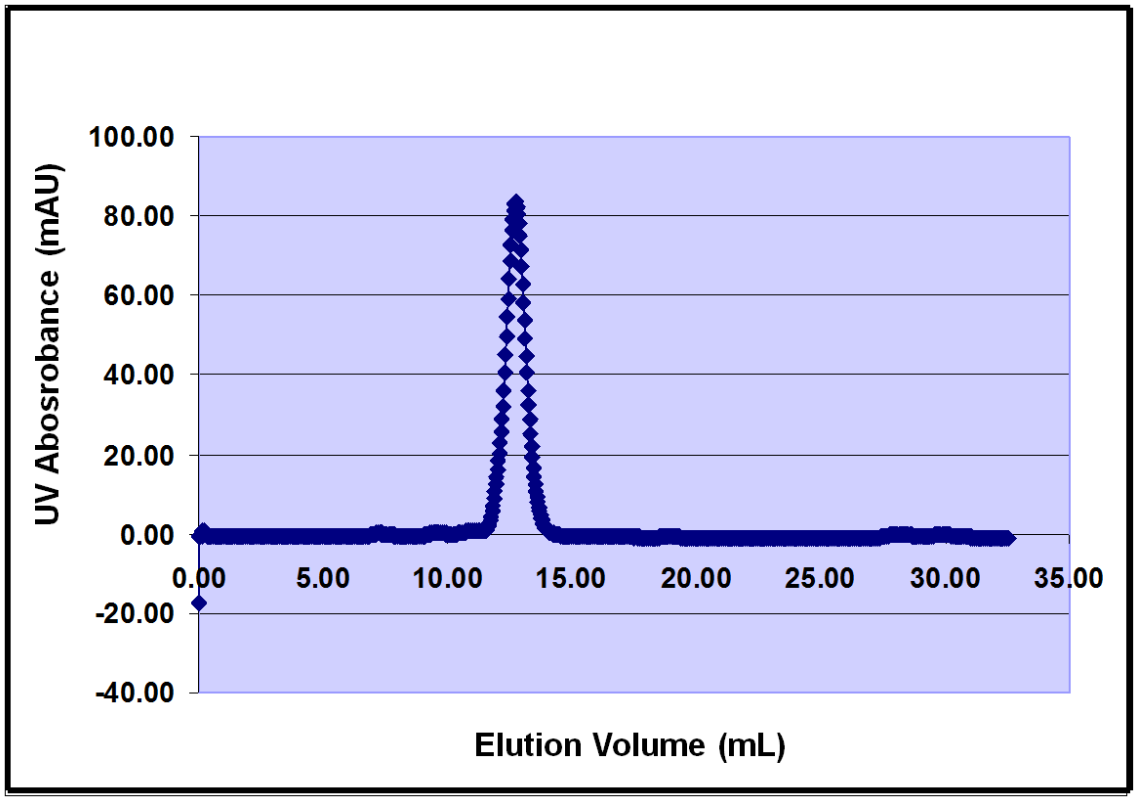

2000.00
1600.00
1200.00
800.00
400.00
0.00
-400.00
UV Abosrobance (mAU)
500.00
400.00
300.00
200.00
100.00
Imidazole (nM)
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
Elution Volume (mL)

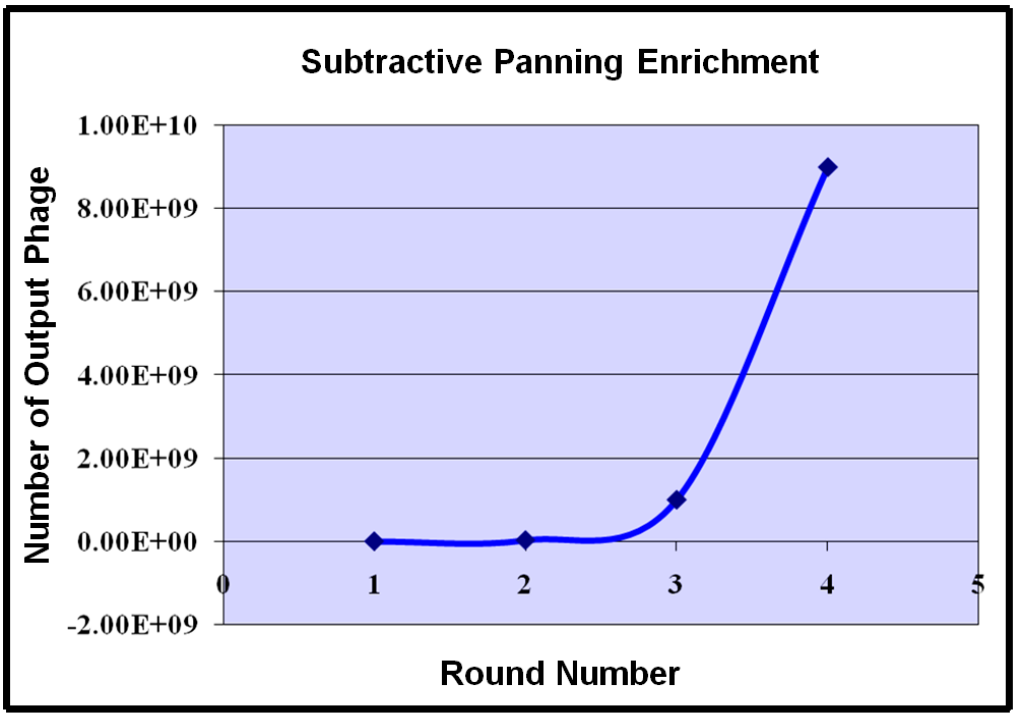
Figure 20: Size exclusion chromatography (SEC) chromatogram for V_L clone W, which was randomly selected from the synthetic library. V_L protein was loaded passed through a Superdex 75 column and eluted in 0.4 mL fractions using buffer consisting of 10 mM HEPES, 500 mM NaCl, and 0.005% P-20 detergent (pH 7.4). The SEC chromatogram displays a symmetrical peak at an elution volume of ~13.5 mL (consistent with a monomeric protein).


100.00
80.00
60.00
40.00
20.00
0.00
-20.00
-40.00
UV Abosrobance (mAU)
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
Elution Volume (mL)

5. Identification and expression of subtractive panning isolates from the constructed synthetic V_L phage display library

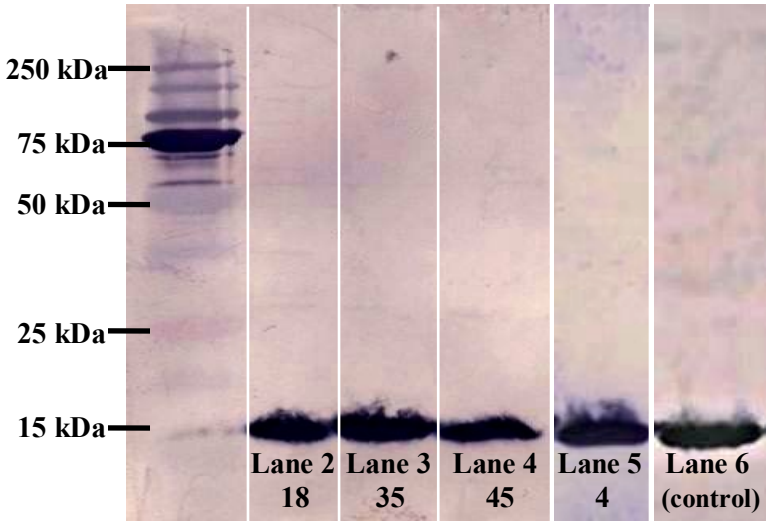
The constructed V_L phage display Ab library was subjected to four rounds of subtractive panning against two different mammalian cell lines for isolation of NRP1 binders. The output phage titre for each round of subtractive panning was monitored as an indication of enrichment for V_Ls with affinity towards cells endogenously expressing the NRP1 epitope. A total of 2 x 10^6 output phage was obtained at the end of round one, while the second round provided a total of 3 x 10^7 output phage, the third round found 1 x 10^9, and the fourth, 9 x 10^9 (Figure 21).

Figure 21: Phage output following four rounds of subtractive panning. Four rounds of subtractive panning were carried out on the constructed synthetic V_L phage display library, and the number of output phage (eluted phage before overnight amplification) was calculated from the titer plates. The exponential curve demonstrates successful enrichment of clones during the selection process.


Subtractive Panning Enrichment
Number of Output Phage
1.00E+10
8.00E+09
6.00E+09
4.00E+09
2.00E+09
0.00E+00
-2.00E+09
0
1
2
3
4
5
Round Number

Sequencing analysis of 44 isolates from the 4th round of subtractive panning revealed 15 unique V_Ls, with four clones (V_Ls 4, 18, 35, and 45) that were enriched to a greater extent compared to the other 11 isolates, as evidenced by the repetition of identical sequences following sequence alignment. The most prevalent sequence (V_L clone 45) repeated 22 times, a second sequence (V_L clone 35) was found to repeat 6 times, and two others (V_L clones 4 and 18) were each observed to repeat two times. V_L DNA inserts were amplified by PCR with gene-specific primers "HVL42 BbsI B" and "HVL89 BamHI F" for introduction of restriction enzyme sites *BpiI* (*BbsI)* and *BamHI* to the 5' and 3' ends. After digestion with *BpiI* (*BbsI)* and *BamHI*, V_Ls were ligated into the pSJF2H vector resulting in the addition of C-terminal c-Myc and $(His)_6$ tags. V_L expression in *E. coli* was carried out using the 5-day M9 minimal media method. Following extraction of the periplasmic V_L proteins, protein levels were analyzed by Western blotting. A total of 10 V_Ls were successfully expressed in *E. coli* and later purified by IMAC. In particular, a significant amount of expression was observed for those clones that were highly enriched (18, 35, 45, and 4), as indicated by the bands at ~15 kDa (Figure 22). V_L clone W, which was previously expressed by the same M9 minimal media method, was also loaded onto the gel as a positive control.

Figure 22: Western blot of V_L subtractive panning isolates 18, 35, 45, and 4. 10 μL of protein supernatant extract samples from the periplasm were loaded onto a 12.5% SDS-PAGE gel and electrophoresed (200V, 30 minutes) followed by transfer to a PVDF membrane. V_L proteins were detected with mouse anti-6His IgG monoclonal Ab (Roche Diagnostics) and alkaline phosphatase (AP) labeled goat anti-mouse IgG conjugate (Cedarlane Laboratories), followed by visualization with an AP conjugate substrate kit (Bio-Rad). Protein expression for all 4 clones was relatively high, as indicated by the bands at ~15 kDa. Lane 1 = Protein Plus Standard Ladder (Bio-Rad), Lane 2 = V_L clone 18, Lane 3 = V_L clone 35, Lane 4 = V_L clone 45, Lane 5 = V_L clone 4, Lane 6 = control V_L.


250 kDa
75 kDa
50 kDa
25 kDa
15 kDa
Lane 2
18
Lane 3
35
Lane 4
45
Lane 5
4
Lane 6
(control)

In total, 10 V_L subtractive panning isolates were successfully expressed by the 5-day 1 L M9S method, their proteins purified by IMAC, and separated by SEC for aggregation analysis. Randomized CDR positions for each of the 15 V_Ls were aligned and are shown in Table 2. Predicted molecular properties for the 15 V_Ls including molecular weight, molar extinction coefficient, isoelectric points, expression yield, and SEC-based properties such as monomericity and elution volume, are displayed in Table 3. "N/A" is designated for those V_L clones that were not successfully expressed.

Table 2: Amino acid alignment of each randomized position for the 15 V_L subtractive panning isolates.

	24	28	30	31	32	34	50	53	54	55	89	91	92	93	94	95a	95b	96
V_L 35	N	S	E	K	R	K	N	T	R	L	A	R	G	T	V	P		G
V_L 45	R	S	S	T	Y	N	Q	L	L	F	Q	S	Y	S	T			R
V_L 4	R	S	S	T	Y	N	E	E	R	W	L	N	G	T	P			R
V_L 18	G	G	G	N	W	E	Q	K	L	D	K	V	A	W	R			L
V_L 8	R	S	N	Q	N	R	A	T	L	Q	M	R	K	L	V			H
V_L 10	E	S	R	K	N	K	A	R	R	P	V	A	P	G	T	Q		T
V_L 19	E	S	L	P	N	N	D	T	R	T	A	P	S	R	R	P		H
V_L 20	N	S	D	K	D	L	E	L	L	F	Q	S	Y	S	T			R
V_L 21	S	S	E	R	R	N	S	T	R	C	Q	S	Y	S	T			R
V_L 25	A	G	D	G	N	E	Y	A	L	S	P	P	A	A	A	L		A
V_L 26	K	S	Y	D	N	L	D	Y	R	W	T	N	G	T	F			G
V_L 29	R	G	G	Q	M	Y	G	L	L	Y	K	D	A	E	S			P
V_L 31	V	S	S	N	N	K	R	S	R	V	V	R	G	Q	T	A		A
V_L 40	M	G	M	N	D	N	R	Q	R	I	G	T	T	P	A	R	K	L
V_L 41	R	S	S	T	Y	N	A	T	L	Q	K	S	N	D	A	N	S	T

Table 3: Predicted protein properties of the 15 V_L subtractive panning isolates including molecular weight, molar extinction coefficients, isoelectric points, expression yield, and SEC-based properties such as monomericity and elution volume. "N/A" or "Not Applicable" denotes those V_Ls that had not been successfully expressed using the M9 minimal media method.

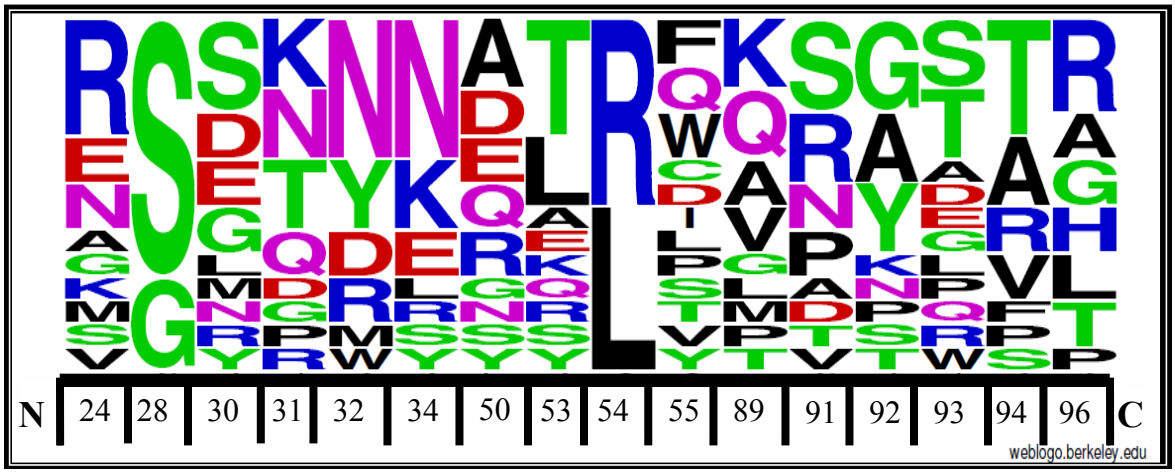
	Molecular Weight (Da)	Molar Extinction Coefficient (M^{-1} cm^{-1})	Isoelectric Point (pI)	Expression yield (mg per L culture)	Monomericity	Elution Volume (mL)
V_L 35	11587.0	10095	9.57	~4.8	Monomeric	~12
V_L 45	11682.0	13075	8.68	~1.1	Monomeric	~14.5
V_L 4	11669.0	17085	8.70	~2.2	Monomeric	~12.5
V_L 18	11560.0	21095	8.68	~1.8	Monomeric	~13.5
V_L 8	11655.2	10095	9.62	~1.7	Monomeric	~12.5
V_L 10	11599.0	10095	9.57	~0.6	Monomeric	~12
V_L 19	11651.0	10095	8.73	N/A	N/A	N/A
V_L 20	11592.9	11585	6.75	N/A	N/A	N/A
V_L 21	11620.9	11585	9.22	~2.1	Monomeric	~12.5
V_L 25	11290.6	11585	5.57	~2.9	Monomeric	~13.5
V_L 26	11605.9	18575	7.97	N/A	N/A	N/A
V_L 29	11487.9	13075	7.97	N/A	N/A	N/A
V_L 31	11532.9	10095	9.62	~0.4	50% aggregating	~7; ~12
V_L 40	11789.3	10095	9.39	~4.4	Monomeric	~12.5
V_L 41	11660.9	11585	8.68	N/A	N/A	N/A

Randomized positions (24, 28, 30, 31 32, 34, 50, 53, 54, 55, 89, 91, 92, 93, 94, and 96) for the 15 clones were analyzed using the WebLogo software (Crooks et al., 2004) which graphically plots amino acids from aligned sequences top to bottom from most to least frequently occurring (Figure 23). The overall height of symbols within the stack indicates the relative frequency of each amino acid at that position. Positions 95a and 95b were excluded for sequence alignment purposes.

The most frequent residue observed at position 24 was arginine (R), followed by glutamic acid (E), and asparagine (N). For position 28, serine predominated over glycine (G). Serine (S), followed by aspartic acid (D) and glutamic acid (E) were the most frequent amino acids at position 30; lysine (K), then asparagine (N) and threonine (T) were most frequently found for position 31; asparagine (N) was enriched at positions 32 and 34. Alanine (A), followed by aspartic acid (D) and glutamic acid (E) were found to be most preferred at position 50; threonine (T) and leucine (L) were most prominent at position 53, while there was an equal distribution of arginine (R) and leucine (L) at position 54; phenylalanine (F), glutamine (Q), and tryptophan (W) were also equally enriched at position 55. Position 89 had a slight preference for lysine (K), whereas position 91 displayed serine (S) and arginine (R) as prominent residues. Glycine (G) and alanine (A) were observed to be most prevalent at position 92; serine (S) and threonine (T) were equally predominant amino acids at position 93; threonine (T) was most frequent followed by alanine (A) at position 94, and arginine (R) for position 96.

A prevalence of polar amino acids and a high occurrence of basic residues were found from sequence analysis of the randomized positions for the 15 isolates (Figure 23).

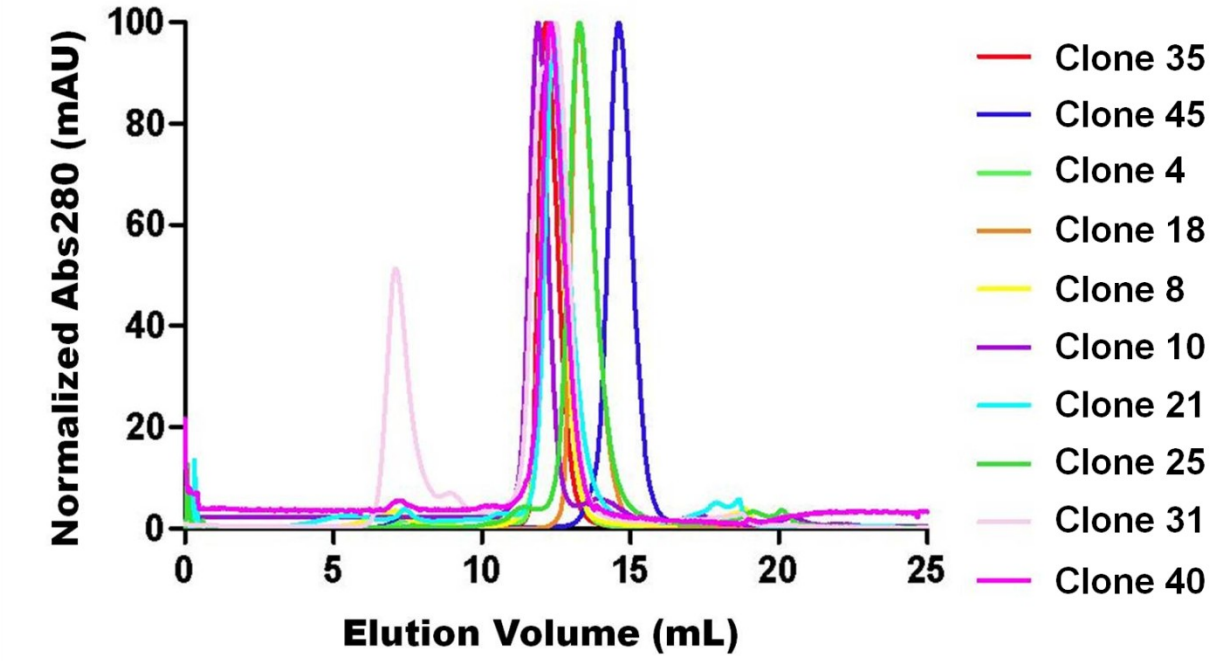
Figure 23: Sequence analysis of randomized positions in the 15 V_L subtractive panning isolates. Fractional occurrence of amino acids in the randomized positions 24, 28, 30, 31 32, 34, 50, 53, 54, 55, 89, 91, 92, 93, 94, and 96 for the 15 V_L clones. Amino acids are arranged top to bottom from most to least frequently occurring. The figure was prepared from the data of Table I, by means of WebLogo (Crooks et al., 2004).


N
24
28
30
31
32
34
50
53
54
55
89
91
92
93
94
96
C
weblogo.berkeley.edu

6. Size exclusion chromatography analysis of V_L subtractive panning isolates

Size exclusion chromatography (SEC) analysis was performed on 10 of the 15 V_L isolates. IMAC-purified proteins were passed through a Superdex-75 column (which optimally resolves proteins in the 3-70 kDa range) and eluted in buffer containing 10 mM HEPES, 500 mM NaCl, and 0.005% P-20 detergent, at pH 7.4. Nine of these 10 proteins were monomeric, while clone 31 demonstrated aggregation tendencies, as evidenced by the (first) peak at an approximate elution volume of 7 mL, and the (second) major peak observed at ~12 mL (Figure 24). The monomeric and aggregate peaks for clone 31 were observed to be at an approximate ratio of 2:1 (monomeric):(aggregate). The remaining 9 clones were free of any aggregation, displaying symmetrical peaks in the range of 12-15 mL. V_L clones 35 and 10 were seen to elute first at ~12 mL, while clones 8, 21 and 40 appeared to elute immediately after, at ~12.5 mL. Clones 18, 25 and 31 each eluted later at ~13.5 mL, and clone 45 was the last of the group with an elution volume of ~14.5 mL. Differences in elution volumes are attributed to the varying apparent molecular weights among the V_Ls.

