Antibody Discovery, Optimization, and Application: Translational Protein Engineering for Precision Medicines

By Bailey Bea Banach

B.S., Biomedical Engineering, University of Iowa, 2016

Submitted to the graduate degree program in Bioengineering and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Co-Chair: Brandon DeKosky

Co-Chair: Kenneth Fischer

Cory Berkland

Jennifer Robinson

Anthony Fehr

Date Defended: January 26th, 2022

The dissertation committee for Bailey Banach certifies that this is the approved version of the following dissertation:

Antibody Discovery, Optimization, and Application: Translational Protein Engineering for Precision Medicine

By Bailey Bea Banach

Co-Chair: Brandon DeKosky

Co-Chair: Kenneth Fischer

Date Approved: January 31st, 2022

Dedication

Commit your work to the Lord, and your plans will be established.
Proverbs 16:3

This dissertation is dedicated to God Almighty: my creator, my strength, and my shield.
He has been my source of rest and determination throughout this program; by His grace I prevail.

Acknowledgements

I am grateful to have been surrounded by a team of smart, capable, and generous people during my graduate career. These individuals have guided me through my graduate career, and to all of you I say, thank you; your dedication, patience, and kindness invested in me has made all the difference.

The undergraduate research mentors I learned from at the University of Iowa, specifically: Budd A. Tucker and Kristan S. Worthington, demonstrated firsthand the excitement of bench-top research and I'm thankful they bestowed this value in me as well. Many thanks and appreciation to my graduate research advisor at the University of Kansas, Brandon J. DeKosky, for playing an instrumental role in developing my independence as a research scientist and for equipping me with a repertoire of research skills to further advance my scientific curiosity. I'm very thankful to all the key collaborators & mentors from multiple partner research institutes who's combined breadth in expertise helped define this research including Peter D. Kwong, Lawrence Shapiro, Robert A. Seder, John R. Mascola, David H. Ho, Timothy A. Whitehead, Marie Pancera, Azza H. Idris, Gabrielle Cerutti, Prabhanshu Tripathi, Chen-Hsiang Shen, Baoshan Zhang, Reda Rawi, Tatsiana Bylund, Nicole Doria-Rose, Sergei Pletnev, Jason Gormon, Kai Xu and others in their research labs at the Vaccine Research Center, Columbia University, University of Colorado, Fred Hutchinson Cancer Research Center.

Thank you to the University of Kansas faculty members who provided guidance and feedback on my thesis as well as advising service on my committee: Kenneth Fischer, Cory Berkland, Jennifer Robinson, and Tony Fehr. I'm sincerely thankful for the numerous people and funding entities who provided financial opportunities for me to engage in this research, specifically Madison "Al" Self and his wife, Lila Self, whose generously endowment of the Self Graduate Fellowship award stipend my research career from 2016 to 2020 while providing professional development training. To the team of administrative faculty and staff at KU who helped me enter and navigate the bioengineering program, the Madison & Lila Self Graduate Fellowship, financial support, and professional development training: I appreciate all your time, engagement, and contributions that expanded my PhD experience far beyond the lab. Thank you so

much to the team of smart, capable lab members and research collaborators at the University of Kansas for providing invaluable advice, assistance, mentorship, and collaborative spirit through-out my graduate career: Ahmed S. Fahad, Bharat Madan, Jacy R. Wolfe, Matheus Oliveria de Souza, Matias Gutiérrez-González, Sheila N. López Acevedo, Xiaoli Pan, Amy D. Laflin, Rajani Nimrania, Natalie Bui, and Thuy Duong Nguyen.

Thank you to my close friends who supported me with encouragement during this program: Laura Shepherd, Katherine Baer, Kelsie Smock, Rachel Beitzel, and Kamryn Kletschke-Fegenbush. Thank you to every barista who made me a cup of coffee these years, especially my dear friends: Christine and Jon Clutton at the Wildway coffee. Thank you to my classmate, SGF cohort-fellow, and bioengineering colleague: Ember Krech, who's faith and joy during graduate school gave me a new appreciation for the importance of friendship. Thank you to my family: Ed, Kelly, Riley, Brittany, Ross, Steve, and Margie Banach and Linda, Lamar, and Titus Grayson for lifting me up through-out this chapter of my life, and all the other ones as well. Finally, I thank my husband, Patrick John Rhoads, who encourages me endlessly and has been affected in every way possible by this endeavor. Dear Patrick, thank you. My love for you can never be quantified. May God eternally bless you.

Antibody Discovery, Optimization, and Application:

Translational Protein Engineering for Precision Medicine

Bailey Bea Banach

The University of Kansas, 2021

Next-generation, “omics-based”, translational research is rapidly characterizing genetically determined pathophysiology at an unprecedented depth. Through the application of high-throughput sequencing technologies and advanced immune engineering methods new insights into adaptive immunity has been generated at an extraordinary rate. However, despite rapid accumulation of new scientific insight, critical gaps in our scientific understanding of antibody-mediated molecular immunity limit the number of novel antibody-based medical interventions that translate past basic scientific discovery and into clinical application. Here, this original work directly addresses outstanding gaps in our scientific knowledge of molecular humoral immunity through the discovery, optimization, enhancement, and biophysical characterization of monoclonal antibody proteins to inform the development of new precision medicines. Experimental monoclonal antibody genes, structure, and functions are assessed using a suite of immune engineering technologies, including: natively-paired antibody heavy:light chain complementary DNA libraries, *in vitro* mutagenesis, recombinant DNA vectors, functional yeast-surface display, fluorescence activated cell screening, high-throughput single-cell next-generation sequencing, and advanced biophysical characterization assays to translate basic research findings into new, clinically relevant, insights. Experimental application of these technologies using functional antibody:antigen screening methods produced the discovery of a novel SARS-CoV-2 neutralizing monoclonal antibody 910-30, and delineated antibody paired heavy:light chain sequence-structure-function signatures contributing to potent SARS-CoV-2 neutralization in public antibody responses. These findings have important applications for understanding communal immunity and developing effective intervention strategies during an acute global pandemic outbreak. This work also describes translational antibody protein engineering methods used to define precise, functionally optimized, genetic sequences from template monoclonal antibodies CIS43 and

VRC34.01, which target the clinically relevant pathogens *plasmodium falciparum* and viral HIV-1, respectively. The antibody insights gained from functional optimization experiments were used to produce improved biomolecular blueprints defining structural mechanisms and forward pathways for highly protective humoral immunity against *P. falciparum* and HIV-1, respectively. Altogether, the scientific outcomes from this research have immediate clinical applications for the development of therapeutic, prophylactic, diagnostic and research reagents for COVID-19, HIV/AIDS, and malaria, as well as broad impact for the development of precision medicines against diseases of clinical relevance.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION AND BACKGROUND

1.1 Antibody Engineering for the Clinical Development of Precision Medicines

Immune engineers leverage features of natural immunity to invent, design, and build targeted solutions which alleviate clinical dilemmas and advance scientific investigation. Collaboration between scientists, researchers, clinical doctors, and public health officials has led to the detailed characterization of immune *antibody* proteins that use gene-specific recognition mechanisms to target and eliminate pathogens with high precision. Through the omics-based expansion of biologic drug development known as *precision medicine*, the genetic features related to desired humoral immunity outcomes are now being determined at unprecedented depth (*1, 2*). Highly informative *antibody-omics* studies map the capacity of humoral immunity genetic features to impact infection outcomes using comprehensive immune protein-interrogation technologies to inform new disease intervention strategies (*3*). The research described here generates a detailed functional knowledge of precise antibody features related to clinically relevant health outcomes translated to the design of targeted, antibody-based medicines.

Modern antibody research began in the late 19th century when the presence of adaptive "anti-toxin" immune components in the blood were confirmed through scientific investigations into *C. diphtheria* toxin immunization led by Shibasaburo Kitasato and Emil Von Behring (*4*). Expanding on these studies, German chemist: Paul Ehrlich, created the first working model of an antibody protein, named for the German phrase: *anti-körper,* "against-body" for the protein's unique ability to recognize substances that were foreign to the body (*5*). Building off interdisciplinary investigations of antibody-mediated immunity, researchers in the mid 20th century described in increasingly molecular detail the developmental and structural features defining this unique class of immune proteins' capacity for foreign body recognition. These insights translated into the invention of *hybridoma* technologies, produced by investigators Kohler and Millstein in the 1970s, allowing for the reproducible isolation of antibodies arising from a single "monoclonal*"* B-cell clone in culture, i.e., *monoclonal antibodies (6)*. Hybridomas marked a significant breakthrough in the development of antibody biologics, allowing for powerful single cell molecular interrogation of the