Figure 24: Size exclusion chromatography (SEC) profiles obtained for 9 V_L subtractive panning isolates. SEC curves for V_L clones 35, 45, 4, 18, 8, 10, 21, 24, 31, and 40; major peaks correspond to the monomeric form of each V_L and retention volumes are shown for each. Nine of the 10 V_Ls are virtually monomeric except for clone 31, which displayed a 2:1 ratio of monomeric to aggregated protein. Elution volumes for each V_L are also displayed. Variations in apparent molecular weights of sdAbs (including the V_Ls seen here) are not uncommon (Jespers et al., 2004; Arbabi-Ghahroudi et al., 2009b).


Normalized Abs280 (mAU)
100
80
60
40
20
0
0
5
10
15
20
25
Elution Volume (mL)
Clone 35
Clone 45
Clone 4
Clone 18
Clone 8
Clone 10
Clone 21
Clone 25
Clone 31
Clone 40

7. Flow cytometric analysis of V_L subtractive panning isolates

The four V_L clones (35, 45, 4, and 18) that had displayed the greatest enrichment and that had also been highly expressed, were tested for their potential binding to both PC12 and HEK293 cells, through flow cytometry. Cells (5 x 10^6/mL) were first washed then incubated under one of the following conditions: 1) alone, 2) with anti-6His PE, 3) with 5 µg of each of the four V_Ls, and 4) with a V_L randomly selected from the library as control. Following a wash, reaction mixtures containing V_Ls were incubated with anti-6His PE then subjected to a second wash. Finally, cell reaction mixtures were exposed to PBS-1% formaldehyde and data acquisition was carried out using a flow cytometer.

Anti-6His PE-conjugated secondary Ab was used for V_L detection and a maximum of 10,000 events was established as part of the gating hierarchy. Controls included incubation of each cell line alone, with anti-6His PE, as well as with a V_L clone not isolated from subtractive panning. Flow cytometric histograms represent a single experimental trial and depict cell count as a function of mean fluorescence intensities (MFIs) for each of the cell reaction mixtures (Figure 25, A-N); MFI values obtained were plotted as bar graphs (Figure 26). PC12 and HEK293 cells were each incubated alone (Figures 25A and 25H), with anti-6His PE (Figures 25B and 25H), with 5 µg of V_L 35 (Figures 25C and 25J), 5 µg of V_L 45 (Figures 25D and 25K), 5 µg of V_L 4 (Figures 25E and 25L), 5 µg of V_L 18 (Figures 25F and 25M), or with 5 µg of a "control V_L" that was randomly selected from the constructed synthetic library and not isolated from subtractive panning (Figures 25G and 25N). A difference in the MFI values was observed between the NRP1-expressing PC12 cells and the NRP1-lacking HEK293 cell line, with respect to each of the four V_Ls (35, 45, 4, and 18) isolated from subtractive panning. Reaction mixtures of PC12 and HEK293 cells each incubated alone (Figure 25, A and H), displayed MFIs of 97 and

64, respectively (Figure 26). PC12 and HEK293 cells each incubated with only anti-6His PE (Figure 25, B and I) yielded MFIs of 106 and 87, respectively (Figure 26). 5 µg of V_Ls 35, 45, 4, and 18, added to PC12 cells (Figure 25, C-F), resulted in MFIs of 278, 291, 237, and 273, respectively (Figure 26). 5 µg of V_Ls 35, 45, 4, and 18, added to HEK293 cells (Figure 25, J-M), yielded MFIs of 142, 167, 158 and 186 (Figure 26). 5 µg of a control V_L (i.e., randomly selected from the constructed synthetic library and not isolated from subtractive panning) was added to both PC12 and HEK293 cells (Figure 25, G and N) and resulted in MFI values of 193 and 156, respectively (Figure 26). Thus, a slight difference in mean fluorescence between the NRP1-expressing PC12 cells and the NRP1-lacking HEK293 cell line was observed with respect to each of the 4 V_L subtractive panning isolates (clones 35, 45, 18, and 4).While this experimental trial indicated binding of V_Ls 35, 45, 4, and 18, to PC12 cells, and not to HEK293 cells, other trials presented inconsistent observations such as binding signals for the "control V_L" and similar mean fluorescent intensities between the HEK293 and PC12 cell reaction mixtures. Therefore, no definitive conclusions can be inferred from these flow cytometry experiments due to the discrepancies and conflicting data.

A) PC12 (no anti-6His PE)
B) PC12 (+ anti-6His PE)
C) PC12 + V_L 35 (+ anti-6His PE)
D) PC12 + V_L 45 (+ anti-6His PE)
E) PC12 + V_L 4 (+ anti-6His PE)
F) PC12 + V_L 18 (+ anti-6His PE)
G) PC12 + control V_L (+ anti-6His PE)
H) HEK293 (no anti-6His PE)
I) HEK293 (+ anti-6His PE)
J) HEK293 + V_L 35 (+ anti-6His PE)
K) HEK293 + V_L 45 (+ anti-6His PE)
L) HEK293 + V_L 4 (+ anti-6His PE)
M) HEK293 + V_L 18 (+ anti-6His PE)
N) HEK293 + control V_L (+ anti-6His PE)

Figure 25: Flow cytometric histograms for 4 V_L subtractive panning isolates (clones 35, 45, 18, and 4). PC12 and HEK293 cells were either incubated alone (A, H), with anti-6His-Pe (B, I), or with 5 μg of pure monoclonal protein corresponding to the V_L clones 35, 45, 4, and 18 isolated from round four of subtractive panning (C-F, J-M). A V_L randomly selected from the constructed synthetic library (not isolated from subtractive panning) was also included as a control (G, N). Anti-6His PE conjugated secondary Ab was used for V_L fluorescent detection. Each cell line was also incubated alone with anti-6His PE, as well as with a V_L control protein.

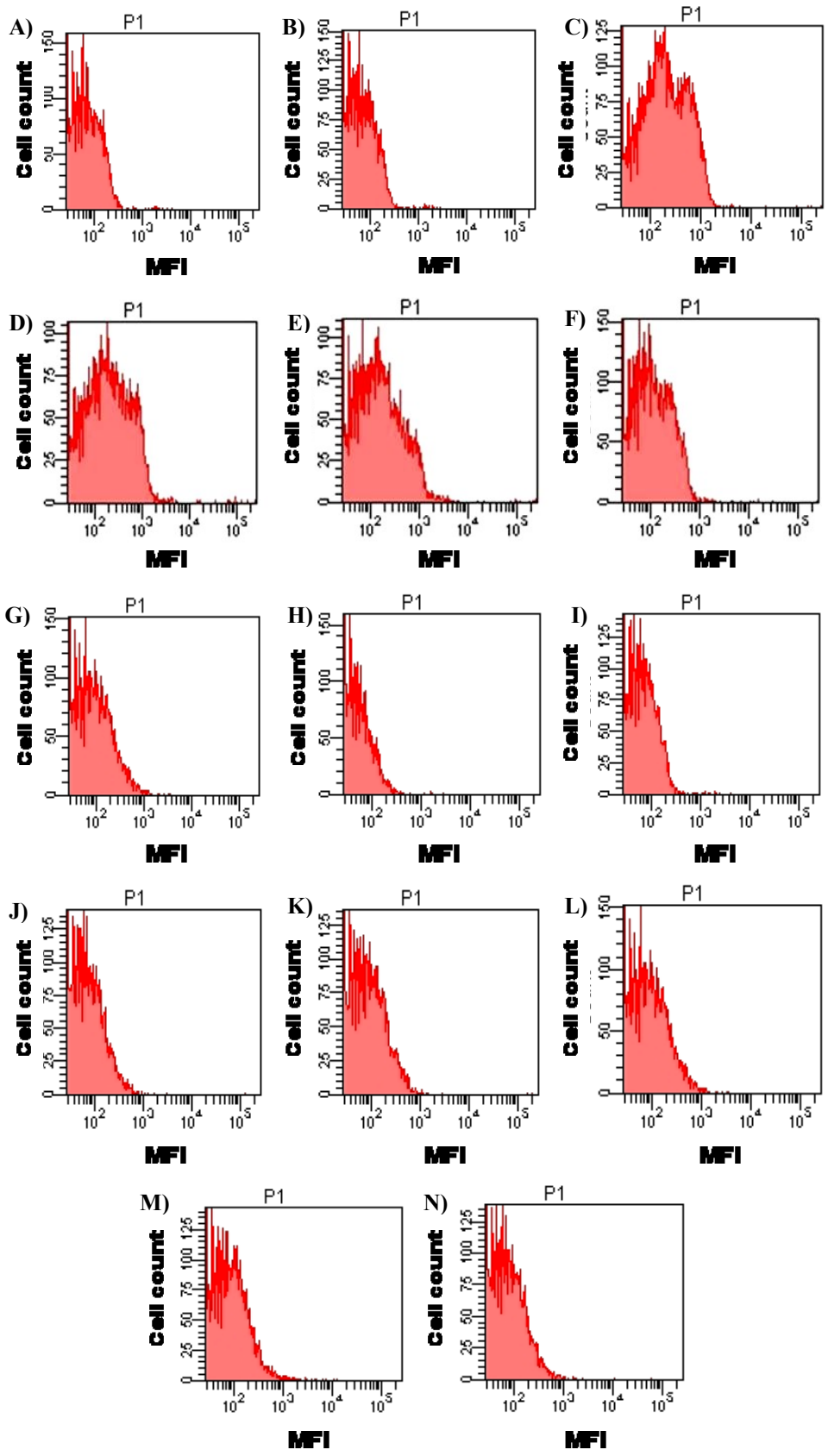
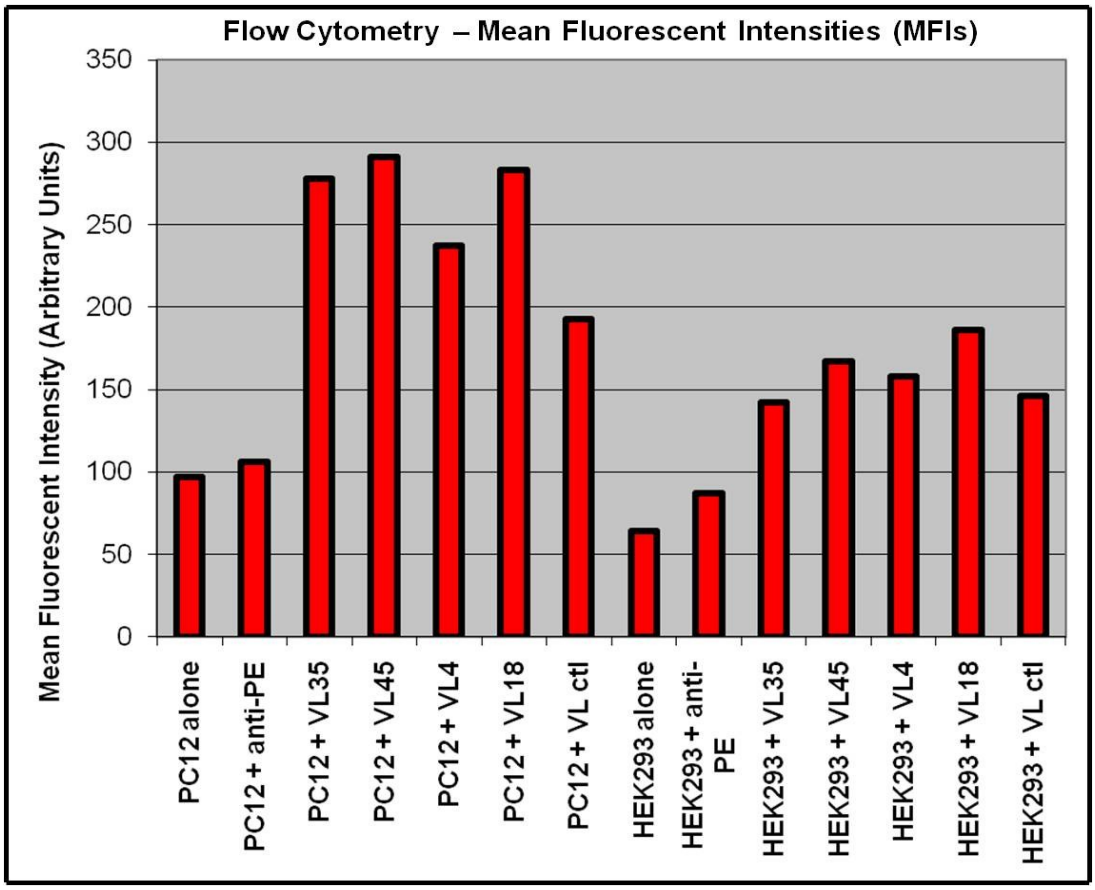

A)
B)
C)
D)
E)
F)
G)
H)
I)
J)
K)
L)
M)
N)
P1
Cell count
MFI

Figure 26: Mean fluorescent intensities obtained from flow cytometric analysis of 4 V_L subtractive panning isolates (clones 35, 45, 18, and 4). Bars depict fluorescent intensities obtained from flow cytometric analysis of the V_L clones 35, 45, 4, and 18 with NRP1-positive PC12 cells and NRP1-negative HEK293 cells.


Flow Cytometry – Mean Fluorescent Intensities (MFIs)
Mean Fluorescent Intensity (Arbitrary Units)
350
300
250
200
150
100
50
0
PC12 alone
PC12 + anti-PE
PC12 + VL35
PC12 + VL45
PC12 + VL4
PC12 + VL18
PC12 + VL ctl
HEK293 alone
HEK293 + anti-PE
HEK293 + VL35
HEK293 + VL45
HEK293 + VL4
HEK293 + VL18
HEK293 + VL ctl

VII. DISCUSSION

The majority of therapeutic Abs approved by the Therapeutic Products Directorate (TPD) of Health Canada and most of those in clinical trials, are full size Abs (i.e., 150 kDa) in conventional IgG format (Dimitrov and Marks, 2009). Scaffolds derived from sdAb domains are ~10-fold smaller than full size Abs and are easier to manufacture and express in microbial cell cultures (Holt et al., 2003; Saerens et al., 2008; Dimitrov, 2009, Chen et al., 2010). The enhanced ability to penetrate tissues and access cryptic epitopes (a hidden antigenic site that is present on buried surface subunits; cryptic epitopes are antigenically active only after dissociation of protein aggregates; Thatcher et al., 2000), as well as a high efficiency of bio-distribution and clearance from the kidneys and blood make sdAbs ideal for use in clinical applications (i.e., drug carriers for targeted therapy) (Yokota et al., 1992; Yokota et al., 1993; Nuttall et al., 2000; Cortez-Retamozo et al., 2002; Liu et al., 2007).

Considerable progress has been made in the construction and characterization of Ab libraries (Griffiths et al., 1994; Nissim et al., 1994; Pini et al., 1998; Viti et al., 2000), ever since the first isolation of human monoclonal Abs from large naïve phage display libraries was reported in 1991 (Marks et al., 1991). The robustness of a library, described as the repeated ability to isolate clones with binding specificities against a wide assortment of Ags, represents one of the most essential requirements in phage display applications. Another important aspect is how well each the of individual clones perform in terms of expression yield, stability, solubility, and monomericity. Studies have consistently shown that increases in functional library size facilitate the isolation of high affinity clones; increases in library size (up to a maximum of ~10^{11} genes) and in library quality have already been realized using phage display (Perelson and Oster, 1979; Bradbury and Marks, 2004; Carter, 2006). Synthetic Ab libraries,

which are created by introducing synthetic DNA into the regions encoding the CDRs of defined frameworks, now rival or even surpass the recognition potential of natural immune systems (Sidhu and Fellouse, 2006). The number of fully human Abs in clinical trials has experienced a rapid increase, with a large proportion obtained via phage display (Holliger and Hudson, 2005; Hoogenboom et al., 2005; Reichert et al., 2005).

Thus far, in contrast to V_Hs, V_L single domains have not been investigated in depth with regard to library construction and screening. V_Ls were not believed to be powerful Ag-binding components until reports of V_L functionality first appeared (Gavish et al., 1977). Since then, there has been a steady rise in V_L-focused studies; one such investigation demonstrated an increase in Ag-binding affinity for the V_L domain of a recombinant immunotoxin (fused to a truncated form of *Pseudomonas* exotoxin) compared to the V_H domain of the same immunotoxin (Brinkmann et al., 1993). In addition, the V_L portion of a ferritin-specific Ab F11 was found to maintain the Ig fold (Martsev et al., 1998; Martsev et al., 2000; Martsev et al., 2002a) and Ag-binding function of its parental, whole Ab (Martsev et al., 2002b). While V_Hs are considered to have a predominant role in forming the Ag binding site and are generally more favourable than V_Ls in terms of structure for the generation of sdAbs as therapeutic agents, V_Hs also display relatively low levels of solubility and higher occurrences of aggregation (Givol, 1991; To et al., 2005). Consequently, the solubility of V_Hs must be increased via genetic manipulations, such as "camelization" and "llamination" (i.e., the incorporation of crucial solubilizing amino acids, derived from camelid sdAbs, into human V_H domains) (Davies and Riechmann, 1994; Davies and Riechmann, 1995; Sheriff and Constantine, 1996; Frenken et al., 2000; Tanha et al., 2001). V_L domains confer a number of advantages, including a high degree of solubility and an increased resistance to aggregation, making them useful for developing treatments against human

diseases (Jäger and Plückthun, 1999; Ohage and Steipe, 1999; Dubnovitsky et al., 2000; Ewert et al., 2003).

The project described herein thus aimed to create a functional, complex synthetic human V_L phage display library in *E. coli*, analyze it by sequence analysis, and to isolate potential NRP1-specific binders by subtractive panning. V_L isolates were identified by sequencing and characterized by SEC and flow cytometric analysis.

1. Efficiency of transformation for electrocompetent cells

The construction of highly diverse phage display Ab libraries largely depends on the transformation efficiency of the electrocompetent *E. coli* host cells, which can vary significantly across different batches of prepared cells, due in large part to the inherent difficulties in standardizing washing steps (Dower et al., 1988; Sambrook and Russell, 2001; Wang et al., 2007). A more recent approach in the preparation of electrocompetent *E. coli* cells, which requires two centrifugation steps through a glycerol/mannitol density cushion, has been reported to give efficiencies of $> 3 \times 10^{10}$/μg DNA for TG1 (Warren, 2011). The electrocompetent *E. coli* TG1 cells prepared here using 10% ultrapure glycerol and HEPES washing solutions displayed transformation efficiencies of 10^8 CFUs/μg for ligation products containing phage heteroduplex DNA and 3×10^9 CFUs/μg for the pUC18 vector. This was in agreement with efficiencies typically reported for the TG1 strain of *E. coli* in the range of 10^7-10^8 CFUs/μg DNA generally obtained for the *E. coli* TG1 strain (most frequently used in the construction of phage display libraries) (Dower et al., 1988; Fiedler and Wirth, 1988; Hanahan et al., 1991; Marks, 1995; Viti et al., 2000; He et al., 2002; Rauchenberger et al., 2003). It was also noted that the freshly prepared batch displayed greater efficiencies compared to the commercially available competent

cells, and lowering the tetracycline concentration on agar plates permitted for greater efficiencies to be obtained.

2. Design, construction and characterization of the synthetic V_L library

Synthetic libraries have proven to be very effective in generating functional, high affinity Abs against a wide variety of Ags (Fellouse et al., 2004; Fellouse et al., 2005; Lee et al., 2007; Liang et al., 2007), as have those based on a single germ line segment (Nissim et al., 1994; Carnemolla et al., 1996; Pini et al., 1997; Borsi et al., 1998). The FRs of synthetic Ab libraries are generally designed according to the consensus sequences of human Ab germline families (Yin et al., 2008). In order to create a stable and highly variable synthetic library of functional V_L sdAbs, an unmodified, naturally occurring human V_L scaffold sequence (i.e., V_L-24; 321 bp; 107 residues) was used as a framework. V_L-24 was selected as template for creation of the synthetic library mainly because of its resistance to aggregation, solubility, high expression yield, and reversible thermal unfolding.

Diversity is generally introduced into a synthetic phage display Ab library by randomizing specific nucleotide positions through the use of synthetic oligonucleotides. The randomization approach of this project relied on trinucleotides consisting of "N", "K", and "R" (where "N" = any of the four nucleotides adenine, thymine, guanine, and cytosine; "K" = thymine or guanine; "R" = adenine or guanine), which minimizes over-representation of the most commonly occurring amino acids in the genetic code (i.e., serine, leucine and arginine, each encoded by six different codons) (Patrick and Firth, 2005). For example, using the "NNN" trinucleotide to introduce sequence degeneracy into a library would result in prevalence of the most commonly occurring sequence variants (i.e., combinations of serine, leucine and arginine)

compared to the least commonly occurring residues (i.e., methionine or tryptophan, each of which are encoded by only one codon) (Patrick and Firth, 2005). Randomization using reduced trinucleotide codon sets (e.g., "NNK") has consistently produced repertoires with superior diversity (Cárcamo et al., 1998; Holt et al., 2000; Kelley and Momany, 2003; Sidhu et al., 2004; Wolkowicz et al., 2005; Liu et al., 2009; Mao et al., 2010; Shukla and Krag, 2010) by maximizing the use of all 20 natural amino acids while allowing for control over the amino acid distributions (Knappik et al., 2000; Yin et al., 2008). The possibility of any termination codons being included is also eliminated through the use of *E. coli* strain TG1 for phage production (Barbas et al., 2001; Krumpe et al., 2007): the TG1 bacterial host encodes a suppressor tRNA that allows for readthrough (i.e., suppression) of the amber stop codon TAG; instead, this codon is translated as glutamine, effectively eliminating the occurrence of truncated Abs within the library (Li et al., 2002; Suzuki et al., 2005; Marcus et al., 2006). While the "NNS" trinucleotide (where "S" = a 50% mix each of cytosine and guanine) is essentially equivalent to "NNK", studies indicate the latter as the favoured choice due to codon usage preferences in *E. coli* and for long-term library maintenance in the bacterial host (Patrick and Firth, 2005).

Ab variable domains have integral roles in maintaining a stable tertiary structure while simultaneously supporting the extensive sequence variability that provides the immune system with a countless variety of binding specificities. Numerous structural studies and statistical analyses of Ab sequences have resulted in the ability to distinguish between the hypervariable sites that influence Ag binding and the conserved frameworks that maintain the tertiary fold (Kabat, 1982; Padlan et al., 1995). Previous studies have also allowed for the determination of those specific positions within the CDRs of V_L-24 that were highly conserved as well as those most likely to tolerate randomization without severely compromising the scaffold's stability,

while significantly contributing to Ag binding. Libraries generally vary considerably in the number of positions randomized: there have been cases in which all CDRs were completely randomized, while in others only selected positions were chosen for diversification (Hayashi et al., 1994; Braunagel, 2003).