antibody sequence-structure-function relationship. These foundational studies paved the way for recombinant antibody engineers to invent new methods to rationally modify and refine these unique immune proteins for translational applications using cutting-edge DNA and cell culture technologies (*7, 8*). Building on these scientific advances, the global antibody-based immune-technology market has grown to a conservative 150 B (USD, 2021) and is projected to increase to ~360 B by 2026.

Since the invention of monoclonal antibodies, engineers have progressed antibody characterization research using advanced high-throughput technologies paired with single-cell microfluidics, library-scale biophysical protein characterization, and next generation sequencing to rapidly define molecular features correlated with immune protection (*9-25*). Yet despite the expedited pace at which this antibody research has developed, translation of these insights into clinical care has lagged in comparison. Development in translational antibody engineering for precision medicines is limited by critical gaps in our fundamental knowledge of disease mechanisms, as well as challenges integrating this mechanistic knowledge into transformative health care strategies (*26-31*). Key requirements to optimize the next generation of precision antibody-based medicines include: (1) A greater understanding of antibody features that enable disease-prevention mechanisms and (2) clear directions on how to apply this understanding for transformative clinical impact.

To address these concerns, this body of work generates research insights related to antibody mediated immunity in communicable disease applications. Functional screening and characterization of the sequence-structure-function relationship defining antibody recognition produced detailed knowledge on features of effective humoral immunity, shedding new light on disease pathogenesis and prevention. Immune protein engineering technologies were applied to translate these insights into promising medical intervention strategies by inventing new antibody-based biomolecular blueprints to inform the design of future precision medicines.

1.1.1 Historical Perspectives on Immune Engineering

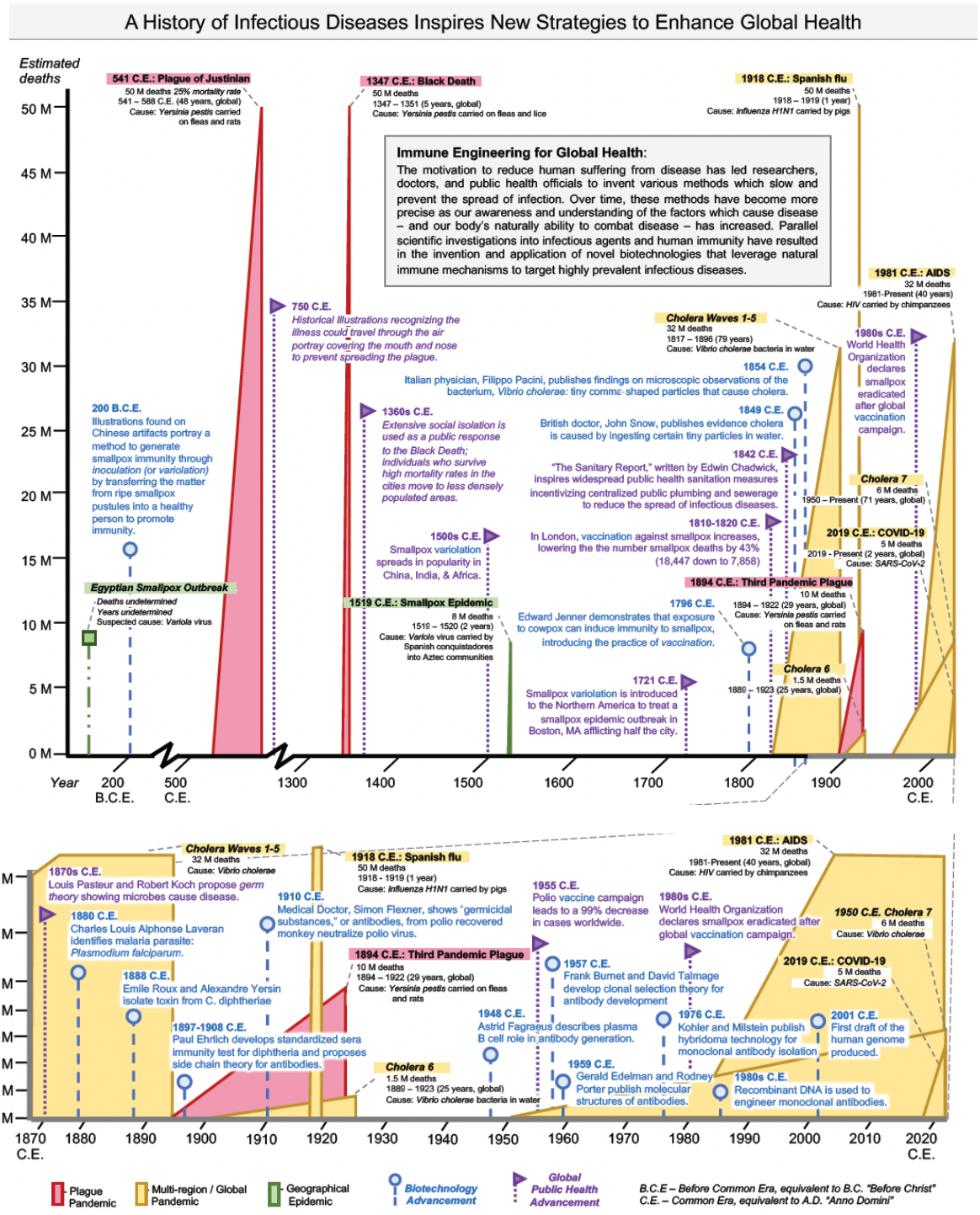


Figure 1. A timeline of human health outcomes from exposure to infectious diseases paralleling historical developments in biotechnology and public health measures related to the study of infectious diseases and immune engineering efforts.

Infectious diseases – and our ability to respond to them – have made a profound impact on human history. From the Plague of Justinian (541-750 A.D.), the first known global pandemic on record, to the most recent global coronavirus pandemic, civilization has been repeatedly impacted by critical encounters with infectious agents (Fig. 1) (*32, 33*). Historically, these small-scale molecular level interactions have been implicated in large scale consequences for human health. The spread of infectious disease across multiple regions or continents has cumulatively impacted billions of people over time and historically caused global suffering in substantial measures. Scientists, researchers, and public health officials in the common era have worked collaboratively to produce powerful tools preventing the spread of global disease. These efforts leverage centuries of biological insight and expand on novel molecular biotechnologies to elicit targeted immune responses that can rapidly improve human health outcomes.

Before the acceptance of germ theory and prior to the common era (CE), preliminary immune engineering efforts relied on empirical methods to achieved desired immune outcomes (*34*). Archeological artifacts provide evidence of smallpox infection caused by the *variola* virus and characterized by a painful rash these small sores, or pustules, often left severe scarring covering the face and body for those who survived. Smallpox-like scarring on Egyptian mummies suggests the disease may have plagued civilization for at least 3,000 years, whereas some accounts in India suggest as long as 5,000 years (*35*). Egyptian illustrations depict early strategies to elicit desired immune outcomes against smallpox disease; inoculation, or *variolation* as it later became named after the virus responsible for the disease was identified, involved exposing a healthy individual to material from smallpox sores, the *variolous* matter, by introducing the material into an incision in the arm or by inhaling as a powder through the nose. Ancient records show inoculation was used to control the spread of smallpox and suggests attempts to elicit immunity to the disease may have been made as early as 200 BCE. The earliest written record of smallpox appeared in 4th

century CE Chinese literature indicating that on average 3 out of every 10 people died from natural infection. Records show transfer of the matter from smallpox sores often produced mild symptoms of smallpox disease such as fever and rash and, importantly, inoculated individuals were more likely to survive smallpox exposure than those who did not receive treatment.