Each CDR contributes differently to Ag affinity and specificity: studies involving crystallographic analysis of Ag/Ab interactions have established that, of the 3 CDRs, CDR3 contributes the most to library diversity and to Ag binding through a large surface area of contact due to its positioning directly in the centre of the Ag-binding site (Chothia et al., 1989; Davies and Cohen, 1996; Liang et al., 2007). CDR3 loops are considered to be the most variable part of the paratope, and are therefore likely to tolerate sequence and length variations (Martineau, 2010). Randomizations of sdAb scaffolds are thus most often carried out on CDR3 and, in many instances using CDR3 of various lengths, something that can significantly improve library diversity (Tanha et al., 2002). Longer CDR3 regions can contribute to an enlarged surface area of interaction with Ags leading to isolation of Abs with potentially higher affinity (Chen et al., 2009). The human immune system naturally produces Abs that differ in CDR length (this is particularly true for CDR3 of both V_Ls and V_Hs), allowing for Abs to adopt novel shapes with which to bind target Ags (Collis et al., 2003). Since CDR3 is at the centre of the Ag-binding pocket and is the most hypervariable region, reduced CDR3 diversity would drastically limit the efficiency of the Ab response. Creating V_Ls with longer CDR3s would therefore contribute to the level of complexity needed for successful isolation of Ag-specific, high affinity lead clones. In addition, previous studies examining protein L-binding V_Ls found a variation in the CDR3 length among the isolates, with most CDR3s consisting of 9 residues, a few comprised of 10 residues, and a small proportion composed of 11 residues. For these reasons, three different

lengths of CDR3 were incorporated into the construction of the V_L library described in this project and, in comparison to CDRs 1 and 2, CDR3 was designed to contain the greatest proportion of residues randomized. While previous studies have omitted CDR1 and CDR2 randomization and limited randomization to CDR3 in order to maintain design simplicity and to avoid theoretical library sizes that exceed practical library sizes, the result is effectively a compromise in library diversity. The V_L library constructed herein effectively maximizes diversity by introducing randomization in all 3 CDRs in addition to incorporating 3 different CDR3 lengths; these features should increase the likelihood of isolating binders to a specific target Ag.

The approach taken in this project involved first constructing a smaller library in which CDR3 alone was randomized. Phage DNA from this smaller library was used to create the second library in which all 3 CDRs were randomized. This approach made it possible to assess randomization efficiency (a factor determining library diversity) through sequencing of select clones from the CDR3-randomized library prior to building the final library. To assess quality and diversity of the constructed V_L library, expected and observed amino acid frequencies were determined for each of the randomized CDR positions of 100 V_Ls arbitrarily selected from titer plates. In determining statistical significance of the V_L library's amino acid frequencies, FDI scores (based on a 95% confidence interval) were calculated for each randomized position (Figure 14, graphs A-R). Interestingly, wild-type amino acids were seen to predominate ($p < 0.05$) at virtually all of the randomized positions. Positions 89 and 91 of CDR3 were the only exceptions. This phenomenon is most likely attributed to annealing efficiency of the oligonucleotide-directed method of mutagenesis. Analysis of the 100 DNA sequences revealed that, of the positions targeted for randomization that shared amino acid identity with V_L-24, 92%

consisted of wild-type V_L-24 nucleotides (Figure 15). Factors such as DNA secondary structure, reaction conditions (i.e., salt concentration and annealing temperature), oligonucleotide purity, and cross-hybridization (annealing of oligonucleotides to incorrect locations) all contribute to a poor annealing efficiency (Wassman et al., 2004). Biological censorship could have also played a role; this is a frequent occurrence in phage libraries whereby residues compromising phage stability or replication tend to be selected against or suppressed within a library, while those considered advantageous to phage are favoured and propagated within the library (Rodi and Makowski, 1999; Rodi et al., 2002; Krumpe et al., 2007).

Despite instances of position-specific and wild-type biases within the CDRs, all 20 natural amino acids were well represented and occurred to a significant extent. Distinguishing characteristics included significant prevalence of "wild-type" amino acids in most of the CDR positions randomized, a consistent preference for proline throughout the central portion of CDR3 as well as near the N-terminus of CDR2, and a higher than expected occurrence for alanine at the N-terminal end of CDR3. There were only three cases of under-representation observed throughout all of the randomized positions; position 28 (the N-terminus of CDR1) displayed a lower than expected frequency for glycine, position 50 (the C-terminus of CDR2) demonstrated under-representation of glutamine, and one of the insertion positions (95b) was significantly lacking in threonine, serine and basic arginine. In comparison to cases of over-representation, the under-representation biases observed were generally considered weak while maintaining statistical significance, as reflected by the corresponding FDI values. The level of significance defined by the P-value is almost always established at 0.05 or smaller in scientific applications. The significance level of 0.05 means there is a 5% chance that the observed test statistic is

caused by random chance (null hypothesis), and the researcher can be 95% sure that the observed value is significantly different from the theoretical value.

It can be inferred from these results that increases and decreases in the numbers of particular residues (i.e., over-representation of proline) likely contribute to the overall stability and integrity of the V_L sdAb. Proline plays crucial roles in chain conformation and protein folding (Tchernychev et al., 1997), commonly occurs in proximity to protein-protein interaction sites (Kini and Evans, 1995; Kay et al., 2000), and has also been observed to occur in the majority of human Vκ (variable domain of the kappa light chain) CDR3s, defining a rigid omega-loop conformation (Ewert et al., 2003). Instances of under-represented residues could also be a result of sequence toxicity. The uneven positional distribution of amino acids can also be attributed to differences in infectivity of individual V_Ls in the library toward the host bacterium and biosynthesis variations of phage particles within *E. coli* during phage assembly and export stages. A decrease in population diversity of library clones (of more than 7-fold) has been noted to occur during phage amplification in bacterial hosts, whereby the occurrence of unique library clones changes during growth of the clones in culture medium due to the different efficiencies of phage assembly and export for different clones (Kuzmicheva et al., 2009). This can limit the number of binding clones identified by panning, particularly those screened against cell targets with multiple binding sites (Derda et al., 2011). However, the synthetic V_L library created here seems to contain a high level of diversity in all 3 CDRs according to sequence analysis. Sequence alignment further revealed complete sequence homology to the V_L-24 template for three of the randomly picked V_Ls; the remaining clones demonstrated a considerable amount of variability at randomized positions. With respect to individual CDRs, it was

interesting to note that 13 clones out of the 100 randomly selected for sequence analysis contained wild-type CDR3 in comparison to the 5 clones that displayed wild-type CDR2 and 3 that had wild-type CDR1. This could also be due to the use of phage DNA that had to be isolated from the primary CDR3-randomized library for subsequent generation of the final repertoire, in which case a certain amount of CDR3 diversity would be lost in the extra step. Alternatively, the annealing efficiency could have been lower during the randomization of the CDR3 library. It was additionally noted that clones containing one insertion ("CDR3a") and two insertions ("CDR3b") in CDR3, were more scarcely found in comparison to those lacking any additional "NNK" insertions in CDR3. This can be attributed to a low efficiency of amino acid insertion or a lower tolerance of the CDR3 loop to accommodate inserted residues in the particular V_L-24 gene. Generally, the CDR3 loop structure of the V_L domain is well suited to insertions, with a classic β-hairpin likely to present multiple conformations (Krykbaev et al., 2002); β-hairpin motifs each consist of two strands adjacent to one another and oriented in an antiparallel arrangement (such that the N-terminal end of one sheet is adjacent to the C-terminal end of the next) and linked by a short loop of 2-5 amino acids (Lee and Shin, 2001).

Due to physical and technical limitations, it is virtually impossible to achieve the entire spectrum of sequence diversity that could theoretically be achieved in a library. The randomization schemes adopted in library construction strategies can have a large impact on the ultimate quality and screening efficiency of repertoires. The redundancy of the genetic code refers to the disproportion or uneven representation of amino acids that are encoded by conventional trinucleotides such as "NNK" (Yin et al., 2008). For example, methionine is encoded just once by "NNK", while arginine is encoded three times. Furthermore, the gene-to-protein ratio increases exponentially with the number of randomized codons (i.e., when

randomizing with "NNN", 64^n genes are required to encode 20^n proteins, where n is the number of positions to be randomized); this exponential increase means that the number of genes rapidly exceeds cloning capabilities, causing libraries to be smaller than what would normally be required for full amino acid representation (Hughes et al., 2003). While "NNK" and "NNS" randomizations ameliorate the situation, an exponential relationship still exists between codon number and the ratio of genes to encoded proteins. To optimize randomization efficiency, the redundancy of codons encoding different amino acids would have to be completely eliminated; this bias, however, is inevitable and reflects the degeneracy of the genetic code. A number of alternative strategies have been proposed to circumvent any limitations imposed by randomization procedures (Knappik et al., 2000; Hughes et al., 2003; Yin et al., 2003; Neylon, 2004; Tabuchi et al., 2004). One such method attempts to eliminate genetic redundancy and inherent amino acid biases through theoretical maximum efficiency ("MAX") randomization involving the cloning of only one codon per encoded protein, thereby eliminating any exponential relationship regardless of the number of randomized codons (Hughes et al., 2003). The randomization method proposed by Hughes et al. (2003), which also eliminates termination codons, involves a template oligo randomized with the "NNN" trinucleotide that is hybridized with a pool of 20 synthetic "selection oligos", each containing the MAX codon for optimal expression of a single amino acid. Any single selection oligonucleotide thus base pairs with its complement in the template. "Selectional hybridization" can also be used to generate combinatorial randomization cassettes, allowing for multiplicative numbers of randomized genes to be obtained. For example, considering three "MAX" codons, a "selection" oligonucleotide encoding the first codon can hybridize alongside another "selection" oligonucleotide encoding the second and third codons; therefore 60 (i.e., 20 + 20 + 20) selection oligonucleotides can

produce 8000 (i.e., 20^3) different genes (Hughes et al., 2003). Although "MAX" technology is expected to rival conventional randomization in terms of diversity, it remains both more complex and expensive than conventional randomization methods. More affordable commercially available oligonucleotides, coupled with the benefits of "MAX" randomization, mean the cost and complexity should no longer present limitations in the near future.

Immunogenicity is a major factor to consider when developing sdAbs as therapeutic agents. Since the constructed V_L library's CDR diversity is derived from a human template sequence, the V_Ls that comprise the library are fully human and not likely to induce any immunogenic response in patients. In addition, the extensive amount of randomization incorporated into the CDRs should facilitate a broad diversity of conformations and consequently, of epitopes recognized (Collis et al., 2003). Kunkel's method of mutagenesis using synthetic oligonucleotides has been widely used in the successful generation of highly complex phage display libraries (Sidhu et al., 2004; Simon et al., 2004; Scholle et al., 2005). While a large amount of variability was found in the constructed synthetic V_L phage display library, there were instances of wild-type bias due to factors that influence oligonucleotide annealing efficiency, in addition to possible biological censoring throughout the randomized positions. Inconsistencies in titer plates made it difficult to accurately predict library size, but transformation efficiencies obtained suggest an estimate of 10^{10} variants. There is no assay to determine library complexity, and transformation into *E. coli* is the experimental bottleneck limiting complexity. Therefore, it is generally considered that the number of independent clones after primary transformation into *E. coli* and before overnight amplification is the most accurate way to measure the maximal number of different Ab genes. Library size titers could be repeated using the frozen phage aliquots however, the true complexity may not be reflected, as various

factors including different folding rates of different Ab fragments, potential cell toxicity of some Abs, incorrect protein folding and/or assembly, and proteolysis of the displayed Ab, can significantly decrease the actual "functional complexity".

The size of the constructed synthetic V_L library compares favourably with other synthetic Ab repertoires, including Fab libraries with sizes of 4-6.5 x 10^{10} clones (Griffiths et al., 1994; Lee et al., 2004), scFv-based libraries ranging from 3 x 10^8 to 2 x 10^{10} unique members (Pini et al., 1998; Viti et al., 2000; Sidhu et al., 2004; Silacci et al., 2005; Cobaugh et al., 2008), IgNAR repertoires with diversities of approximately 1 x 10^8 (Liu et al., 2007, Shao et al., 2007), and V_H libraries reported to comprise between 2 x 10^8 and 2 x 10^{10} clones (Davies and Riechmann, 1995; Reiter et al., 1999; Jespers et al., 2004; Arbabi-Ghahroudi et al., 2009a; Arbabi-Ghahroudi et al., 2009b). In addition, a sdAb library with a size of ~10^{10} was created by combining CDR3s from heavy and light chains in a V_H-based scaffold (Chen et al., 2010).

3. V_L expression analysis

Small-scale expression is typically performed before proceeding with a large-scale preparation and the use of small expression cultures provides a quick way to judge the effects of varied growth conditions on expression levels and solubility of recombinant proteins. Expression capacities differ between colonies of freshly transformed cells, and small-scale preparations allow for the selection of clones demonstrating optimal expression levels. *E. coli* strains can be grown to high densities in common media such as "Luria Broth" (LB) (Curless et al., 1990; Li et al., 1999), the M9 minimal medium with added supplements (denoted as "M9S" herein) (Marley et al., 2001; Kobe et al., 2008), "Terrific Broth" (TB) (Zanette et al., 1998; Lim et al. 2000) and "Super Broth" (SB) (Madurawe et al., 2000).

In this project, attempts to solubly express periplasmic proteins of mutant V_Ls randomly selected from the library were unsuccessful using the small-scale B2xYT/Amp method; only the wild-type V_L-24 clone seemed to display periplasmic protein expression using the B2xYT/Amp method (Figure 17). It was also determined that in many cases, protein was produced in the cell pellets indicating failure of the protein to localize in the periplasm.

V_Ls seemed to respond more favourably to the conditions of the large scale M9S method as evidenced by the high level of expression for clone W (Figure 18); clones isolated from subtractive panning were also abundantly expressed in M9S with yields ranging from 0.4-4.8 mg of protein per L of culture (Table 3; Figure 22). Similar V_L protein yields have been described elsewhere (Ward et al., 1989; Raffen et al., 1998; Wörn and Plückthun, 1998).

The M9S method employs glucose and casamino acids as carbon sources for bacterial growth (*E. coli* grows at a quicker rate on glucose than on any other carbon source) and supplies essential vitamins and minerals (i.e., $MgCl_2$, $CaCl_2$, and thiamine-HCl) as well as induction medium containing bacto-tryptone and bacto-yeast extract, which improve the protein production capacity of the medium, permit cultivation of auxotrophic bacterial strains, and relieve protease production during synthesis of recombinant protein (Lee and Keasling, 2006). In addition to providing phosphates, which are important for attaining high-cell densities in *E. coli,* the phosphate salts of the M9S medium also serve as a buffer to prevent pH fluctuations that can negativery impact normal metabolic activity (Korz et al. 1995; Lim et al. 2000; Manderson et al., 2005). The presence of glucose helps minimize the substantial burden placed on bacterial cells during the initial growth phase (i.e., prior to IPTG induction) (Donovan et al., 2000).

One study examining expression of the V_L domain of ferritin-binding mouse monoclonal Ab F11, found that V_Ls reached maximum expression levels at 28-32°C (the temperature range used throughout the M9S method of expression) following induction with IPTG; it was also suggested that changes in the culture growth temperature prior to induction affected only the rate at which the bacterial cells would grow (Dubnovitsky et al., 2000). Bacterial culture growth temperature optimization during expression of the ferritin-binding V_L resulted in yields of up to 60 mg per L of bacterial culture, although protein was accumulated in insoluble form as bacterial inclusion bodies (Dubnovitsky et al., 2000) whereas V_Ls described herein were recovered as soluble proteins from the periplasm. It is possible that the V_Ls of this project can be successfully expressed using the LB/Amp or other commonly used expression techniques, however optimization of the temperature and related conditions would likely be needed in order to obtain an adequate amount of protein expression.

4. Subtractive panning for NRP1-specific V_L isolates

Expression cloning, Ab perturbation, and genetic manipulation studies have firmly established a significant role for NRPs in neurodegeneration by mediating Sema3A-induced repulsion of sensory axons and growth cone collapse both *in vitro* and *in vivo* (Kitsukawa et al., 1997; He and Tessier-Lavigne, 1997; Kolodkin et al., 2007; Jiang et al., 2007; Hou et al., 2008); as such, anti-NRP1 sdAbs are logical agents to explore in the treatment of neurodegenerative disorders. NRP1 additionally promotes VEGF-mediated tumor angiogenesis, cell migration, and tumorigenicity (Fu et al., 2000; Whitaker et al., 2001; Roche et al., 2002; Murga et al., 2005).

After the fourth round of subtractive panning, 15 unique V_Ls were isolated and their randomized positions were aligned for statistical analysis with WebLogo (Figure 23). The

relatively high sequence variability observed for the 15 clones reflects the size and diversity of the library from which they were isolated. An overall increase in polar amino acids was found to be accompanied by an increase in basicity, a trend also evident in the pI values of the isolated V_Ls; a large majority of the isolates had (predicted) basic pIs (clones 35, 45, 4, 18, 8, 10, 19, 21, 31, 40, and 41; pI > 8) while only one was acidic (clone 25, pI 5.57) (Table 3). Interestingly, crystal structures suggest that the "b1" loop domain of NRP1 consists primarily of polar residues, with an electronegative portion serving as a site for interaction with Sema3A, which is predominantly basic (Lee et al., 2003; Geretti et al., 2007). Therefore, the enrichment of V_Ls with a predominantly polar and basic nature could be explained by the fact that the clones are able to compete with Sema3A for recognition of the NRP1 "b1" domain.

SEC analysis revealed monomericity for 9 of the 10 V_Ls that demonstrated sufficient expression levels, as evidenced by the sharp, symmetrical peaks formed within the expected elution volume range (~12-15 mL) for a monomer (Figure 24). SEC analysis of the 10 clones provides promising results for the V_L library, indicating that most isolates should be monomeric, making them useful in the development of potential therapeutics.

Flow cytometric analysis was not entirely conclusive, as multiple trials (not shown) showed a mix of results. However, the trial described herein revealed the 4 V_Ls to have a slight increase in affinity for PC12 cells (in comparison to HEK293 cells) through anti-6His-PE detection. Reaction mixtures containing V_Ls and NRP1-positive PC12 cells (Figures 25C-F and 26) displayed an average 2.5-fold increase in MFI over cells mixed with fluorescent label (Figure 25B). In comparison, reaction mixtures containing V_Ls and NRP1-negative HEK293 cells (Figures 25J-M and 26) displayed an average 1.8-fold increase in MFI over cells with fluorescent

label (Figures 25I and 26). While the overall difference in cell binding was not dramatic, the 4 clones yielded subtle increases in MFI for PC12 cell reactions compared to HEK293 cell reactions. However, it remains to be seen if the NRP1 epitope is recognized by the V_L clones. In both cases of PC12 and HEK293 cell reactions, the control V_L revealed a 1.8-fold MFI increase (Figures 25G, 25N, and 26) compared to cells incubated with fluorescent label alone (Figures 25B, 25I and 26).

Ideally, those V_Ls with the greatest affinity to its target should be revealed by panning; in practice, however, these clones may be absent from the selected pool of phage because the theoretical diversity of the library exceeds the actual observed diversity. For example, Kuzmicheva et al. (2009) predicted 1.3×10^{11} possible peptide structures for a randomized 9-mer library, which was ~60-fold greater than the actual number of clones observed in the constructed repertoire. The identification and isolation of irrelevant or non-specific binders is always a possible occurrence in any panning protocol, as propagation and enrichment steps can often result in a bias towards clones with growth and expression advantages in *E. coli*. In such a case, high affinity clones with poor growth characteristics would not survive in later selection rounds. Since a given V_L is genetically coupled to a biological entity (i.e., the phage), the availability of this V_L for selection is dependent upon the growth capacities of its carrier phage. For example, a phage clone may be a moderate binder but yet be selected and propagated by virtue of its better amplification potential. Thus, a higher abundance of lower affinity phage clones would more likely be selected.

Alterations could be made for any future projects concerning the screening of sdAb libraries for NRP1-specific binders. For example, the use of two identical mammalian cell lines

with one transfected to express NRP1 would likely provide for a simple and direct approach avoid any potential ambiguities that more commonly arises from the use of two cell lines derived from closely related but different species. However, this could prove to be a challenge due to the large size of the NRP1 gene (932 residues) which would likely be a limiting factor in successful and efficient gene transfection. Some of the unfavourable outcomes that may arise from transfection include unexpected and atypical morphologies, as well as target cell irregularities not previously observed, and possible cell mortality (Hardie, 1999). Nevertheless, gene transfection has resulted in the isolation of binders to numerous targets via phage display; examples include peptide ligands to the urokinase receptor, C3a receptor-specific scFvs, peptides directed against MARCO (macrophage receptor with collagenous structure), and peptide binders to the IgG receptor FcγRI (Goodson et al., 1994; Hawlisch et al., 1998; Chen et al., 2006; Berntzen et al., 2006). Another alternative approach to the subtractive panning used here would be the use of a NRP1-derived peptide in competitively eluting those V_Ls displaying affinity to the NRP1 epitope on the surface of the cells. The Sema3A-derived peptide that was used in this project eluted those V_Ls with affinity for the cell-surface displayed NRP1 epitope, by interacting with the b1/b2 domains of NRP1. In contrast, the NRP1-derived peptide would theoretically retain in solution those V_Ls that recognize and bind the NRP1 epitope on the surface of PC12 cells, by direct association.

A number of studies have been successful in isolating binders to NRP1 for the purpose of investigating either Sema3A- or VEGFR-mediated effects. A synthetic phage Fab library with an estimated 10^{10} variants that utilized a single human framework containing human consensus CDRs, was reported to yield two binders that cross-react with human and murine NRP1 in murine tumor models (Liang et al., 2007). Similar to this project's mutagenesis design, Liang et

al. introduced library diversity at select CDR positions via synthetic degenerate and trinucleotide codons that mimic natural human Abs. One clone was found to bind the CUB domains (a1/a2) of NRP1 leading to disruption of Sema3A induced neuron collapse; another isolate was found to recognize the coagulation factor V/VIII domains (b1/b2) of NRP1, inhibiting tumor growth in animal xenograft models and blocking VEGF induced cell migration. Giordano et al. obtained cyclic peptide dual binders to VEGFR and NRP1 after three rounds of subtractive panning against VEGFR-expressing human umbilical vein endothelial cells (HUVECs) and VEGFR-negative HUVECs (Giordano et al., 2001; Giordano et al., 2005). In addition, research groups associated with Ark Therapeutics (London, United Kingdom) have been actively investigating NRP1 as a target of angiogenesis, tumor growth, and metastatis (Jarvis et al., 2010; Jia et al., 2010). The same company has recently reported that a small NRP1 antagonist successfully completed preclinical trials in a murine model of lung cancer. The antagonist, which exhibits low nanomolar affinity for NRP1, is believed to directly impair tumour cell growth, inhibit the development of new blood vessels that stimulate tumor cell growth, and reduce cancer cell mobility, without any signs of toxicity.

VIII. CONCLUSIONS AND FUTURE DIRECTIONS

The data presented herein suggest the construction of a complex synthetic V_L phage display library and the isolation of potential V_L binders to NRP1. However, further studies are needed to elucidate the ultimate roles of the isolated clones in neuroprotection and recovery from neurodegenerative diseases.

Selection from the V_L repertoire was accomplished by subtractively panning the library against a cell line (PC12) endogenously expressing NRP1, and a cell line (HEK293) lacking NRP1. NRP1-specific clones were recovered by competitive elution using a synthetic Sema3A-derived peptide recognizing the b1/b2 domains of NRP1. Fifteen V_L clones were isolated and cloned into the *E. coli* expression vector pSJF2H, their proteins expressed, and purified by IMAC. Out of 10 expressed V_Ls, 9 were determined to be non-aggregating by SEC. Four of the V_L isolates displaying high expression levels were analyzed by flow cytometry which indicated a moderate level of binding to target cells, with overall increases in fluorescent intensities compared to control cells. Additional studies will determine the potential therapeutic use of these V_L sdAbs as agents in recovery from stroke and neuron degeneration.

While flow cytometry data was inconsistent, there was indication of cell binding for four of the isolated V_Ls to the NRP1-expressing cell line. It remains to be seen, however, whether the V_Ls will recognize and bind the NRP1 epitope displayed on the surface of PC12 cells, and if they would effectively prevent Sema3A-mediated neuronal degeneration in ischemic/stroke-induced brain models. To investigate any potential impact on the regeneration of adult cortical neurons and induction of axonal outgrowth, an *in vitro* assay involving mouse cortical tissue blocks and the V_Ls should be carried out. The non-toxic dose would first need to be determined for each of

the V_L proteins, followed by administration of each clone for assessment of any potential prevention of neurite outgrowth in stroke-induced animal models. To further study the functional relevance of these V_L sdAbs, specific binding assays to the NRP1 protein would need to be carried out.

Determination of any affinity and specificity toward the NRP1 epitope (which is endogenously expressed by the mammalian rat adrenal PC12 cell line) would be important in clarifying the potential applications of the isolated V_Ls. This could be accomplished by surface plasmon resonance (SPR) binding assays, which would involve the immobilization of NRP1 on a sensor chip surface for measurements of any association and dissociation rate constants (k_a, and k_d, respectively) on a sensorgram plot. ELISA could also be used to detect any binding by coating a plastic surface with the NRP1 protein, either in free or biotinylated form (the latter requiring an additional surface coating step with streptavidin). Additional *in vitro* assays could be performed to examine the potential role of the isolated V_Ls in disrupting the Sema3A/NRP1 interaction and inducing neuron regeneration following stroke. One such assay involves incubation of mouse cortical tissue blocks (that express NRP1 and Sema3A) with each of the V_L isolates for any potential neuroprotection. V_Ls identified here or in future projects experiments could also be further characterized in terms of their thermal stability, refolding efficiency and protease resistance.

With respect to neurotherapeutic approaches, V_Ls could be administered systemically to target NRP1, as they are unlikely to induce any immunogenic response in patients due to the human origin of the stable V_L framework, and would thus be well tolerated in humans. Enhanced targeting efficacy should be obtainable by increasing their intrinsic affinity and avidity

through chimeric formats of anti-NRP1 V_Ls linked to an Fc domain. V_HH sdAbs (i.e., FC5) that bind brain endothelial cells and transmigrate across the BBB *in vitro* and *in vivo*, were discovered by subtractively panning a llama sdAb phage display library against human lung and brain endothelial cells (Muruganandam et al., 2002; Tanha et al., 2003; Abulrob et al., 2005). A suitable approach in therapeutic delivery would be to cross-link the V_Ls with FC5, which could transmigrate across the blood-brain barrier (BBB) and act as a "nanocarrier" to target the brain *in vivo*. Biodistribution studies have indicated significant localization to the brain and rapid clearance from the kidney and liver for FC5 and its derivatives (Muruganandam et al., 2002; Harris and Chess, 2003; Chen et al., 2004a). Methods aimed at improving the pharmacokinetic and pharmacodynamic properties of BBB-targeting sdAbs (such as "PEGylation", the covalent attachment of polyethylene glycol polymer chains to a drug or therapeutic protein) would help maximize *in vivo* brain targeting efficiency (Harris and Chess, 2003; Chen et al., 2004a). With considerable resistance to proteases and efficient export from the brain via the Fc receptor-mediated efflux system of the BBB (unlike IgGs), sdAbs are a valuable alternative to current strategies designed to target drugs to the brain by employing vesicular transendothelial transport (Zhang and Pardridge, 2001; Schlachetzki et al., 2002).

Highly diverse phage display Ab libraries and in particular, synthetic Ab repertoires based on naturally occurring scaffolds, have proven to be rich sources of specific binders and have made significant contributions toward the discovery and development of novel drug treatments. An enhanced understanding of Ab structure and function, as well as which frameworks are ideal for supporting naïve diversity, and what types of chemical diversity should be introduced at select positions within the CDRs is needed to successfully generate highly functional synthetic libraries. Further *in vitro* optimization strategies and improvements in the

design, construction and screening of synthetic Ab libraries, will certainly extend the capacity of synthetic Abs as therapeutic agents to virtually any conceivable target.

With a rapidly expanding list of sequenced genomes and disease-related gene products, large repertoires such as the synthetic human V_L phage display library described herein will be very useful in generating functional Abs against a wide variety of therapeutically relevant target Ags for biomedical research and pharmaceutical applications. Binders to NRP1 and other targets could be further developed as potential therapeutics, diagnostics and research reagents.

IX. REFERENCES

Abulrob, A., H. Sprong, P. V. Bergen en Henegouwen, and D. Stanimirovic. 2005. "The blood-brain barrier transmigrating single domain antibody: mechanisms of transport and antigenic epitopes in human brain endothelial cells." *Journal of Neurochemistry, 95: 1201-1214.*

Adey, N. B., A. H. Mataragnon, J. E. Rider, J. M. Carter, B. and K. Kay. 1995. "Characterization of phage that bind plastic from phage-displayed random peptide libraries." *Gene, 156: 27–31*.

Akerström, B., and L. Björck. 1989. "Protein L: an immunoglobulin light chain-binding bacterial protein. Characterization of binding and physiochemical properties." *J. Biol. Chem., 264: 19740-19746.*

Anand, N. N., G. Dubuc, J. Phipps, C. R. MacKenzie, J. Sadowska, N. M. Young, D. R. Bundle, and S. A. Narang. 1991. "Synthesis and expression in *Escherichia coli* of cistronic DNA encoding an antibody fragment specific for a *Salmonella* serogroup B Oantigen." *Gene 100, 39-44.*

Andris, J. S., S. R. Abraham, V. Pascual, M. P. Pistillo, S. Mantero, G. B. Ferrara, and J. D. Capra. 1995. "The human antibody repertoire: heavy and light chain variable region gene usage in six alloantibodies specific for human HLA class I and class II alloantigens." *Mol. Immunol., 32(14-15): 1105-22*.

Arbabi-Ghahroudi, M., A. Desmyter, L. Wyns, R. Hamers, and S. Muyldermans. 1997. "Selection and identification of single domain antibody fragments from camel heavy-chain antibodies." *FEBS Lett., 414: 521–526.*

Arbabi-Ghahroudi, M**.**, J. Tanha, and R. MacKenzie. 2005. "Prokaryotic expression of antibodies." *Cancer and Metastasis Reviews, 24: 501–519*.

Arbabi-Ghahroudi, M., J. Tanha, and R. MacKenzie. 2009. "Isolation of Monoclonal Antibody Fragments from Phage Display Libraries." *Methods Mol. Biol., 502: 341-64.*

Arbabi-Ghahroudi, M., R. MacKenzie, and J. Tanha. 2009a. "Selection of non-aggregating VH binders from synthetic VH phage-display libraries." *Methods Mol. Biol., 525: 187-216, xiii.*

Arbabi-Ghahroudi, M., R. To, N. Gaudette, T. Hirama, W. Ding, R. MacKenzie, and J. Tanha. 2009b. "Aggregation-resistant VHs selected by *in vitro* evolution tend to have disulfide-bonded loops and acidic isoelectric points." *Protein Eng. Des Sel., 22(2): 59-66.*

Atwell, J. L., K. A. Breheney, L. J. Lawrence, A. J. McCoy, A. A. Kortt, and P. J. Hudson. 1999. "scFv multimers of the anti-neuraminidase antibody NC10: length of the linker between VH and VL domains dictates precisely the transition between diabodies and triabodies." *Protein Eng., 12: 597-604.*

Azzazy, H. M. E., and E. W Highsmith. 2002. "Phage display technology: clinical applications and recent innovations." *Clinical Biochemistry, 35: 425–445.*

Barbas CF III, Burton DR, Scott JK, Silverman GJ. Appendix 1, in Phage display: a laboratory manual. Cold Spring Harbor Laboratory Press. New York: 2001, A1.1-A1.2.

Beck, F., R. Sommer, E. A. Auerswald, C. Kurz., B. Zink, G. Osterburg, H. Schaller, K. Sugimoto, K. Sugisak, T. Okamoto, and M. Takanami. 1978. "Nucleotide sequence of bacteriophage fd DNA." *Nucleic Acids Research, 5: 4495-4503.*

Beck, H., T. Acker, A. W. Puschel, H. Fujisawa, P. Carmeliet, and K. H. Plate. 2002. "Cell type-specific expression of neuropilins in an MCA-occlusion model in mice suggests a potential role in post-ischemic brain remodeling." *J. Neuropathol. Exp. Neurol., 61: 339-350.*

Bell, A., Z. J. Wang, M. A. Ghahroudi, T. A. Chang, Y. Durocher, U. Trojahn, J. Baardsnes, M. L. Jaramillo, S. Li, T. N. Baral, M. O'Connor-McCourt, R. MacKenzie, and J. Zhang. 2010. "Differential tumor-targeting abilities of three single-domain antibody formats." *Cancer Letters, 289: 81–90.*

Belaid, Z., F. Hubint, C. Humblet, J. Boniver, B. Nusgens, and M. P. Defresne. 2005. "Differential expression of vascular endothelial growth factor and its receptors in hematopoietic and fatty bone marrow: evidence that neuropilin-1 is produced by fat cells." *Haematologica., 90(3): 400-401.*

Ben-Zvi, A., O. Manor, M. Schachner, A. Yaron, M. Tessier-Lavigne, and O. Behar. 2008. "The Semaphorin Receptor PlexinA3 Mediates Neuronal Apoptosis during Dorsal Root Ganglia Development." *J.Neurosci., 28, 12427-12432.*

Benhar, I. 2001. "Biotechnological applications of phage and cell display." *Biotechnol. Adv., 19(1): 1-33.*

Benhar, I. 2007. "Design of synthetic antibody libraries." *Expert Opin. Biol. Ther., 7(5): 76-779.*

Berntzen, G., O. H. Brekke, S. A. Mousavi, J. T. Andersen, T. E. Michaelsen, T. Berg, I. Sandlie, and V. Lauvrak. 2006. "Characterization of an FcgammaRI-binding peptide selected by phage display." *Protein Eng. Des. Sel., 19(3): 121-128.*

Bessette, P. H., J. J. Rice, and P. S. Daugherty. 2004. "Rapid isolation of high-affinity protein binding peptides using bacterial display." *Protein Eng. Des. Sel., 17(10):731-9.*

Better, M., C. P. Chang, R. R. Robinson, and A.H. Horwitz. 1988. "Escherichia coli secretion of an active chimeric antibody fragment." *Science, 240: 1041–1043.*

Bird, R. E., K. D. Hardman, J. W. Jacobson, S. Johnson, B. M. Kaufman, S. M. Lee, S. H. Lee, G. S. Pope, G. S. Riordan, and M. Whitlow. 1989. "Single-chain antigen-binding proteins." *Science, 242: 423-426.*

Borsi, L., P. Castellani, G. Allemanni, and L. Zardi. 1998. "Preparation of phage antibodies to the ED-A domain of human fibronectin." *Exp. Cell Res. 240: 244–251.*

Bothmann, H., and A. Plückthun. 1998. "Selection for a periplasmic factor improving phage display and functional periplasmic expression." *Nat. Biotechnol., 16: 376–380.*

Bothmann, H., and A. Plückthun. 2000. "The periplasmic *Escherichia coli* peptidylprolyl cis,trans-isomerase FkpA. I. Increased functional expression of antibody fragments with and without cis-prolines. *J. Biol. Chem., 275: 17100–17105*.

Bradbury, A. R., and J. D. Marks. 2004. "Antibodies from phage antibody libraries." *J. Immunol. Methods,* 290: 29-49.

Bradbury, A. R. 2010. "The Use of Phage Display in Neurobiology." *Curr. Protoc. Neurosci.,. Chapter 5, Unit 5.12.*

Braunagel, M. 2003. "Construction of semisynthetic antibody libraries." *Methods Mol. Biol., 207: 123-132.*

Bredesen, D. E.,, R. V. Rao, and P. Mehlen. 2006. "Cell death in the nervous system." *Nature, 443(7113): 796-802.*

Brinkmann, U., B. K. Lee, and I. Pastan. 1993. "Recombinant immunotoxins containing the VH or VL domain of monoclonal antibody B3 fused to Pseudomonas exotoxin." *J. Immunol., 150, 2774-2782.*

Burton, D. R., C. F. Barbas, and M. A. Persson. 1991. "A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals." *Proc. Natl. Acad. Sci., 88(22): 10134-10137*.

Burton, D. R. 1995. "Phage display." *Immunotechnology, 1: 87-94.*

Campbell, D. S., and C. E. Holt. 2003. "Apoptotic pathway and MAPKs differentially regulate chemotropic responses in retinal growth cones." *Neuron, 37: 939-952*.

Cárcamo, J., M. W. Ravera, R. Brissette, O. Dedova, J. R. Beasley, A. Alam-Moghé, C. Wan, A. Blume, and W. Mandecki. 1998. "Unexpected frameshifts from gene to expressed protein in a phage-displayed peptide library." *Proc. Natl. Acad. Sci., 95(19): 11146-51.*

Carmen, S., and L. Jermutus. 2002. "Concepts in antibody phage display." *Briefings in Functional Genomics and Proteomics, 1(2): 189–203*.

Carnemolla, B., D. Neri, P. Castellani, A. Leprini, G. Neri, A. Pini, G. Winter, and L. Zardi. 1996. "Phage antibodies with pan-species recognition of the oncofoetal angiogenesis marker fibronectin ED-B domain." *Int. J. Cancer, 68: 397–405.*

Carter, P., R. F. Kelley, M. L. Rodrigues, B. Snedecor, M. Covarrubias, M. D. Velligan, W. L. Wong, A. M. Rowland, C. E. Kotts, and M. E. Carver. 1992. "High level Escherichia coli expression and production of a bivalent humanized antibody fragment." *Biotechnology, 10: 163–167*.

Charbit, A., A. Molla, W. Saurin, and M. Hofnung**.** 1988. "Versatility of a vector for expressing foreign polypeptides at the surface of gram-negative bacteria." *Gene 70: 181-189*.

Chen, H., A. Chedotal, Z. He, C. S. Goodman, and M. Tessier-Lavigne. 1997. "Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III." *Neuron, 19: 547-559*.

Chen, H., Z. He, and A. Bagri. 1998. "Semaphorin-neuropilin interactions underlying sympathetic axon responses to class III semaphorins." *Neuron, 21(6): 1283–1290*.

Chen, C., B. Snedecor, J. C. Nishihara, J. C. Joly, N. McFarland, D. C. Andersen, J. E Battersby, and K. M. Champion. 2004. "Highlevel accumulation of a recombinant antibody fragment in the periplasm of Escherichia coli requires a triplemutant (degP prc spr) host strain." *Biotechnol. Bioeng., 85: 463–474*.

Chen, X., R. Park, A. H. Shahinian, J. R. Bading, and P. S. Conti. 2004a. "Pharmacokinetics and tumor retention of 125I-labeled RGD peptide are improved by PEGylation." *Nucl. Med. Biol., 31: 11-19.*

Chen, Y., M. Sankala, J.R. Ojala, Y. Sun, A. Tuuttila, D. E. Isenman, K. Tryggvason, and T. Pikkarainen. 2006. "A phage display screen and binding studies with acetylated low density lipoprotein provide evidence for the importance of the scavenger receptor cysteine-rich (SRCR) domain in the ligand-binding function of MARCO." *J. Biol. Chem., 281(18): 12767-12775.*

Chen, W., Z. Zhu, Y. Feng, X. Xiao, and D.S. Dimitrov. 2008. "Construction of a large phage-displayed human antibody domain library with a scaffold based on a newly identified highly soluble, stable heavy chain variable domain." *J. Mol. Biol., 382***:** *779-789.*

Chen, W., Z. Zhua, Y. Fenga, and D. S. Dimitrov. 2010. "A large human domain antibody library combining heavy and light chain CDR3 diversity." *Molecular Immunology, 47: 912-921.*

Chisti, Y. 1998. "Strategies in downstream processing." In: Subramanian, G. (ed.) Bioseparation and bioprocessing: a handbook, vol 2. *Wiley-VCH, New York: pp 3-30*.

Choi, G. H., D. H. Lee, W. K. Min, Y. J. Cho, D. H. Kweon, D. H. Son, K. Park, and J. H. Seo. 2004. "Cloning, expression, and characterization of single-chain variable fragment antibody against mycotoxin deoxynivalenol in recombinant Escherichia coli." *Protein Expr. Purif., 35(1): 84-92.*

Chothia, C., and A. M. Lesk. 1987. "Canonical structures for the hypervariable regions of immunoglobulins." *J. Mol. Biol., 196(4): 901-917.*

Chothia, C., A. M. Lesk, A. Tramontano, M. Levitt, S. J. Smith-Gill, and G. Air. 1989. "Conformations of immunoglobulin hypervariable regions." *Nature, 342: 877–883.*

Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter. 1991. "Making antibody fragments using phage display libraries." *Nature, 352(6336): 624-628.*

Click, E. M., and R. E. Webster. 1997. "Filamentous phage infection: required interactions with the ToIA protein." *J. Bacteriol., 179: 6464 6471.*

Click, E. M., and R. E. Webster. 1998. "The TolQRA proteins are required for membrane insertion of the major capsid protein of the filamentous phage f1 during infection." *J. Bacteriol., 180: 1723–1728.*

Cobaugh, C. W., J. C. Almagro, M. Pogson, B. Iverson, and G. Georgiou. 2008. "Synthetic Antibody Libraries Focused Towards Peptide Ligands." *J Mol Biol. 378(3): 622-633.*

Cohen, T., Z. Gluzman-Poltorak, A. Brodzky, V. Meytal, E. Sabo, I. Misselevich, M. Hassoun, J. H. Boss, M. Resnick, D. Shneyvas, S. Eldar, and G. Neufeld. 2001. "Neuroendocrine cells along the digestive tract express neuropilin-2." *Biochem. Biophys. Res. Commun., 284(2): 395-403.*

Colcher, D., R. Bird, M. Roselli, K. D. Hardman, S. Johnson, S. Pope, S. W. Dodd, M. W. Pantolianio, D. Milenic, and J. Schlom. 1990. "In vivo tumor targeting of a recombinant single-chain antigen binding protein." *J. Natl Cancer Inst. 82, 1191-1197.*

Collis, A. V. J., A. P. Brouwer, and A. C. R. Martin. 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.*

O'Connor, C. 2008. "Meiosis, genetic recombination, and sexual reproduction." *Nature, Education 1(1).*

Cook, G. P., and I. M. Tomlinson. 1995. "The human immunoglobulin VH repertoire." *Immunol. Today, 16(5): 237-42.*

Corchero, J. L., R. Cubarsi, P. Vila, A. Aris, and A. Villaverde. 2001. "Cell lysis in Escherichia coli cultures stimulates growth and biosynthesis of recombinant proteins in surviving cells." Microbiol. Res., 156: 13-18.

Cortez-Retamozo, V., M. Lauwereys, G. Hassanzadeh, M. Gobert, K. Conrath, S. Muyldermans, P. De Baetselier, and H. Revets. "Efficient tumor targeting by single domain antibody fragments of camels." *Int. J. Cancer, 98: 456–462.*

Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner. 2004. "WebLogo: a sequence logo generator." *Genome Res., 14: 1188–1190.*

Curless, C., J. Pope, and L. Tsai. 1990. "Effect of preinduction specific growth rate on recombinant alpha consensus interferon synthesis in Escherichia coli." *Biotechnol. Prog., 6: 149-152.*

Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. 1990. "Peptides on phage: A vast library of peptides for identifying ligands." *Proc. Natl. Acad. Sci., 87: 6378-6382*.

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

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

Davies, J., and L., Riechmann. 1994. "'Camelising' human antibody fragments: NMR studies on VH domains." *FEBS Lett. Feb 21;339(3):285-90.*

Davies, J., and L. Riechmann. 1995. "Antibody VH domains as small recognition units." *Biotechnology (N. Y.), 13: 475-479.*

Davies, D. R., and G. H. Cohen. 1996. "Interactions of protein antigens with antibodies." *Proc. Natl. Acad. Sci., 93: 7-12*.

de Bruin, R., K. Spelt, J. Mol, R. Koes, and F. Quattrocchio. 1999. "Selection of high-affinity phage antibodies from phage display libraries." *Nat. Biotechnol., 17(4): 397-9*.

de Haardt, H. J., N. van Neer, A. Reurst, S. E. Hufton, R. C. Roovers, P. Henderikx, A. P. de Bruine, J.-W. Arends, and H. R. Hoogenboom. 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 Marco, A. 2011. "Biotechnological applications of recombinant single-domain antibody fragments." *Microbial Cell Factories, 10:44.*

Deng, L. W., P. Malik, and R. N. Perham. 1999. "Interaction of the globular domain pIII and the F-pilus of Escherichia coli." *Virology, 253: 271-277*.

Derda, R., S. K. Tang, S. C. Li, S. Ng, W. Matochko, and M. R. Jafari. 2011. **"**Diversity of phage-displayed libraries of peptides during panning and amplification.**"** *Molecules, 16(2): 1776-1803.*

Desmyter, A., T. R. Transue, M. A. Ghahroudi, M. H. Thi, F. Poortmans, R. Hamers, S. Muyldermans L. Wyns. 1996. "Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme." *Nat. Struct. Biol., 3(9): 803-11.*

Desmyter, A., S. Spinelli, F. Payan, M. Lauwereys, L. Wyn, S. Muyldermans, and C. J. Cambillau. 2002. "Three camelid V_HH domains in complex with porcine pancreatic alpha-amylase: Inhibition and versatility of binding topology." *J. Biol. Chem., 277(26): 23645-23650.*

De Winter, F., A. J. Holtmaat, and J. Verhaagen. 2002. "Neuropilin and class 3 semaphorins in nervous system regeneration." *Adv. Exp. Med. Biol., 515: 115-139.*

De Winter, F., M. Oudega, A. J. Lankhorst, F. P. Hamers, B. Blits, and M. J. Ruitenberg. 2002. "Injury-induced class 3 semaphorin expression in the rat spinal cord." *Exp. Neurol., 175: 61-75.*

Dimitrov, D. S. 2009. "Engineered CH2 domains (nanoantibodies)." *mAbs, 1: 26-28.*

Dimitrov, D. S., and J. D. Marks. 2009. "Therapeutic antibodies: current state and future trends - is a paradigm change coming soon?" *Methods Mol. Biol., 525: 1–27.*

Dower, W. J., J.F. Miller, and C.W. Ragsdale. 1988. "High efficiency transformation of *E.coli* by high voltage electroporation." *Nucleic Acids Res., 16: 6127-6145.*

Dubnovitsky, A. P., Z. I. Kravchuk, A. A. Chumanevich, A. Cozzi, P. Arosio, and S. P. Martsev. 2000. "Expression, Refolding, and Ferritin-Binding Activity of the Isolated VL-Domain of Monoclonal Antibody F11." *Biochemistry (Moscow), 65(9): 1011-1018.*

Duenas, M., M. Ayala, J. Vazquez, M. Ohlin, E. Soderlind, C. A. Borrebaeck, and J. V. Gavilondo. 1995. "A point mutation in a murine immunoglobulin V-region strongly influences the antibody yield in *Escherichia coli*." *Gene, 158: 61–66.*

Dumoulin, M., K. Conrath, A. V. Meirhaeghe, F. Meersman, K. Heremans, and L. G. Frenken. 2002. "Single-domain antibody fragments with high conformational stability." *Protein Sci., 11: 500–515.*

Dzionek, A., A. Fuchs, P. Schmidt, S. Cremer, M. Zysk, S. Miltenyi, D. W. Buck, and J. Schmitz. 2000. "BDCA-2, BDCA-3, and BDCA-4: three markers for distinct subsets of dendritic cells in human peripheral blood." *J. Immunol., 165(11):6037-6046.*

Dzionek, A., Y. Inagaki, K. Okawa, J. Nagafune, J. Röck, Y. Sohma, G. Winkels, M. Zysk, Y. Yamaguchi, and J. Schmitz. 2002. "Plasmacytoid dendritic cells: from specific surface markers to specific cellular functions." *Hum. Immunol., 63(12): 1133-1148.*

Ellis, L. M. 2006. "The role of neuropilins in cancer." *Mol. Cancer Ther., 5(5): 1099-1107.*

Endemann, H., and P. Model. 1995. "Location of filamentous phage minor coat proteins in phage and in infected cells." *J. Mol. Biol., 250: 496–506*.

Ewert, S., T. Huber, A. Honegger, and A. Plückthun. 2003. "Biophysical Properties of Human Antibody Variable Domains." J. Mol. Biol., 325(3): 531-553.

Felici, F., L. Castagnoli, A. Musacchio, R. Jappelli, G. Cesareni. 1991. "Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector." *J. Mol. Biol., 222: 301-310*.

Fellouse, F. A., K. Esaki, S. Birtalan, D. Raptis, V. J. Cancasci, A. Koide, P. Jhurani, M. Vasser, C. Wiesmann, A. A. Kossiakoff, S. Koide, and S. S. Sidhu. 2007. "High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries." *J. Mol. Biol., 373(4): 924-40.*

Fiedler, S., and R. Wirth. 1988. "Transformation of bacteria with plasmid DNA by electroporation." *Anal. Biochem. 170: 38-44.*

Filpula, D. 2007. "Antibody engineering and modification technologies." *Review Biomolec. Eng., 24: 201-215.*

Finnern, R., J. M. Bye, K. M. Dolman, M. H. Zhao, A. Short J. D. Marks, M. C. Lockwood, and W. H. Ouwehand. 1995. "Molecular characteristics of anti-self antibody fragments against neutrophil cytoplasmic antigens from human V gene phage display libraries." *Clin. Exp. Immunol., 102: 566–574.*

Forrer, P., S. Jung, and A. Plückthun. 1999. "Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins." *Curr. Opin. Struct. Biol., 9(4): 514-20*.

Frenken, L. G., R. H. van der Linden, P. W. Hermans, J. W. Bos, R. C. Ruuls, B. de Geus, C. T. Verrips. 2000. "Isolation of antigen specific llama VHH antibody fragments and their high level secretion by *Saccharomyces cerevisiae*." *J Biotechnol. 78(1): 11-21.*

Fuh, G., K. C. Garcia, and A. M. de Vos. 2000. "The interaction of neuropilin-1 with vascular endothelial growth factor and its receptor Flt-1." *J. Biol. Chem., 275: 26690-26695*.

Fuh, G. 2007. "Synthetic antibodies as therapeutics." *Expert Opin. Biol. Ther., 7(1): 73-87.*

Fujisawa, H., T. Kitsukawa, A. Kawakami, S. Takagi, M. Shimizu, and T. Hirata. 1997. "Roles of a neuronal cell-surface molecule, neuropilin, in nerve fiber fasciculation and guidance." *Cell Tissue Res., 290: 465–470.*

Fujita, H., B. Zhang, K. Sato, J. Tanaka, and M. Sakanaka. 2001. "Expressions of neuropilin-1, neuropilin-2 and semaphorin 3A mRNA in the rat brain after middle cerebral artery occlusion." *Brain Res., 914: 1-14*.

Furumatsu, T., Z. N. Shen, A. Kawai, K. Nishida, H. Manabe, T. Oohashi, H. Inoue, and Y. Ninomiya. 2003. "Vascular endothelial growth factor principally acts as the main angiogenic factor in the early stage of human osteoblastogenesis." *J. Biochem. (Tokyo), 133: 633–639.*

Gagliardini, V., and C. Fankhauser. 1999. "Semaphorin III can induce death in sensory neurons." *Mol. Cell Neurosci. 14, 301-316*.

Gál, P., J. Dobó, P. Závodszky, and R. B. M. Sim. 2009. "Early complement proteases: C1r, C1s and MASPs. A structural insight into activation and functions." *Mol. Immun., 46(14): 2745-2752.*

Gao, C., S. Mao, G. F. Kaufmann, P. Wirsching, R. A. Lerner, and K. D. Janda. 1999. "Making artificial antibodies: A format for phage display of combinatorial heterodimeric arrays." *Proc. Natl. Acad. Sci*., *96: 6025-6030.*

Gavish, M., R. A. Dwek, and D. Givol. 1977. "Comparison of the fine specificity of anti-dinitrophenyl-combining site composed of either VL dimer or VL and VH of protein 315." *Biochemistry, 16(14): 3154-3159.*

Genst, E. D., D. Saerens, S. Muyldermans, K. Conrath. 2006. "Antibody repertoire development in camelids." *Developmental and Comparative Immunology, 30: 187–198*.

Geretti, E., and M. Klagsbrun. 2007. "Neuropilins: Novel Targets for Anti-Angiogenesis Therapies." *Cell Adh. Migr., 1(2): 56–61.*

Geretti, E., A. Shimizu, P. Kurschat, and M. Klagsbrun. 2007. "Site-directed mutagenesis in the B-neuropilin-2 domain selectively enhances its affinity to VEGF165, but not to semaphorin 3F." *J. Biol. Chem., 282(35): 25698-25707.*

Giger, R. J., E. R. Urquhart, S. K. H. Gillespie, D. V. Levengood, D. D. Ginty, and A. L. Kolodkin. 1998. "Neuropilin-2 is a receptor for semaphorin IV: insight into the structural basis of receptor function and specificity." *Neuron 21, 1079–1092.*

Giordano, R. J., M. Cardo-Vila, J. Lahdenranta, R. Pasqualini, and W. Arap. 2001. "Biopanning and rapid analysis of selective interactive ligands." *Nat. Med., 7: 1249-1253.*

Giordano, R. J., C. D. Anobom, M. Cardo-Vila, J.Kalil, A. P. Valente, R. Pasqualini, F. C. Almeida, and W. Arap. 2005. "Structural basis for the interaction of a vascular endothelial growth factor mimic peptide motif and its corresponding receptors." *Chem. Biol., 12: 1075-1083.*

Givol, D. 1991. "The minimal antigen-binding fragment of antibodies--Fv fragment." *Mol. Immunol.*, *28: 1379-1386.*

Goldman, E.R., G. P. Anderson, J. L. Liu, J. B. Delehanty, L. J. Sherwood, L. E. Osborn, L. B. Cummins, and A. Hayhurst. 2006. "Facile generation of heat-stable antiviral and antitoxin single domain antibodies from a semisynthetic llama library." *Anal Chem., 78(24): 8245-55.*

Goldsby, R., T. J. Kindt, B. A. Osborne, and J. Kuby. 2003. "Antigens (Chapter 3)." Immunology (5th ed.) *W. H. Freeman and Company, New York: pp. 57-75.*

Goodson, R. J., M. V. Doyle, S. E. Kaufman, and S. Rosenberg. 1994. "High-affinity urokinase receptor antagonists identified with bacteriophage peptide display." *Proc. Natl. Acad. Sci., 91(15): 7129-7133.*

Gosling, J. 1995. "Introductory Statistics." *Pascal Press: Australia, pp. 78–79.*

Grandclement, C., and C. Borg. 2011. "Neuropilins: A New Target for Cancer Therapy." *Cancers*, *3: 1899-1928.*

Greenwood, J., A. E. Willis, and R. N. Perham. 1991. "Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from Plasmodium falciparum circumsporozoite protein as antigens." *J. Mol. Biol., 220, 821-827.*

Griffiths, A.D. 1993. "Production of human antibodies using bacteriophage." *Curr. Opin. Immunol., 5(2): 263-267*

Griffiths, A. D., S. C. Williams, O. Hartley, I. M. Tomlinson, P. Waterhouse, W. L. Crosby, R. E. Kontermann, P. T. Jones, N. M. Low, and T. J. Allison. 1994. "Isolation of high affinity human antibodies directly from large synthetic repertoires." *EMBO J., 13(14): 3245-3260.*

Gu, C., B. J. Limberg, G. B. Whitaker, B. Perman, D. J. Leahy, J. S. Rosenbaum, D. D. Ginty, and A. L. Kolodkin. 2002. "Characterization of neuropilin-1 structural features that confer binding to semaphorin 3A and vascular endothelial growth factor 165." *J. Biol. Chem., 277(20): 18069-76.*

Gu, C., E. R. Rodriguez, D. V. Reimert, T. Shu, B. Fritzsch, L. J. Richards, A. L. Kolodkin, and D. D. Ginty. 2003. "Neuropilin-1 conveys semaphorin and VEGF signaling during neural and cardiovascular development." *Dev. Cell, 5: 45-57.*

Hamers-Casterman, C., T. Atarhouch, S. Muyldermans, G. Robinson, C. Hamers, and E. B. Songa. 1993. "Naturally occurring antibodies devoid of light chains." *Nature, 363: 446–448*.

Hanahan, D., J. Jessee, and F. R. Bloom. 1991. "Plasmid transformation of *Escherichia coli* and other bacteria." *Methods Enzymol., 201: 63-113.*

Handa, P., S. Thanedar, and U. Varshney. 2002. "Rapid and reliable site-directed mutagenesis using Kunkel's approach." *Methods Mol Biol. 182: 1-6.*

Hanes, J., and A. Plückthun. 1997. "In vitro selection and evolution of functional proteins by using ribosome display." *Proc. Natl. Acad. Sci., 94(10): 4937-42*.

Hardie, D. G. (Ed.). 1999. Protein Phosphorylation: A Practical Approach (2nd ed.). *Oxford University Press, New York.*

Harding, S. E., E. Longman, B. Carrasco, A. Ortega, and J. Garcia. 2004. "Studying Antibody Conformations by Ultracentrifugation and Hydrodynamic Modeling." In: Antibody Engineering: Methods and Protocols. B. K. C. Lo (ed). *Humana Press: New Jersey, USA, pp. 93-113.*

Harper, S. J., C. Y. Xing, C. Whittle, R. Parry, D. Gillatt, D. Peat, and P. W. Mathieson. 2001. "Expression of neuropilin-1 by human glomerular epithelial cells in vitro and in vivo." *Clin. Sci. (Lond)., 101(4): 439-446.*

Harris, J. M., and R. B. Chess. 2003. "Effect of pegylation on pharmaceuticals." *Nat. Rev. Drug Discov., 2: 214-221.*

Harrison, J. L., S. C. Williams, G. Winter, and A. Nissim. 1996. "Screening of phage antibody libraries." *Methods Enzymol., 267: 83–109.*

Hawlisch, H., R. Frank, M. Hennecke, M. Baensch, B. Sohns, L. Arseniev, W. Bautsch, A. Kola, A. Klos, and J. Köhl. 1998. "Site-directed C3a receptor antibodies from phage display libraries." *J. Immunol., 160(6): 2947-2958.*

Hawkins, R. E., S. J. Russell, and G. Winter. 1992. "Selection of phage antibodies by binding affinity. Mimicking affinity maturation." J. Mol. Biol., 226: 889–96.

Hayashi, N., M. Welschoff, M. Zewe, M. Braunagel, S. Dübel, F. Breitling, and M. Little. 1994. "Simultaneous mutagenesis of antibody CDR regions by overlap extension and PCR." *Biotechniques, 17: 310-316.*

He, Z., and M. Tessier-Lavigne. 1997. "Neuropilin is a receptor for the axonal chemorepellent Semaphorin III." *Cell, 90: 739-751.*

Hellmuth, K., D. J. Korz, E. A. Sanders, and W. D. Deckwer. 1994. "Effect of growth rate on stability and gene expression of recombinant plasmids during continuous and high cell density cultivation of Escherichia coli TG1." *J. Biotechnol., 32: 289-298.*

Holliger, P., and P. J. Hudson. 2005. "Engineered antibody fragments and the rise of the single domains." *Natl. Biotech., 23: 1126-1136.*

Holt, L. J., K. Büssow, G. Walter, and I. M. Tomlinson. 2000. "By-passing selection: direct screening for antibody-antigen interactions using protein arrays." *Nucleic Acids Res., 28(15): E72.*

Holt, L. J., C. Herring, L. S. Jespers, B. P. Woolven, and I. M. Tomlinson. 2003. "Domain antibodies: proteins for therapy." *Trends Biotechnol., 21: 484-490.*

Hoogenboom, H. R. 1997. "Designing and optimizing library selection strategies for generating high-affinity antibodies." *Trends Biotechnol., 15(2): 62-70.*

Hoogenboom, H. R. 2005. "Selecting and screening recombinant antibody libraries." *Nature Biotechnol.,* 23: 1105-1116.

Hou,, S. T., D. Callaghan, M. C. Fournier, I. Hill, L. Kang, B. Massie, P. Morley, C. Murray, I. Rasquinha, R. Slack, and J. P. MacManus. 2000. "The transcription factor E2F1 modulates apoptosis of neurons." *J. Neurochem., 75: 91-100*.

Hou, S. T., S. X. Jiang, and R. A. Smith. 2008. "Permissive and repulsive cues and signalling pathways of axonal outgrowth and regeneration." *Int. Rev. Cell Mol. Biol., 267: 125-181*.

Hou, S. T., A. Keklikian, J. Slinn, M. O'Hare, S. X. Jiang, and A. Aylsworth. 2008. "Sustained up-regulation of semaphorin 3A, Neuropilin1, and doublecortin expression in ischemic mouse brain during long-term recovery." *Biochem. Biophys. Res. Commun. 367: 109-115*.

Hughes, M., D. Nagel, A. Santos, A. Sutherland, and A. Hine. 2003. "Removing the redundancy from randomized gene libraries." *J. Mol. Biol., 331: 973-979.*

Hughes-Jones, N. C., B. D. Gorick, J. M. Bye, R. Finnern M. L. Scott, D. Voak, J. D. Marks, and W. H. Ouwehand. 1994. "Characterization of human blood group scFv antibodies derived from a V gene phage-display library." *Br. J. Haematol., 88: 180–186.*

Hughes-Jones, N. C., J. M. Bye, B. D. Gorick, D. Marks, and W. H. Ouwehand. 1999. "Synthesis of Rh Fv phage-antibodies using V_H and V_L germline genes." *Br. J. Haematol., 105: 811–816.*

Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mees, D. R. Burton, S. J. Benkovic, R. A. Lerner. 1989. "Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda." *Science, 246(4935): 1275-81.*

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

Huston, J. S., M. Mudgett-Hunter, M. S. Tai, W. McCartney, F. Warren, E. Haber, and H. Opperman. 1991. "Protein engineering of single-chain Fv analogs and fusion proteins." *Meth. Enzymol., 203: 46-48.*

Huston, J. S., M. N. Margolies, and E. Haber. 1996. "Antibody binding sites." *Adv. Prot. Chem., 49: 329-450.*

Ilyichev, A. A., O. O. Minenkova, S. I. Tatkov, N. N. Karpyshev, A. M. Eroshkin, V. A. Petrenko, and L. S. Sandakhchiev. 1989. "Construction of M13 viable bacteriophage with the insert of foreign peptides into the major coat protein." *Dokl. Biochem. (Proc. Acad. Sci. USSR)-Engl. Tr., 307: 196–198.*

Jäger, M., and A. Plückthun. 1999. "Folding and assembly of an antibody Fv fragment, a heterodimer stabilized by antigen." *J. Mol. Biol., 285: 2005-2019.*

Jarvis, A., C. K. Allerston, H. Jia, B. Herzog, A. Garza-Garcia, N. Winfield, K. Ellard, R. Aqil, R. Lynch, C. Chapman, B. Hartzoulakis, J. Nally, M. Stewart, L. Cheng, M. Menon, M. Tickner, S. Djordjevic, P. C. Driscoll, I. Zachary, and D. L. Selwood. 2010. "Small Molecule Inhibitors of the Neuropilin-1 Vascular Endothelial Growth Factor A (VEGF-A) Interaction." *J. Med. Chem., 53(5): 2215-2226.*

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

Jemmerson, R. 1987. "Multiple overlapping epitopes in the three antigenic regions of horse cytochrome c1." *J Immunol, 138: 213-219.*

Jespers, L., O. Schon, K. Famm, and G. Winter. 2004. "Aggregation-resistant domain antibodies selected on phage by heat denaturation." *Nat. Biotechnol., 22(9): 1161-1165.*

Jia, H., L. Cheng, M. Tickner, A. Bagherzadeh, D. Selwood, and I. Zachary. 2010. "Neuropilin-1 antagonism in human carcinoma cells inhibits migration and enhances chemosensitivity." *Br. J. Cancer., 102(3): 541-552.*

Jiang, S. X., M. Sheldrick, A. Desbois, J. Slinn, and S. T. Hou. 2007. "Neuropilin-1 Is a Direct Target of the Transcription Factor E2F1 during Cerebral Ischemia-Induced Neuronal Death In Vivo." *Mol. Cell Biol. 27: 1696-1705.*

Jiang, S.X., S. Whitehead, A. Aylsworth, J. Slinn, B. Zurakowsk, K. Chan, J. Li, and S. T. Hou. 2010. "Neuropilin-1 Directly Interacts with Fer Kinase to Mediate Semaphorin3A-induced Death of Cortical Neurons." *J. Biol. Chem., 285(13): 9908-18*.

He, J., G. Zhou, K. D. Liu, and X. Y. Qin. 2002. "Construction and Preliminary Screening of a Human Phage Single-Chain Antibody Library Associated with Gastric Cancer." *Journal of Surgical Research, 102: 150-155*.

Jönsson, U., L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Löfas, B. Persson, H. Roos, I. Rönnberg, S. Sjölander, E. Stenberg, R. Stahlberg, C. Urbaniczky, H. Ostlin, and M. Malmqvist. 1991. "Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology." *BioTechniques, 11: 620.*

Kabat, E. A. 1982. "Antibody diversity versus antibody complementarity." *Pharmacol. Rev., 34: 23-38*.

Kabat, E.A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. "Sequences of Proteins of Immunological Interest." *US Department of Health and Human Services, US Public Health Service, Bethesda, MD.*

Kaneko, S., A. Iwanami, M. Nakamura, A. Kishino, K. Kikuchi, S. Shibata, H. J. Okano, T. Ikegami, A. Moriya, O. Konishi, C. Nakayama, K. Kumagai, T. Kimura, Y. Sato, Y. Goshima, M. Taniguchi, M. Ito, Z. He, Y. Toyama, and H. Okano. 2004. "Axonal Regeneration and Functional Recovery by Administration of Strong and Selective Semaphorin3A Inhibitor into the Injured Spinal Cord." *Nat. Med., 12: 1380-1389*.

Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1991. "Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces." *Proc. Natl. Acad. Sci., 88: 4363–6*.

Karjalainen, K., D. E. Jaalouk, C. E. Bueso-Ramos, A. J. Zurita, A. Kuniyasu, B. L. Eckhardt, F. C. Marini, B. Lichtiger, S. O'Brien, H. M. Kantarjian, J.E. Cortes, E. Koivunen, W. Arap, and R. Pasqualini. 2011. "Targeting neuropilin-1 in human leukemia and lymphoma.**"** *Blood, 117(3): 920-927.*

Kawasaki, T., T. Kitsukawa, Y. Bekku, Y. Matsuda, M. Sanbo, T. Yagi, and H. Fujisawa. 1999. "A requirement for neuropilin-1 in embryonic vessel formation." *Development., 126(21): 4895-4902.*

Kay, B. K., M. P. Williamson, and M. Sudol. 2000. "The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains." *FASEB J., 14(2): 231-41.*

Kehoe, J. W., and B. K. Kay. 2005. "Filamentous phage display in the new millennium." *Chem. Rev., 105(11): 4056-72.*

Kelley, L. L. C., and C. Momany. 2003. "Generation of a phagemid mouse recombinant antibody fragment library by multisite-directed mutagenesis." *BioTechniques, 35: 750-758.*

Kini, R. M., and H. J. Evans. "A hypothetical structural role for proline residues in the flanking segments of protein-protein interaction sites." *Biochem. Biophys. Res. Commun., 212(3): 1115-1124*.

Kitsukawa, T., M. Shimizu, and M. Sanbo. 1997. "Neuropilin-semaphorin III/D-mediated chemorepulsive signals play a crucial role in peripheral nerve projection in mice." *Neuron, 19(5): 995-1005*.

Klagsbrun, M., and P. A. D'Amore. 1996. "Vascular endothelial growth factor and its receptors." *Cytokine Growth Factor Rev., 7(3): 259-270.*

Klagsbrun, M., and A. Eichmann. 2005. "A role for axon guidance receptors and ligands in blood vessel development and tumor angiogenesis." *Cytokine Growth Factor Rev., 16: 535-548*.

Klostermann, A., M. Lohrum, and R. H. Adams. 1998. "The chemorepulsive activity of the axonal guidance signal semaphorin D requires dimerization." *J Biol Chem., 273(13): 7326–7331*.

Knappik, A., L. Ge, A. Honegger, P. Pack, M. Fischer, G. Wellnhofer, A. Hoess, J. Wölle, A. Plückthun, and B. Virnekäs. 2000. "Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides." J. Mol. Biol., 296(1): 57-86.

Kobe, B., M. Guss, and T. Huber. 2008. "Structural proteomics: high-throughput methods (Preface)." *Methods Mol Biol., 426: v-vi.*

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

Kolodkin, A. L., D. V. Levengood, E. G. Rowe, Y. T. Tai, R. J. Giger, and D. D. Ginty. 1997. "Neuropilin is a semaphorin III receptor." *Cell, 90: 753-762*.

Krebs B, Rauchenberger R, Reiffert S, Rothe C, Tesar M, Thomassen E, Cao M, Dreier T, Fischer D, Höss A, Inge L, Knappik A, Marget M, Pack P, Meng XQ, Schier R, Söhlemann P, Winter J, Wölle J, Kretzschmar T. 2001. "High-throughput generation and engineering of recombinant human antibodies." *J. Immunol. Methods, 254(1-2): 67-84*.

Krumpe, L. R., K. M. Schumacher, J. B. McMahon, L. Makowski, and T. Mori. 2007. "Trinucleotide cassettes increase diversity of T7 phage-displayed peptide library." *BMC Biotechnol., 7: 65-73.*

Krykbaev, R. A., P. Tsantili, P. D. Jeffrey, and M. N. Margolies. 2002. "Modifying specificity of antidigoxin antibodies using insertional mutagenesis." *Protein Sci., 11(12): 2899-2908.*

Kunkel, T. A. 1985. "Rapid and efficient site-specific mutagenesis without phenotypic selection." *Proc. Natl. Acad. Sci., 82(2): 488-92.*

Kunkel, T. A., J. D. Roberts, and R. A. Zakour. 1987. "Rapid and efficient site-specific mutagenesis without phenotypic selection." *Meth. Enzymol., 154: 367.*

Kupsch, J. M., N. H. Tidman, N. V. Kang, et al. 1999. "Isolation of human tumor-specific antibodies by selection of an antibody phage library on melanoma cells." Clin. Cancer Res., 5: 925–31.

Kuzmicheva, G. A., P. K. Jayanna, I. B. Sorokulova, and V. A. Petrenko. 2009. "Diversity and censoring of landscape phage libraries." *Protein Engineering, Design & Selection, 22(1): 9-18.*

Kweon, D. H., N. S. Han, K. M. Park, and J. H. Seo. 2001. "Overproduction of Phytolacca insularis protein in batch and fed-batch culture of recombinant Escherichia coli." *Proc. Biochem., 36: 537-542.*

Ladenson, R. C., D. L. Crimmins, Y. Landt, and J. H. Ladenson. 2006. "Isolation and characterization of a thermally stable recombinant anti-caffeine heavy-chain antibody fragment." *Anal Chem., 78(13): 4501-8.*

Laemmli, U.K. 1970. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature (London) 227, 680-685.*

Lee, J., and S. Shin. 2001. "Understanding beta-hairpin formation by molecular dynamics simulations of unfolding." *Biophys. J., 81(5): 2507-16.*

Lee, S., J. H. Kim, C. S. Lee, J. H. Kim, Y. Kim, K. Heo, Y. Ihara, Y. Goshima, P. G. Suh, and S. H. Ryu. 2002. "Collapsin Response Mediator Protein-2 Inhibits Neuronal Phospholipase D2 Activity by Direct Interaction." *Journal of Bio. Chem., 277(8): 6542–6549.*

Lee, C. C., A. Kreusch, D. McMullan, K. Ng, and G. Spraggon. 2003. "Crystal structure of the human neuropilin-1 b1 domain." *Structure, 11(1): 99-108.*

Lee, C. V., W.-C. Liang, M. S. Dennis, C. Eigenbrot, S. S. Sidhu, and G. Fuh. 2004. "High-affinity Human Antibodies from Phage-displayed Synthetic Fab Libraries with a Single Framework Scaffold." *J. Mol. Biol., 340: 1073-1093.*

Lee, S. K., and J. D. Keasling. 2006. "Effect of Glucose or Glycerol as the Sole Carbon Source on Gene Expression from the *Salmonella prpBCDE* Promoter in *Escherichia coli." Biotechnol. Prog., 22: 1547-1551.*

Lee, C. M. Y., N. Iorno, F. Sierro, and D. Christ. 2007. "Selection of human antibody fragments by phage display." *Nature Protocols, 2(11): 3001-3008*.

Li, N., L. J. Qu, Y. Liu, Q. Li, H. Gu, and Z. Chen. 1999. "The refolding, purification, and activity analysis of a rice Bowman-Birk inhibitor expressed in Escherichia coli." *Protein Express. Pur., 15: 99-104.*

Li, M., 2000. "Applications of display technology in protein analysis." *Nat. Biotechnol., 18(12): 1251-6.*

Li, B., H. Zhao, Q. Liu, R. Murali, M. I. Greene, and H. Zhang. 2002. "Deoxycholate-based Method to screen phage display clones for uninterrupted open reading frames." *BioTechniques, 33(2): 294-296.*

Liang, W.-C., M. S. Dennis, S. Stawicki, Y. Chanthery, Q. Pan, Y. Chen, C. Eigenbrot, J. Yin, A.W. Koch, X. Wu, N. Ferrara, A. Bagri, M. Tessier-Lavigne, R.J. Watts, and Y. Wu. 2007. "Function Blocking Antibodies to Neuropilin-1 Generated from a Designed Human Synthetic Antibody Phage Library." *J. Mol. Biol., 366: 815–829*.

Lim, H. K, K. H. Jung, D. H. Park, and S. I. Chung. 2000. "Production characteristics of interferon-a using an l-arabinose promoter system in a high-cell-density culture." *Appl. Microbiol. Biotechnol., 53: 201-208.*

Lindner, P., et al. 1992. "Purification of Native Proteins from the Cytoplasm and Periplasm of *Escherichia coli* Using IMAC and Histidine Tails: A Comparison of Proteins and Protocols." *METHODS: A Companion to Methods in Enzymology, 4: 41–56*.

Liu, J. L., G. P. Anderson, and E. R. Goldman. 2007. "Isolation of anti-toxin single domain antibodies from a semi-synthetic spiny dogfish shark display library." *BMC Biotechnol., 7: 78-88.*

Liu, C. C., A. V. Mack, E. M. Brustad, J. H. Mills, D. Groff, V. V. Smider, and P. G. Schultz. 2009. "Evolution of proteins with genetically encoded "chemical warheads"." *J. Am. Chem. Soc., 131(28): 9616-9617.*

Lowman, B. H. 2007. "Bacteriophage Display and Discovery of Peptide Leads for Drug Development." Annu. *Rev. Biophys. Biomol. Struct., 26: 401–24*.

Lu, Z., K. S. Murray, V. Van Cleave, E. R. LaVallie, M. L. Stahl, and J. M. McCoy**.** 1995. "Expression of thioredoxin random peptide libraries on the *Escherichia coli* cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions." *Bio/Technology, 13: 366-372.*

Lubkowski, J., F. Hennecke, A. Plückthun, and A. Wlodawer. 1999. "Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA." *Structure. 7(6): 711-22*.

MacKenzie, C.R., V. Sharma, D. Brummell, D. Bilous, G. Dubuc, J. Sadowska, N. M. Young, D. R. Bundle, and S. A. Narang. 1994. "Effect of C lambda-C kappa domain switching on Fab activity and yield in *Escherichia coli*: synthesis and expression of genes encoding two anti-carbohydrate Fabs." *Biotechnology 12, 390-395.*

MacKenzie, R., and R. To. 1998. "The role of valency in the selection of anti-carbohydrate single-chain Fvs from phage display libraries." *J. Immunol. Methods, 220: 39.*

MacManus, J. P., M. Jian, E. Preston, I. Rasquinha, J. Webster, and B. Zurakowski. 2003. "Absence of the transcription factor E2F1 attenuates brain injury and improves behavior after focal ischemia in mice." *J. Cereb. Blood Flow Metab., 23: 1020-1028.*

Madurawe, R. D., T.E. Chase, E. I. Tsao, and W. E. Bentley. 2000. "A recombinant lipoprotein antigen against Lyme disease expressed in E. coli: fermentor operating strategies for improved yield." *Biotechnol. Progress, 16: 571-576.*

Makrides, S. C. 1996. "Strategies for achieving high-level expression of genes in Escherichia coli." *Microbiol. Rev., 60: 512-538.*

Malmborg, A. C., M. Duenas, M. Ohlin, E. Soderlind, and C. A. Borrebaeck. 1996. "Selection of binders from phage displayed antibody libraries using the BIAcore biosensor." J. Immunol. Methods, 198: 51–7.

Manderson, D., R. Dempster, and Y. Chisti. 2006. "A recombinant vaccine against hydatidosis: production of the antigen in Escherichia coli." *J. Ind. Microbiol. Biotechnol., 33(3): 173-82.*

Mao, H., J. J. Graziano, T. M. Chase, C. A. Bentley, O. A. Bazirgan, N. P. Reddy, B. D. Song, and V. V. Smider. "Spatially addressed combinatorial protein libraries for recombinant antibody discovery and optimization." *Nat. Biotechnol., 28(11): 1195-1202.*

Marcus, W. D., L. M. Stuart, and M. R. Sierks. 2006. "Identification and repair of positive binding antibodies containing randomly generated amber codons from synthetic phage display libraries." *Biotechnol. Prog. , 22: 919-922.*

Marks, J.D., H. R. Hoogenboom, and T. P. Bonnert. 1991. "By-passing immunization. Human antibodies from V-gene libraries displayed on phage." *J. Mol. Biol., 222(3): 581-597.*

Marks, J. D., W. H. Ouwehand, J. M. Bye, R. Finnern, B. D. Gorick, D. Voak, S. J. Thorpe, N. C. Hughes Jones, and G. Winter. 1993. "Human antibody fragments specific for human blood group antigens from a phage display library." *Biotechnology, 11: 1145–1149.*

Marks, J. D. 1995. Human monoclonal antibodies from V-gene repertoires expressed on bacteriophage, in: C.A.K. Borrebaeck (Ed.) *Antibody Engineering, Oxford University Press, New York: pp. 53-88.*

Marks, J. D. 2004. "Antibody affinity maturation by chain shuffling." *Methods Mol Biol*, 248:327-343.

Marley, J., M. Lu, and C. Bracken. 2001. "A method for efficient isotopic labeling of recombinant proteins." *J. Biomol. NMR, 20(1): 71-75.*

Marroni, M., K. M. Kight, M. Hossain, L. Cucullo, S. Y. Desai, and D. Janigro. 2003. "Dynamic in vitro model of the blood-brain barrier. Gene profiling using cDNA microarray analysis." *Methods Mol. Med.;89: 419-434.*

Martineau, P. 2010. Antibody Engineering Volume 1. Kontermann, R. E., and S. Dübel (Eds.) "Chapter 6: Synthetic Antibody Libraries." *Springer-Verlag: Berlin, Germany, pp. 85-99.*

Martsev, S. P., Z. I. Kravchuk, A. A. Chumanevich, A. P. Vlasov, A. P. Dubnovitsky, I. A. Bespalov, P. Arosio, and S. M. Deyev. 1998. "Antiferritin single-chain antibody: a functional protein with incomplete folding?" *FEBS Lett., 441(3): 458-462.*

Martsev, S. P., A. A. Chumanevich, A. P. Vlasov, A.P. Dubnovitsky, Y. I. Tsybovsky, S. M. Deyev, A. Cozzi, P. Arosio, and Z. I. Kravchuk. 2000. "Antiferritin single-chain Fv fragment is a functional protein with properties of a partially structured state: comparison with the completely folded V(L) domain." *Biochemistry, 39(27): 8047-8057.*

Martsev, S. P., A. P. Dubnovitsky, A. P. Vlasov, M. Hoshino, K. Hasegawa, H. Naiki, and Y. Goto. 2002a. "Amyloid fibril formation of the mouse V(L) domain at acidic pH." *Biochemistry, 41(10): 3389-3395.*

Martsev, S. P., A. P. Dubnovitsky, O. A. Stremovsky, A. A. Chumanevich, Y. I. Tsybovsky, Z. I. Kravchuk, and S. M. Deyev. 2002b. "Partially structured state of the functional VH domain of the mouse anti-ferritin antibody F11." *FEBS Lett., 518(1-3): 177-182.*

Martsev, S. P., Y. I. Tsybovsky, O. A. Stremovskiy, S. G. Odintsov, T. G. Balandin, P. Arosio, Z. I. Kravchuk, and S. M. Deyev. 2004. "Fusion of the antiferritin antibody VL domain to barnase results in enhanced solubility and altered pH stability." *Protein Eng. Des. Sel., 17(1): 85-93.*

Marvin, D. A. 1998. "Filamentous phage structure, infection and assembly." *Curr. Opin. Struct. Biol., 8: 150–158.*

Maynard, J., and G. Georgiou. 2000. "Antibody engineering." *Annu. Rev. Biomed. Eng., 2: 339-376.*

Mayr-Wohlfart, U., J. Waltenberger, H. Hausser, S. Kessler, K. P. Günther, C. Dehio, W. Puhl, and R. E. Brenner. 2002. "Vascular endothelial growth factor stimulates chemotactic migration of primary human osteoblasts." *Bone, 30(3): 472-477.*

McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell. 1990. "Phage antibodies: filamentous phage displaying antibody variable domains." *Nature, 348(6301): 552-554.*

Mikule, K., J. C. Gatlin, B. A. de la Houssaye, and K. H. Pfenninger. 2002. "Growth Cone Collapse Induced by Semaphorin 3A Requires 12/15-Lipoxygenase." *Journal of Neuroscience, 22(12): 4932-4941.*

Mülhardt, C. 2007. Molecular Biology and Genomics. *Elsevier Inc: Oxford, UK.*

Murga, M., O. Fernandez-Capetillo, and G. Tosato. 2005. "Neuropilin-1 regulates attachment in human endothelial cells independentlyof vascular endothelial growth factor receptor-2." *Blood, 105: 1992-1999.*

Muruganandam, A., J. Tanha, S. Narang, and D. Stanimirovic. 2002. "Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium." *FASEB J., 16(2): 240-242.*

Muyldermans, S. 2001. "Single domain camel antibodies." *Current Status. Rev. Mol. Biotech., 74: 277-302.*

Nakamura, F., M. Tanaka, T. Takahashi, R. G. Kalb, and S. M. Strittmatter. 1998. "Neuropilin-1 extracellular domains mediate semaphorin D/III-induced growth cone collapse." *Neuron 21, 1093–1100.*

Nakamura, F., R. G. Kalb, and S. M. Strittmatter. 2000. "Molecular basis of semaphorin-mediated axon guidance." *J Neurobiol., 44(2): 219–229.*

Nakamura, F., and Y. Goshima. 2002. "Structural and functional relations of neuropilins." *Adv. Exp. Med. Biol., 515: 55–69.*

Nasarre, C., M. Roth, L. Jacob, L. Roth, E. Koncina, A. Thien, G. Labourdette, P/ Poulet, P. Hubert, G. Crémel, G. Roussel, D. Aunis, and D. Bagnard. 2010. "Peptide-based interference of the transmembrane domain of neuropilin-1 inhibits glioma growth in vivo." *Oncogene, 29(16): 381-392.*

Neri, D., B. Carnemolla, A. Nissim, E. Balza, A. Leprini, G. Querze, A. Pini, L. Tarli, C. Halin, P. Neri, L. Zardi, and G. Winter. 1997. "Targeting by affinity-matured recombinant antibody fragments of an angiogenesis associated fibronectin isoform." *Nat. Biotechnol., 15: 1271–1275.*

Neylon, C. 2004. "Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution." *Nucleic Acids Res. l32, 1448-1459.*

Nilson, B. H., A. Solomon, L. Björck, and B. Akerström. 1992. "Protein L from Peptostreptococcus magnus binds to the kappa light chain variable domain." *The Journal of Biological Chemistry, 267, 2234-2239.*

Nissim, A., H. R. Hoogenboom, I. M. Tomlinson, G. Flynn, C. Midgley, D. Lane, and G. Winter. 1994. "Antibody fragments from a 'single pot' phage display library as immunochemical reagents." *EMBO J., 13: 692–698.*

Nuttall, S. D., R. A. Irving, and P.J. Hudson. 2000. "Immunoglobulin VH domains and beyond: design and selection of single-domain binding and targeting reagents." *Curr. Pharm. Biotechnol., 1(3): 253-263.*

Ohage, E., and B. Steipe. 1999. "Intrabody construction and expression. I. The critical role of VL domain stability." *J. Mol. Biol., 291: 1119-1128.*

Orlandi, R., D. H. Gussow, P. T. Jones, and G. Winter. 1989. "Cloning immunoglobulin variable domains for expression by the polymerase chain reaction." *Proc. Natl Acad. Sci., 86(10): 3833-3837.*

Orum, H., P. S. Andersen, A. Oster, L. K. Johansen, E. Riise, M. Bjørnvad, I. Svendsen, and J. Engberg. 1993 "Efficient method for constructing comprehensive murine Fab antibody libraries displayed on phage." *Nucleic Acids Res., 21(19): 4491-8.*

Padlan, E. A., C. Abergel, and J.P. Tipper. 1995. "Identification of specificity-determining residues in antibodies." *FASEB J., 9: 133-139.*

Pande, J., M. M. Szewczyk, and A. K. Grover. 2010. "Phage display: Concept, innovations, applications and future." *Biotechnol Adv., (EPub ahead of print).*

Paschke, M. 2006. "Phage display systems and their applications." *Appl. Microbiol. Biotechnol., 70: 2–11.*

Pasterkamp, R.J., R. J. Giger, and J. Verhaagen. 1998. "Regulation of semaphorin III/collapsin-I gene expression during peripheral nerve regeneration." *Exp. Neurol., 153: 313-327.*

Pasterkamp, R. J., and J. Verhaagen. 2001. "Emerging roles for semaphorins in neural regeneration." *Brain Res. Brain Res. Rev., 35: 36-54.*

Pasterkamp, R. J., and A. L. Kolodkin. 2003. "Semaphorin junction: making tracks toward neural connectivity." *Curr. Opin. Neurobiol., 13: 79– 89.*

Pasterkamp, J., and R. J. Giger. 2009. "Semaphorin function in neural plasticity and disease." *Curr. Opin. Neurobiol., 19: 263-274.*

Patrick, W. M., and A. E. Firth. 2005. "Strategies and computational tools for improving randomized protein libraries." *Biomolecular Engineering, 22(4): 105-112.*

Perelson, A. S., and G. F. Oster. 1979. "Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self non-self discrimination." *J. Theor. Biol., 81: 645-70.*

Persson, M. A., R. H. Caothien, and D. R. Burton. 1991. "Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning." *Proc. Natl. Acad. Sci., 88(6): 2432-6.*

Pini, A., A. Spreafico, R. Botti, D. Neri, and P. Neri. 1997. "Hierarchical affinity maturation of a phage library derived antibody for the selective removal of cytomegalovirus from plasma." *J. Immunol. Methods, 206: 171–182.*

Pini, A., F. Viti, A. Santucci, B. Carnemolla, L. Zardi, P. Neri, and D. Neri. 1998. "Design and use of a phage display library. Human antibodies with subnanomolar affinity against a marker of angiogenesis eluted from a two-dimensional gel." *J Biol Chem, 273: 21769-21776.*

Plückthun, A. 1994. "*Escherichia coli* producing recombinant antibodies." *Bioprocess Technol., 19: 233–252.*

Porath, J. 1992. "Immobilized Metal Ion Affinity Chromatography." *Protein Expression and Purification, 3: 263–281.*

Pugsley, A. P. 1993. "The complete general secretory pathway in gram-negative bacteria." *Microbiol. Rev., 57: 50–108.*

Raff, M. C., A. V. Whitmore, and J. T. Finn. 2002. "Axonal self-destruction and neurodegeneration." *Science, 296: 868-871.*

Rakonjac, J., J. N. Feng, and P. Model. 1999. "Filamentous phage assembly is terminated by two-step mechanism involving a short C-terminal fragment of pIII." *J. Mol. Biol. 289: 1253-1265.*

Rauchenberger, R., E. Borges, E. Thomassen-Wolf, E. Rom, R. Adar, Y. Yaniv, M. Malka, I. Chumakov, S. Kotzer, D. Resnitzky, A. Knappik, S. Reiffert, J. Prassler, K. Jury, D. Waldherr, S. Bauer, T. Kretzschmar, A. Yayon, and C. Rothe. 2003. "Human combinatorial Fab library yielding specific and functional antibodies against the human fibroblast growth factor receptor 3." *J. Biol. Chem., 278: 38194-38205.*

Reichert, J. M., C. J. Rosensweig, L. B. Faden, and M. C. Dewitz. 2005. "Monoclonal antibody successes in the clinic." *Nature Biotechnol.,* 23: 1073-1078.

Reiter, Y., P. Schuck, L. F. Boyd, and D. Plaksin. 1999. "An antibody single-domain phage display library of a native heavy chain variable region: isolation of functional single-domain VH molecules with a unique interface." *J. Mol. Biol., 290(3): 685-698.*

Renzi, M. J., L. Feiner, A. M. Koppel, and J. A. Raper. 1999. "A Dominant Negative Receptor for Specific Secreted Semaphorins Is Generated by Deleting an Extracellular Domain from Neuropilin-1." *J. Neurosci., 19, 7870–7880*.

Revets, H., P. D. Baetselier, and S. Muyldermans. 2005. "Nanobodies as novel agents for cancer therapy." *Expert Opin. Invest. Drugs, 5: 111–124*.

Riechmann, L., and P. Holliger. 1997. "The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*." *Cell, 90: 351-360*.

Roberts B. L., W. Markland, K. Siranosian, M. J. Saxena, S. K. Guterman, and R. C. Ladner. 1992. "Protease inhibitor display M13 phage: selection of high-affinity neutrophil elastase inhibitors." *Gene, 121: 9–15*.

Robinson, C. R., and S. T. Sauer. 1998. "Optimizing the stability of single-chain proteins by linker length and composition mutagenesis." *Proc. Natl. Acad. Sci. 95: 5929-5934.*

Roche, J., H. Drabkin, and E. Brambilla. 2002. "Neuropilin and its ligands in normal lung and cancer." *Adv. Exp. Med. Biol., 515: 103-14*.

Rodi, D. J., and L. Makowski L. 1999. "Phage-display technology–finding a needle in a vast molecular haystack." *Curr. Opin. Biotechnol.*, *10: 87-93*.

Rodi, D. J., A. S. Soares, and L. Makowski. 2002. "Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries." *J. Mol. Biol., 322(5): 1039-1052.*

Rosner, B. 2006. Fundamentals of Biostatistics. *Thomson Brooks/Cole, Belmont, CA*.

Rossignol, M., M. L. Gagnon,and M. Klagsbrun. 2000. "Genomic organization of human neuropilin-1 and neuropilin-2 genes: identification and distribution of splice variants and soluble isoforms." *Genomics, 70(2): 211–222.*

Rother, R. P., S. A. Rollins, C. F. Mojcik, R. A. Brodsky and L. Bell. 2007. "Discovery and development of the complement inhibitor eculizumab for the treatment of paroxysmal nocturnal hemoglobinuria." *Nature Biotechnology, 25: 1256-1264.*

Russel, M., H. B. Lowman, and T. Clackson. 2004. "Introduction to Phage Biology and Phage Display", in Phage Display: A Practical Approach. *Oxford University Press, New York: pp 1-24*.

Saerens, D., G. H. Ghassabeh, and S. Muyldermans. 2008. "Single-domain antibodies as building blocks for novel therapeutics." *Curr. Opin. Pharmacol., 8: 600-608.*

Sambrook, J., and D. W. Russell. 2001. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press: New York, USA*.

Sanchez-Carbayo, M., N. D. Socci, J. J. Lozano, W. Li, E. Charytonowicz, T. J. Belbin, M. B. Prystowsky, A. R. Ortiz, G. Childs, and C. Cordon-Cardo. 2003. "Gene discovery in bladder cancer progression using cDNA microarrays." *Am. J. Pathol., 163(2): 505-516.*

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. "DNA sequencing with chain terminating inhibitors." *Proc. Natl. Acad. Sci., USA 74: 5463-5467*.

Sanna, P. P., R. A. Williamson, A. De Logu, F. E. Bloom, and D. R. Burton. 1995. "Directed selection of recombinant human monoclonal antibodies to herpes simplex virus glycoproteins from phage display libraries." Proc. Natl. Acad. Sci., 92: 6439–43.

Sanz, I. 1991. "Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions." *J. Immunol., 147: 1720–1729.*

Scarlato, M., J. Ara, P. Bannerman, S. Scherer, and D. Pleasure. 2003. "Induction of neuropilins-1 and -2 and their ligands, Sema3A, Sema3F, and VEGF, during Wallerian degeneration in the peripheral nervous system." *Experimental Neurology, 183: 489–498*.

Schervish, M. J. 1996. "P Values: What They Are and What They Are Not." *The American Statistician, 50(3): 203-206.*

Schier, R., J. D. Marks, E. J. Wolf, G. Apell, C. Wong, J. E. McCartney, M. A. Bookman, J. S. Huston, L. L. Houston, and L. M. Weiner. 1995. "*In vitro* and *in vivo* characterization of a human anti-c-erbB-2 single-chain Fv isolated from a filamentous phage antibody library." *Immunotechnology, 1(1): 73-81.*

Schier, R., R. F. Balint, A. McCall, G. Apell, J. W. Larrick, and J. D. Marks. 1996. "Identification of functional and structural amino-acid residues by parsimonious mutagenesis." *Gene, 169: 147–155.*

Schlachetzki, F., C. Zhu. and W. M. Pardridge. 2002. "Expression of the neonatal Fc receptor (FcRn) at the blood-brain barrier." *J. Neurochem., 81: 203-206*.

Schmidbauer, S. B., and O. K. Strobel. 1997. "Rapid Purification of Histidine-Tagged Glutathione S-Transferase Fusion Protein by Metal Chelate POROS Perfusion Chromatography Media." *Biochemica, 1: 28-29*.

Scholle, M. D., J. W. Kehoe, and B. K. Kay. 2005. "Efficient construction of a large collection of phage-displayed combinatorial peptide libraries." *Comb. Chem. High Throughput Screen., 8(6): 545-551.*

Schuch, G., M. Machluf, G. Bartsch Jr., M. Nomi, H. Richard, A. Atala, and S. Soker. 2002. "In vivo administration of vascular endothelial growth factor (VEGF) and its antagonist, soluble neuropilin-1, predicts a role of VEGF in the progression of acute myeloid leukemia in vivo." *Blood, 100: 4622-4628*.

Scott, J.K., and G. P. Smith. 1990. "Searching for peptide ligands with an epitope library." *Science, 249: 386-390.*

Semaphorin Nomenclature Committee. 1999. "Unified nomenclature for the semaphorins/ collapsins." *Cell, 97(5): 551–552.*

Shao, C. Y., C. J. Secombes, and A. J. Porter. 2007. "Rapid isolation of IgNAR variable single-domain antibody fragments from a shark synthetic library." *Mol. Immunol., 44(4): 656-665.*

Sheriff, S., and K. L. Constantine. 1996. "Redefining the minimal antigen-binding fragment." *Nature Struct. Biol., 3: 733-736.*

Shirvan, A., I. Ziv, G. Fleminger, R. Shina, Z. He, I. Brudo, E. Melamed, and A. Barzilai. 1999. "Semaphorins as mediators of neuronal apoptosis**.**" *J. Neurochem., 73: 961-971.*

Shukla, G. S., and D. N. Krag. 2010. "Phage-displayed combinatorial peptide libraries in fusion to beta-lactamase as reporter for an accelerated clone screening: Potential uses of selected enzyme-linked affinity reagents in downstream applications." Comb. Chem. High Throughput Screen., 13(1): 75-87.

Sidhu, S. S., H. B. Lowman, B. C. Cunningham, and J. A. Wells. 2000. "Phage display for selection of novel binding peptides." *Methods Enzymol., 328: 333-363.*

Sidhu, S. S., B. Li, Y. Chen, F. A. Fellouse, C. Eigenbrot, and G. Fuh. 2004. "Phage-displayed Antibody Libraries of Synthetic Heavy Chain Complementarity Determining Regions." *J. Mol. Biol. (2004) 338, 299-310.*

Sidhu, S. S., and F. A. Fellouse. 2006. "Synthetic therapeutic antibodies." Nat. Chem. Biol., 2(12): 682-688.

Silacci, M., S. Brack, G. Schirru, J. Mårlind, A. Ettorre, A. Merlo, F. Viti, and D. Neri. 2005. "Design, construction, and characterization of a large synthetic human antibody phage display library." *Proteomics, 5: 2340–2350.*

Simon, M. D., K. Sato, G. A. Weiss, and K. M. Shokat. 2004. "A phage display selection of engrailed homeodomain mutants and the importance of residue Q50." *Nucleic Acids Res., 32(12): 3623-3631.*

Skerra, A., and A. Plückthun. 1988. "Assembly of a functional immunoglobulin Fv fragment in Escherichia coli." *Science, 240: 1038–1041.*

Skerra, A. 1994. "Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*." *Gene, 151: 131–135.*

Smith, G. P. 1985. "Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface." *Science, 228: 1315–1317.*

Smith, G. P., and V. A. Petrenko. 1997. "Phage Display." *Chem. Rev., 97, 391–410.*

Soker, S., H. Fidder, G. Neufeld, and M. Klagsbrun. 1996. "Characterization of novel vascular endothelial growth factor (VEGF) receptors on tumor cells that bind VEGF165 via its exon 7-encoded domain." *J. Biol. Chem., 271(10): 5761-7.*

Soker, S., S. Gollamudi-Payne, H. Fidder, H. Charmahelli, and M. Klagsbrun. 1997. "Inhibition of vascular endothelial growth factor (VEGF)-induced endothelial cell proliferation by a peptide corresponding to the exon 7-encoded domain of VEGF165." *J. Biol. Chem., 272: 31582–31588.*

Soker, S., S. Takashima, and H. Q. Miao. 1998. "Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor." *Cell, 92(6): 735–745*.

Songsivilai, S., J. M. Bye, J. D. Marks, and N. C. Hughes-Jones. 1990. "Cloning and sequencing of human lambda immunoglobulin genes by the polymerase chain reaction." *Eur. J. Immunol., 20(12): 2661-6.*

Sørensen, H. P., and K. K. Mortensen. 2005. "Advanced genetic strategies for recombinant expression in *Escherichia coli*." *J. Biotechnol., 115: 113-128.*

Sørensen, H. P., and K. K. Mortensen. 2005. "Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli." Microb. Cell Fact., 4(1): 1-8.*

Stephenson, J. M., S. Banerjee, N. K. Saxena, R. Cherian, and S. K. Banerjee. 2002. **"**Neuropilin-1 is differentially expressed in myoepithelial cells and vascular smooth muscle cells in preneoplastic and neoplastic human breast: a possible marker for the progression of breast cancer.**"** *Int. J. Cancer, 101(5): 409-414.*

Suzuki, Y., S. Ito, K. Otsuka, E. Iwasawa, M. Nakajima, and I. Yamaguchi. 2005. "Preparation of functional single-chain antibodies against bioactive gibberellins by utilizing randomly mutagenized phage-display libraries." *Biosci. Biotechnol. Biochem., 69(3): 610-9.*

Tabuchi, I., S. Soramoto, S.Ueno, and Y. Husimi. 2004. "Multi-line split DNA synthesis: a novel combinatorial method to make high quality peptide libraries." *BMC Biotechnol., 4: 19-26.*

Takahashi, T., A. Fournier, and F. Nakamura. 1999. "Plexin-neuropilin-1 complexes form functional semaphorin-3A receptors." *Cell, 99(1): 59–69*.

Takashima, S., M. Kitakaze, M. Asakura, H. Asanuma, S. Sanada, F. Tashiro, H. Niwa, J. Miyazaki Ji, S. Hirota, Y. Kitamura, T. Kitsukawa, H. Fujisawa, M. Klagsbrun, and M. Hori. 2002. "Targeting of both mouse neuropilin-1 and neuropilin-2 genes severely impairs

developmental yolk sac and embryonic angiogenesis." *Proc. Natl. Acad. Sci., 99(6): 3657-3662.*

Tamagnone, L., S. Artigiani, H. Chen, Z. He, G. I. Ming, H. Song, A. Chedotal, M. L. Winberg, C. S. Goodman, M. Poo, M. Tessier-Lavigne, and P. M. Comoglio. 1999. "Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates." *Cell, 99: 71–80.*

Tanha, J., P. Xu, Z. G. Chen, F. Ni, H. Kaplan, S. A. Narang, and C. R. MacKenzie. 2001. "Optimal design features of camelized human single-domain antibody libraries." *J. Biol. Chem. 276, 24774-24780.*

Tanha, J., G. Dubuc, T. Hirama, S. A. Narang, and C.R. MacKenzie. 2002. "Selection by phage display of llama conventional VH fragments with heavy chain antibody VHH properties." *J. Immunol. Methods, 263: 97–109.*

Tanha, J., A. Muruganadam, and D. Stanimirovic. 2003. "Phage display technology for identifying specific antigens on brain endothelial cells." *Meth. Mol. Med., 89: 435-449.*

Taniguchi, M., S. Yuasa, and H. Fujisawa. 1997. "Disruption of semaphorin III/D gene causes severe abnormality in peripheral nerve projection." *Neuron, 19(3): 519–530.*

Tchernychev, B., S. Cabilly, and M. Wilchek. 1997. "The epitopes for natural polyreactive antibodies are rich in proline." *Proc. Natl. Acad. Sci., 94(12): 6335-6339.*

Thakur, C. S., M. E. Brown, J. N. Sama, M. E. Jackson, and T. K. Dayie. 2010. "Growth of wildtype and mutant E. coli strains in minimal media for optimal production of nucleic acids for preparing labeled nucleotides." *Appl. Microbiol. Biotechnol., 88(3): 771-779.*

Thatcher, T. H., D. P. O'Brien, S. Altuwaijri, and R. K. Barth. 2000. "Increasing the frequency of T-cell precursors specific for a cryptic epitope of hen-egg lysozyme converts it to an immunodominant epitope." *Immunology. 99(2): 235-242.*

To, R., T. Hirama, M. Arbabi-Ghahroudi, R. MacKenzie, P. Wang, P. Xu, F. Ni, and J. Tanha. 2005. "Isolation of monomeric human VHs by a phage selection." *J Biol Chem. 280(50): 41395-41403.*

Tonegawa, S. 1983. "Somatic generation of antibody diversity." *Nature, 302(5909): 575-581.*

Tordjman, R., N. Ortéga, L. Coulombel, J. Plouët, P. H. Roméo, and V. Lemarchandel. 1999. "Neuropilin-1 is expressed on bone marrow stromal cells: a novel interaction with hematopoietic cells?" *Blood, 94(7): 2301-2309.*

Turner, D. J., M. A. Ritter, and A. J. George. 1997. "Importance of the linker in expression of single-chain Fv antibody fragments: optimization of peptide sequence using phage display technology." *J. Immunol. Meth., 205: 43-54.*

Vaughan, T. J., A. J. Williams, and K. Pritchard. 1996. "Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library." *Nat. Biotechnol., 14(3): 309-314*.

Verma, R., E. Boleti, A. J. George. 1998. "Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems." *J. Immunol. Methods, 216(1-2): 165-181*.

Vianello, F., E. Righi, and M. C. Poznansky. 2010. "Methods for quantitation of leukocyte chemotaxis and fugetaxis." *Methods Mol. Biol., 616: 115-124.*

Villaverde, A., and M. M. Carrio. 2003. "Protein aggregation in recombinant bacteria: biological role of inclusion bodies." *Biotechno.l Lett., 25: 1385-1395.*

Viti, F., F. Nilsson, S. Demartis, A. Huber, and D. Neri. 2000. "Design and use of phage display libraries for the selection of antibodies and enzymes." *Methods Enzymol., 326: 480–505.*

Wang, S., Y. Luo, X. Yi, W. Yu, Z. Xu, X. Ma, J. He, and Q. Liu. 2007. "A highly efficient and highly reliable protocol for transformation of *Escherichia coli* by electroporation." *J. Rapid Methods Autom. Microbiol., 15: 253-258.*

Wang, S. H., X. Y. Du, Y. M. Huang, D. S. Lin, P. L. Hart, and Z. H. Wang. 2007. "Detection of deoxynivalenol based on a single-chain fragment variable of the antideoxynivalenol antibody." *FEMS Microbiol. Lett., 272(2): 214-219.*

Ward, E. S., D. Gussow, A. D. Griffiths, P. T. Jones, and G. Winter. 1989. "Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*." *Nature, 341(6242): 544-546.*

Ward, R.L., M. A. Clark, J. Lees, and N. J. Hawkins. 1996. "Retrieval of human antibodies from phage-display libraries using enzymatic cleavage." *J. Immunol. Methods, 189: 73–82*.

Warnes, A., J. R. Stephenson, A. R. Fooks, and J. Melling. 1991. "Expression of recombinant protein A from the lac promoter in Escherichia coli JM83 is not subject to catabolite repression when grown under specific conditions of continuous culture." *Biotechnol. Bioeng., 38: 1050-1058.*

Warren, D. J. 2011. "Preparation of highly efficient electrocompetent *Escherichia coli* using glycerol/mannitol density-step centrifugation." *Anal. Biochem., [Epub ahead of print]*.

Wassaf, D., G. Kuang, K. Kopacz, Q. L. Wu, Q. Nguyen, M. Toews, J. Cosic, J. Jacques, S. Wiltshire, J. Lambert, C. C. Pazmany, S. Hogan, R. C. Ladner, A. E. Nixon, and D. J. Sexton. 2006. "High-throughput affinity ranking of antibodies using surface plasmon resonance microarrays." *Analytical Biochemistry, 351: 241–253*.

Wassman, C. D., P. Y. Tam, R. H. Lathrop, and G. A. Weiss. 2004. "Predicting oligonucleotide-directed mutagenesis failures in protein engineering." *Nucl. Acids Res. 32(21): 6407-6413.*

Watkins, N. A., and W. H. Ouwehand. 2000. "Introduction to Antibody Engineering and Phage Display." *Vox Sanguinis, 78(2): 72-79.*

Webster, R. E. 1996. "Biology of the filamentous bacteriophage", in "Phage Display of Peptides and Proteins". *Academic Press Inc., San Diego, CA. pp. 1–20.*

Whitaker, G. B., B. J. Limberg, and J. S. Rosenbaum. 2001. "Vascular endothelial growth factor receptor-2 and neuropilin-1 form a receptor complex that is responsible for the differential signaling potencyo f VEGF(165) and VEGF(121)." *J. Biol. Chem., 276: 25520-25531.*

Willats, W. G. 2002. "Phage display: practicalities and prospects." *Plant Mol. Biol.*, *50(6): 837-854.*

Williams, G., B. J. Eickholt, P. Maison, R. Prinjha, F. S. Walsh, and P. Doherty. 2005. "A complementary peptide approach applied to the design of novel semaphorin/neuropilin antagonists." *J. Neurochem, 92: 1180-1190.*

Winter, G., and C. Milstein. 1991. "Man-made antibodies." *Nature, 349: 293-299.*

Winter, G., A. D. Griffiths, R. E. Hawkins, and H. R. Hoogenboom. 1994. "Making antibodies by phage display technology." *Annu. Rev. Immunol., 12: 433-455.*

Winter, G. 1998. "Synthetic human antibodies and a strategy for protein engineering." *FEBS Lett., 430(1-2): 92-94.*

Wolkowicz, R., G. C. Jager, and G. P. Nolan. 2005. "A random peptide library fused to CCR5 for selection of mimetopes expressed on the mammalian cell surface via retroviral vectors." *J. Biol. Chem., 280(15):15195-201.*

Woo, J. H., Y. Y. Liu, A. Mathias, S. Stavrou, Z. Wang, J. Thompson, and D. M. Neville. 2002. "Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in Pichia pastoris." *Protein Expr. Purif., 25: 270–282.*

Yang, W. P., K. Green, S. Pinz-Sweeney, A. T. Briones, D. R. Burton, and C. F. Barbas III. 1995. "CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range." *J. Mol. Biol. 254: 392–403.*

Yang, H., M. Li, H. Chai, S. Yan, P. Lin, A. B. Lumsden, Q. Yao, and C. Chen. 2005. "Effects of cyclophilin A on cell proliferation and gene expressions in human vascular smooth muscle cells and endothelial cells." *J. Surg. Res., 123(2): 312-319.*

Yang, L., and J.P. Nolan. 2007. "High-Throughput Screening and Characterization of Clones Selected from Phage Display Libraries." *Cytometry Part A, 71A: 625-631.*

Yau, K. Y., H. Lee, and J. C. Hall. 2003. "Emerging trends in the synthesis and improvement of hapten-specific recombinant antibodies." *Biotechnol. Adv., 21(7): 599-637.*

Yin, C.C., J. Liu, Z.H. Lin, S. Z. Wang, X. B. Wang, and H. L. Huang. 2003. "An efficient approach to the synthesis of a high-quality random peptide repertoire." *Acta Biochem. Biophys. Sin., 35: 921-924.*

Yin, C. C., L. L. Ren, L. L. Zhu, X. B. Wang, Z. Zhang, H. L. Huang, and X. Y. Yan. 2008. "Construction of a fully synthetic human scFv antibody library with CDR3 regions randomized by a split-mix-split method and its application." *J. Biochem., 144(5): 591-598.*

Yokota, T., D. E. Milenic, M. Whitlow, and J. Schlom. 1992. "Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms." *Cancer Res., 52: 3402-3408.*

Yokota, T., Milenic, D.E., Whitlow, M., Wood, J.F., Hubert, S.L., Schlom, J., 1993. "Microautoradiographic analysis of the normal organ distribution of radioiodinated single-chain Fv and other immunoglobulin forms." *Cancer Res., 53: 3776-3783.*

Zacher, A. N. C. A. Stock, J. W. Golden, and G. P. Smith. 1980. "A New Filamentous Phage Cloning Vector: fd-tet." *Gene, 9:127-140.*

Zanette, D., W. Dundon, A. Soffientini, C. Sottani, F. Marinelli, A. Akeson, and E. Sarubbi. 1998. "Human IL-1 receptor antagonist from Escherichia coli: large-scale microbial growth and protein purification." *J. Biotechnol., 64: 187-196.*

Zhang, Y., and W. M. Pardridge. 2001. "Mediated efflux of IgG molecules from brain to blood across the blood–brain barrier." *J. Neuroimmunol., 114: 168-172.*

X. SUPPLEMENTAL

Table S1: Expected and Actual (Observed) Amino Acid Frequencies (%), as well as calculated FDI scores for 100 clones randomly selected from the constructed synthetic V_L phage display library.

Position 24 (R)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	2	-0.64420436
	T (Threonine)	6.25	7	0.30983867
	R (Arginine)	9.37	22	4.3340864
	L (Leucine)	9.37	11	0.55934765
	N (Asparagine)	3.12	1	-1.21938683
	S (Serine)	9.37	8	-0.47012655
	W (Tryptophan)	3.12	2	-0.64420436
	P (Proline)	6.25	5	-0.51639778
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	8	0.72295689
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	6	-0.10327956
	A (Alanine)	6.25	7	0.30983867
	D (Aspartic acid)	3.12	1	-1.21938683
	I (Isoleucine)	3.12	1	-1.21938683
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	1	-1.21938683
	C (Cysteine)	3.12	1	-1.21938683
	H (Histidine)	3.12	2	-0.64420436

Position 28 (S)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	S (Serine)	50	58	1.6
	G (Glycine)	50	39	-2.2

Position 30 (S)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	2	-0.64420436
	T (Threonine)	6.25	6	-0.10327956
	R (Arginine)	9.37	8	-0.47012655
	L (Leucine)	9.37	6	-1.15644269
	N (Asparagine)	3.12	5	1.08134303
	S (Serine)	9.37	14	1.58882186

	W (Tryptophan)	3.12	2	-0.64420436
	P (Proline)	6.25	8	0.72295689
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	8	0.72295689
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	8	0.72295689
	A (Alanine)	6.25	4	-0.929516
	D (Aspartic acid)	3.12	4	0.50616057
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	2	-0.64420436
	C (Cysteine)	3.12	3	-0.0690219
	H (Histidine)	3.12	2	-0.64420436

Position 31 (T)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	3	-0.0690219
	T (Threonine)	6.25	11	1.96231156
	R (Arginine)	9.37	10	0.21618958
	L (Leucine)	9.37	6	-1.15644269
	N (Asparagine)	3.12	4	0.50616057
	S (Serine)	9.37	9	-0.12696849
	W (Tryptophan)	3.12	4	0.50616057
	P (Proline)	6.25	7	0.30983867
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	4	-0.929516
	E (Glutamic acid)	3.12	2	-0.64420436
	V (Valine)	6.25	6	-0.10327956
	A (Alanine)	6.25	8	0.72295689
	D (Aspartic acid)	3.12	3	-0.0690219
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	2	-0.64420436
	C (Cysteine)	3.12	3	-0.0690219
	H (Histidine)	3.12	3	-0.0690219

Position 32 (Y)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	2	-0.64420436
	T (Threonine)	6.25	5	-0.51639778
	R (Arginine)	9.37	8	-0.47012655
	L (Leucine)	9.37	7	-0.81328462

	N (Asparagine)	3.12	3	-0.0690219
	S (Serine)	9.37	9	-0.12696849
	W (Tryptophan)	3.12	2	-0.64420436
	P (Proline)	6.25	6	-0.10327956
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	5	-0.51639778
	E (Glutamic acid)	3.12	4	0.50616057
	V (Valine)	6.25	4	-0.929516
	A (Alanine)	6.25	6	-0.10327956
	D (Aspartic acid)	3.12	6	1.6565255
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	4	0.50616057
	Y (Tyrosine)	3.12	16	7.40835015
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	1	-1.21938683

Position 34 (N)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	6	1.6565255
	M (Methionine)	3.12	4	0.50616057
	T (Threonine)	6.25	3	-1.34263423
	R (Arginine)	9.37	8	-0.47012655
	L (Leucine)	9.37	10	0.21618958
	N (Asparagine)	3.12	17	7.98353262
	S (Serine)	9.37	7	-0.81328462
	W (Tryptophan)	3.12	3	-0.0690219
	P (Proline)	6.25	4	-0.929516
	Q (Glutamine)	6.25	2	-1.75575245
	G (Glycine)	6.25	5	-0.51639778
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	6	-0.10327956
	A (Alanine)	6.25	2	-1.75575245
	D (Aspartic acid)	3.12	2	-0.64420436
	I (Isoleucine)	3.12	6	1.6565255
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	1	-1.21938683
	C (Cysteine)	3.12	3	-0.0690219
	H (Histidine)	3.12	2	-0.64420436

Position 50 (A)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	1	-1.21938683
	M (Methionine)	3.12	3	-0.0690219
	T (Threonine)	6.25	8	0.72295689

	R (Arginine)	9.37	8	-0.47012655
	L (Leucine)	9.37	9	-0.12696849
	N (Asparagine)	3.12	3	-0.0690219
	S (Serine)	9.37	10	0.21618958
	W (Tryptophan)	3.12	1	-1.21938683
	P (Proline)	6.25	16	4.02790268
	Q (Glutamine)	6.25	1	-2.16887067
	G (Glycine)	6.25	5	-0.51639778
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	7	0.30983867
	A (Alanine)	6.25	13	2.78854801
	D (Aspartic acid)	3.12	1	-1.21938683
	I (Isoleucine)	3.12	2	-0.64420436
	F (Phenylalanine)	3.12	2	-0.64420436
	Y (Tyrosine)	3.12	3	-0.0690219
	C (Cysteine)	3.12	1	-1.21938683
	H (Histidine)	3.12	0	-1.79456929

Position 53 (T)	**Codon**	**Expected Frequency (%)**	**Actual Frequency (%)**	**FDI**
	K (Lysine)	3.12	2	-0.64420436
	M (Methionine)	3.12	1	-1.21938683
	T (Threonine)	6.25	17	4.4410209
	R (Arginine)	9.37	7	-0.81328462
	L (Leucine)	9.37	7	-0.81328462
	N (Asparagine)	3.12	2	-0.64420436
	S (Serine)	9.37	5	-1.49960076
	W (Tryptophan)	3.12	3	-0.0690219
	P (Proline)	6.25	16	4.02790268
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	7	0.30983867
	E (Glutamic acid)	3.12	2	-0.64420436
	V (Valine)	6.25	4	-0.929516
	A (Alanine)	6.25	4	-0.929516
	D (Aspartic acid)	3.12	2	-0.64420436
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	3	-0.0690219
	C (Cysteine)	3.12	2	-0.64420436
	H (Histidine)	3.12	4	0.50616057

Position 54 (L)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	R (Arginine)	50	47	-0.6
	L (Leucine)	50	50	0

Position 55 (Q)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	2	-0.64420436
	T (Threonine)	6.25	5	-0.51639778
	R (Arginine)	9.37	9	-0.12696849
	L (Leucine)	9.37	11	0.55934765
	N (Asparagine)	3.12	3	-0.0690219
	S (Serine)	9.37	8	-0.47012655
	W (Tryptophan)	3.12	2	-0.64420436
	P (Proline)	6.25	9	1.13607511
	Q (Glutamine)	6.25	16	4.02790268
	G (Glycine)	6.25	5	-0.51639778
	E (Glutamic acid)	3.12	1	-1.21938683
	V (Valine)	6.25	3	-1.34263423
	A (Alanine)	6.25	7	0.30983867
	D (Aspartic acid)	3.12	2	-0.64420436
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	2	-0.64420436
	Y (Tyrosine)	3.12	2	-0.64420436
	C (Cysteine)	3.12	3	-0.0690219
	H (Histidine)	3.12	1	-1.21938683

Position 89 (Q)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	3	-0.0690219
	T (Threonine)	6.25	6	-0.10327956
	R (Arginine)	9.37	10	0.21618958
	L (Leucine)	9.37	10	0.21618958
	N (Asparagine)	3.12	1	-1.21938683
	S (Serine)	9.37	9	-0.12696849
	W (Tryptophan)	3.12	1	-1.21938683
	P (Proline)	6.25	6	-0.10327956
	Q (Glutamine)	6.25	5	-0.51639778
	G (Glycine)	6.25	8	0.72295689
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	6	-0.10327956

	A (Alanine)	6.25	15	3.61478446
	D (Aspartic acid)	3.12	3	-0.0690219
	I (Isoleucine)	3.12	1	-1.21938683
	F (Phenylalanine)	3.12	2	-0.64420436
	Y (Tyrosine)	3.12	2	-0.64420436
	C (Cysteine)	3.12	2	-0.64420436
	H (Histidine)	3.12	1	-1.21938683

Position 91 (S)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	4	0.50616057
	T (Threonine)	6.25	6	-0.10327956
	R (Arginine)	9.37	7	-0.81328462
	L (Leucine)	9.37	5	-1.49960076
	N (Asparagine)	3.12	3	-0.0690219
	S (Serine)	9.37	25	5.36356061
	W (Tryptophan)	3.12	1	-1.21938683
	P (Proline)	6.25	9	1.13607511
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	9	1.13607511
	E (Glutamic acid)	3.12	0	-1.79456929
	V (Valine)	6.25	6	-0.10327956
	A (Alanine)	6.25	10	1.54919334
	D (Aspartic acid)	3.12	0	-1.79456929
	I (Isoleucine)	3.12	1	-1.21938683
	F (Phenylalanine)	3.12	2	-0.64420436
	Y (Tyrosine)	3.12	0	-1.79456929
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	3	-0.0690219

Position 92 (Y)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	3	-0.0690219
	T (Threonine)	6.25	5	-0.51639778
	R (Arginine)	9.37	12	0.90250572
	L (Leucine)	9.37	4	-1.84275883
	N (Asparagine)	3.12	1	-1.21938683
	S (Serine)	9.37	9	-0.12696849
	W (Tryptophan)	3.12	0	-1.79456929
	P (Proline)	6.25	11	1.96231156
	Q (Glutamine)	6.25	2	-1.75575245

	G (Glycine)	6.25	4	-0.929516
	E (Glutamic acid)	3.12	2	-0.64420436
	V (Valine)	6.25	2	-1.75575245
	A (Alanine)	6.25	9	1.13607511
	D (Aspartic acid)	3.12	1	-1.21938683
	I (Isoleucine)	3.12	1	-1.21938683
	F (Phenylalanine)	3.12	0	-1.79456929
	Y (Tyrosine)	3.12	24	12.0098099
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	4	0.50616057

Position 93 (S)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	3	-0.0690219
	M (Methionine)	3.12	3	-0.0690219
	T (Threonine)	6.25	3	-1.34263423
	R (Arginine)	9.37	10	0.21618958
	L (Leucine)	9.37	5	-1.49960076
	N (Asparagine)	3.12	0	-1.79456929
	S (Serine)	9.37	24	5.02040254
	W (Tryptophan)	3.12	5	1.08134303
	P (Proline)	6.25	8	0.72295689
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	4	-0.929516
	E (Glutamic acid)	3.12	1	-1.21938683
	V (Valine)	6.25	7	0.30983867
	A (Alanine)	6.25	6	-0.10327956
	D (Aspartic acid)	3.12	1	-1.21938683
	I (Isoleucine)	3.12	4	0.50616057
	F (Phenylalanine)	3.12	1	-1.21938683
	Y (Tyrosine)	3.12	3	-0.0690219
	C (Cysteine)	3.12	3	-0.0690219
	H (Histidine)	3.12	3	-0.0690219

Position 94 (T)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	0	-1.79456929
	M (Methionine)	3.12	1	-1.21938683
	T (Threonine)	6.25	29	9.39843959
	R (Arginine)	9.37	3	-2.1859169
	L (Leucine)	9.37	7	-0.81328462
	N (Asparagine)	3.12	4	0.50616057
	S (Serine)	9.37	6	-1.15644269

	W (Tryptophan)	3.12	0	-1.79456929
	P (Proline)	6.25	11	1.96231156
	Q (Glutamine)	6.25	3	-1.34263423
	G (Glycine)	6.25	8	0.72295689
	E (Glutamic acid)	3.12	2	-0.64420436
	V (Valine)	6.25	3	-1.34263423
	A (Alanine)	6.25	5	-0.51639778
	D (Aspartic acid)	3.12	0	-1.79456929
	I (Isoleucine)	3.12	3	-0.0690219
	F (Phenylalanine)	3.12	3	-0.0690219
	Y (Tyrosine)	3.12	3	-0.0690219
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	6	1.6565255

Position 95a (insertion)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	10	3.95725536
	M (Methionine)	3.12	0	-1.79456929
	T (Threonine)	6.25	5	-0.51639778
	R (Arginine)	9.37	12	0.90250572
	L (Leucine)	9.37	10	0.21618958
	N (Asparagine)	3.12	0	-1.79456929
	S (Serine)	9.37	10	0.21618958
	W (Tryptophan)	3.12	0	-1.79456929
	P (Proline)	6.25	14	3.20166623
	Q (Glutamine)	6.25	2	-1.75575245
	G (Glycine)	6.25	10	1.54919334
	E (Glutamic acid)	3.12	2	-0.64420436
	V (Valine)	6.25	2	-1.75575245
	A (Alanine)	6.25	12	2.37542979
	D (Aspartic acid)	3.12	0	-1.79456929
	I (Isoleucine)	3.12	2	-0.64420436
	F (Phenylalanine)	3.12	5	1.08134303
	Y (Tyrosine)	3.12	2	-0.64420436
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	2	-0.64420436

Position 95b (insertion)	Codon	Expected Frequency (%)	Actual Frequency (%)	FDI
	K (Lysine)	3.12	6	1.6565255
	M (Methionine)	3.12	0	-1.79456929
	T (Threonine)	6.25	0	-2.5819889

	R (Arginine)	9.37	0	-3.2153911
	L (Leucine)	9.37	6	-1.15644269
	N (Asparagine)	3.12	6	1.6565255
	S (Serine)	9.37	0	-3.2153911
	W (Tryptophan)	3.12	0	-1.79456929
	P (Proline)	6.25	12	2.37542979
	Q (Glutamine)	6.25	12	2.37542979
	G (Glycine)	6.25	31	10.224676
	E (Glutamic acid)	3.12	6	1.6565255
	V (Valine)	6.25	6	-0.10327956
	A (Alanine)	6.25	6	-0.10327956
	D (Aspartic acid)	3.12	0	-1.79456929
	I (Isoleucine)	3.12	0	-1.79456929
	F (Phenylalanine)	3.12	0	-1.79456929
	Y (Tyrosine)	3.12	6	1.6565255
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	0	-1.79456929

Position 96 (R)	**Codon**	**Expected Frequency (%)**	**Actual Frequency (%)**	**FDI**
	K (Lysine)	3.12	2	-0.64420436
	M (Methionine)	3.12	5	1.08134303
	T (Threonine)	6.25	4	-0.929516
	R (Arginine)	9.37	34	8.45198322
	L (Leucine)	9.37	5	-1.49960076
	N (Asparagine)	3.12	0	-1.79456929
	S (Serine)	9.37	8	-0.47012655
	W (Tryptophan)	3.12	1	-1.21938683
	P (Proline)	6.25	7	0.30983867
	Q (Glutamine)	6.25	2	-1.75575245
	G (Glycine)	6.25	5	-0.51639778
	E (Glutamic acid)	3.12	3	-0.0690219
	V (Valine)	6.25	7	0.30983867
	A (Alanine)	6.25	4	-0.929516
	D (Aspartic acid)	3.12	3	-0.0690219
	I (Isoleucine)	3.12	0	-1.79456929
	F (Phenylalanine)	3.12	1	-1.21938683
	Y (Tyrosine)	3.12	1	-1.21938683
	C (Cysteine)	3.12	0	-1.79456929
	H (Histidine)	3.12	1	-1.21938683

Figure S1: Amino acid sequence for Sema3A (771 amino acids).

```
        10         20         30         40         50         60
MGWLTRIVCL FWGVLLTARA NYQNGKNNVP RLKLSYKEML ESNNVITFNG LANSSSYHTF

        70         80         90        100        110        120
LLDEERSRLY VGAKDHIFSF DLVNIKDFQK IVWPVSYTRR DECKWAGKDI LKECANFIKV

       130        140        150        160        170        180
LKAYNQTHLY ACGTGAFHPI CTYIEIGHHP EDNIFKLENS HFENGRGKSP YDPKLLTASL

       190        200        210        220        230        240
LIDGELYSGT AADFMGRDFA IFRTLGHHHP IRTEQHDSRW LNDPKFISAH LISESDNPED

       250        260        270        280        290        300
DKVYFFFREN AIDGEHSGKA THARIGQICK NDFGGHRSLV NKWTTFLKAR LICSVPGPNG

       310        320        330        340        350        360
IDTHFDELQD VFLMNFKDPK NPVVYGVFTT SSNIFKGSAV CMYSMSDVRR VFLGPYAHRD

       370        380        390        400        410        420
GPNYQWVPYQ GRVPYPRPGT CPSKTFGGFD STKDLPDDVI TFARSHPAMY NPVFPMNNRP

       430        440        450        460        470        480
IVIKTDVNYQ FTQIVVDRVD AEDGQYDVMF IGTDVGTVLK VVSIPKETWY DLEEVLLEEM

       490        500        510        520        530        540
TVFREPTAIS AMELSTKQQQ LYIGSTAGVA QLPLHRCDIY GKACAECCLA RDPYCAWDGS

       550        560        570        580        590        600
ACSRYFPTAK RRTRRQDIRN GDPLTHCSDL HHDNHHGHSP EERIIYGVEN SSTFLECSPK

       610        620        630        640        650        660
SQRALVYWQF QRRNEERKEE IRVDDHIIRT DQGLLLRSLQ QKDSGNYLCH AVEHGFIQTL

       670        680        690        700        710        720
LKVTLEVIDT EHLEELLHKD DDGDGSKTKE MSNSMTPSQK VWYRDFMQLI NHPNLNTMDE
```

```
       730        740        750        760        770
FCEQVWKRDR KQRRQRPGHT PGNSNKWKHL QENKKGRNRR THEFERAPRS V
```

Figure S2: Amino acid sequence for NRP1 (923 amino acids).

```
        10         20         30         40         50         60
MERGLPLLCA VLALVLAPAG AFRNDKCGDT IKIESPGYLT SPGYPHSYHP SEKCEWLIQA

        70         80         90        100        110        120
PDPYQRIMIN FNPHFDLEDR DCKYDYVEVF DGENENGHFR GKFCGKIAPP PVVSSGPFLF

       130        140        150        160        170        180
IKFVSDYETH GAGFSIRYEI FKRGPECSQN YTTPSGVIKS PGFPEKYPNS LECTYIVFVP

       190        200        210        220        230        240
KMSEIILEFE SFDLEPDSNP PGGMFCRYDR LEIWDGFPDV GPHIGRYCGQ KTPGRIRSSS

       250        260        270        280        290        300
GILSMVFYTD SAIAKEGFSA NYSVLQSSVS EDFKCMEALG MESGEIHSDQ ITASSQYSTN

       310        320        330        340        350        360
WSAERSRLNY PENGWTPGED SYREWIQVDL GLLRFVTAVG TQGAISKETK KKYYVKTYKI

       370        380        390        400        410        420
DVSSNGEDWI TIKEGNKPVL FQGNTNPTDV VVAVFPKPLI TRFVRIKPAT WETGISMRFE

       430        440        450        460        470        480
VYGCKITDYP CSGMLGMVSG LISDSQITSS NQGDRNWMPE NIRLVTSRSG WALPPAPHSY

       490        500        510        520        530        540
INEWLQIDLG EEKIVRGIII QGGKHRENKV FMRKFKIGYS NNGSDWKMIM DDSKRKAKSF

       550        560        570        580        590        600
EGNNNYDTPE LRTFPALSTR FIRIYPERAT HGGLGLRMEL LGCEVEAPTA GPTTPNGNLV

       610        620        630        640        650        660
DECDDDQANC HSGTGDDFQL TGGTTVLATE KPTVIDSTIQ SEFPTYGFNC EFGWGSHKTF

       670        680        690        700        710        720
CHWEHDNHVQ LKWSVLTSKT GPIQDHTGDG NFIYSQADEN QKGKVARLVS PVVYSQNSAH
```

```
       730        740        750        760        770        780
CMTFWYHMSG SHVGTLRVKL RYQKPEEYDQ LVWMAIGHQG DHWKEGRVLL HKSLKLYQVI

       790        800        810        820        830        840
FEGEIGKGNL GGIAVDDISI NNHISQEDCA KPADLDKKNP EIKIDETGST PGYEGEGEGD

       850        860        870        880        890        900
KNISRKPGNV LKTLDPILIT IIAMSALGVL LGAVCGVVLY CACWHNGMSE RNLSALENYN

       910        920
FELVDGVKLK KDKLNTQSTY SEA
```

XI. CURRICULUM VITAE

EDUCATION

Sep 2008 – Aug 2011	**University of Ottawa**	**Ottawa, ON**
	Master's of Science, Microbiology & Immunology	
Sep 2003 – Apr 2008	**University of Ottawa**	**Ottawa, ON**
	Bachelor's of Science, Biochemistry (Honours)	
Sep 1998 – Jun 2003	**Lisgar Collegiate Institute**	**Ottawa, ON**

Specialized Courses:

National Research Council Canada – Workplace Hazardous Materials Information System

PUBLICATIONS

Hou, S. T., **A. Keklikian**, J. Slinn, M. O'Hare, S. X. Jiang, and A. Aylsworth. 2008. "Sustained up-regulation of semaphorin 3A, Neuropilin1, and doublecortin expression in ischemic mouse brain during long-term recovery." *Biochem. Biophys. Res. Commun., 367: 109-115.*

STRENGTHS

- Learn new skills and methodologies quickly and competently
- Possess strong organizational and communication skills
- Work in a diligent manner independently and cooperatively
- Show focus and perseverance

EXTRA-CURRICULAR ACTIVITIES

- Member of Ottawa Travellers Adult Hockey League (2007-2011)
- Member of Ottawa Senior Men's Hockey League (2006-2007)
- Member of Sandy Hill + Ottawa Centre Juvenile hockey teams (2003-2006)
- Member of Microbiology & Immunology Journal Club (University of Ottawa)
- Assistant Manager for Sandy Hill Minor Hockey Team (2004)

WORK EXPERIENCE

University of Ottawa, Department of Microbiology & Immunology **Sep 2008 – Aug 2010**
with the Institute for Biological Sciences, **Ottawa, ON**
National Research Council of Canada

Master's Research Project: "Construction of a synthetic human V_L phage display library and isolation of potential Neuropilin-1-specific V_L therapeutics from the library"

- A diverse synthetic V_L single domain antibody phage display library was constructed in *E. coli* and characterized by sequence and statistical analysis

- The V_L library was screened against a target of therapeutic relevance (Neuropilin-1, which interacts with Sema3A to promote neurodegeneration) for the isolation of potential binders that would ultimately inhibit the NRP1/Sema3A interaction

- V_L isolates were purified by Immobilized Metal Affinity Chromatography and subsequently characterized by Size Exclusion Chromatography for aggregation tendencies as well as Flow Cytometry for cell binding

Techniques: Cell culturing; Colony-PCR; Agarose gel electrophoresis; Western blot; SDS-PAGE; DNA purification; Bacterial expression; Protein extraction; ELISA; Phage panning; Plating for bacterial colony growth; *in vitro* DNA mutagenesis; Cloning; Size exclusion chromatography (gel filtration using a Superdex75 column and FPLC system); Immobilized metal affinity chromatography (IMAC) using a HiTrap Chelating column and FPLC system; Flow cytometry

National Research Council of Canada **May – Aug 2008**
with the Institute for Biological Sciences **Ottawa, ON**
Summer Student

- Worked with Antibody Engineering Group to carry out binding assays for characterization of V_HH single domain antibody isolates obtained against Fractalkine (antigenic target of interest)

Techniques: Colony-PCR; Agarose gel electrophoresis; DNA purification; Bacterial expression; Protein extraction; ELISA; Plating for bacterial colony growth

National Research Council of Canada **May – Aug 2007**
with the Institute for Biological Sciences **Ottawa, ON**
Summer Student

- Worked with the Experimental Stroke Group to analyze expression patterns of the major repulsive guidance cue Semaphorin 3A (Sema3A) as well as its receptor Neuropilin1 (NRP1) in the ischemic brains of stroke-induced adult mice

- Demonstrated the sustained elevation of Sema3A/NRP1 expression in the ischemic area, which contributes to the prevention of new neurons from regenerating and the inhibition of neurogenesis

Techniques: Tissue sectioning (Cryostat); Immunostaining (double labelling); Fluorescent microscopy