Following an increase in global connectivity, the practice of inoculation spread from Africa, China, India, and Turkey to Europe and the Americas. Inoculation helped control the spread of the disease, yet sometimes exposure to the matter from ripe pustules could prove fatal. A new method to protect against the deadly smallpox was introduced in May of 1796 through Edward Jenner's investigations into acquired immunity (*36*). Based on the observation that another disease: cowpox, which incited similar but milder forms of smallpox-like symptoms, Jenner experimentally transferred pustule material from a milkmaid, Sarah Nelmes, who he believed was infected with cowpox into an 8 yr. old boy, James Phipps, hypothesizing the boy would develop immunity against smallpox. Two months later Jenner challenged Phipps with matter from a fresh smallpox sore, and Phipps remained healthy. To distinguish this practice from inoculation this new method of acquiring immunity was coined vaccination, from the Latin phrase: *vacca*, for cow. The inception of Jenner's smallpox vaccine in the late 1700s bolstered numerous public health accomplishments at the turn of the century, advancing scientific investigations into adaptive immunity. In parallel to these new investigations, enhanced public healthcare measures such as vector control, sanitized water, improved hygiene, and vaccination campaigns became widespread effective strategies for mitigating disease damage.

While the factors that defined smallpox vaccine immunity were unknown at the time of Jenner's experiments his new method of vaccination proved safer in practice and more effective than alternative variolation methods and eventually vaccination was globally adopted and standardized. It would take scientists over 150 years following the introduction of inoculation in north America to recognize the critical role of microbes as the infectious agents of disease (*37*); in the 1870s Louis Pasteur and Robert Koch revealed parallel findings of ubiquitous disease-causing microscopic microbes. a Polish physician, Koch

was studying artificial anthrax bacteria microscopically to document the bacteria's natural replication cycle at the same time French chemist Pasteur's experimental with the biological process of yeast fermentation provided evidence showing such microorganisms were a product of normal biological reproduction. These studies increased awareness of molecular-level microbes and transformed our perspective on molecular immunity, as well as our intervention strategies against infectious agents. The discovery of antibodies was derived amid the scientific germ-theory revolution from investigations in the late 19th century led by Shibasaburo Kitasato and Emil Von Behring. Their research demonstrated that the blood product, i.e., sera or serum, of immunized guinea pig prevented the destructive effects of *C. diphtheriae* toxin through an "anti-toxin" protection mechanism (*38, 39*). Emile Roux and Alexandre Yersin later confirmed the presence of these anti-toxin immune components, laying the translational microbiology groundwork for what would later become serum therapies for diphtheria and tetanus (*40*).

Effective clinical application of serum therapies required standardization protocols to ensure reproducibility between doses, which became the work of German chemist: Paul Ehrlich (*41*). While working with diphtheria antitoxin, Ehrlich hypothesized that immune cells had molecular side chains like the dye molecules he had studied in his early work as a chemist. Insights from chemical dyes interactions with select tissues led Ehrlich to the hypothesis these cellular side chain receptors on immune cells consisted of multiple sites that achieved selectivity for a given pathogen. He proposed that these cellular side chains were the anti-toxin immune components in serum therapies and could detach from cells and act as "magic bullets" targeting a particular pathogen in circulation. Follow up investigations describing the molecular components of immune function helped confirm Ehrlich's side chain theory and led to the first working model of the antibody (*42*). Named after the German phrase: *anti-körper* "against body," scientific investigations from the mid-1900's into antibody-mediated immunity go on to describe in increasingly molecular detail the powerful *in vivo* mutagenesis developmental process and structurally guided features allowing natural selection of antigen-specific B-cell genes recognizing virtually unlimited foreign body diversity with extreme molecular specificity; In 1948 Astrid Fagraeus deduced plasma B cells involvement

in antibody generation, followed by Frank Burnet and David Talmage in 1957 who develop clonal selection theory to explain how an antibody's molecular specificity develops prior to antigen encounter (*43*). Two years later in 1959, Gerald Edelman and Rodney Porter produced molecular structure analyses of antibody proteins (*44*). These key scientific investigations, combined with many others, paved the way for Köhler & Milstein to develop their hybridoma technology in 1976 (*6*). Hybridomas allow reproducible isolation of antibodies arising from a single *monoclonal* B-cell culture, representing a new modern era for translational antibody research.

While Watson & Crick famously solved the structure of DNA in 1953 using crystallographic data from Rosalind Franklin and Maurice Wilkins, it wasn't until the recombinant DNA revolution starting in the 1970 that genetic engineering antibody took precedence (*45*). Given the early success of hybridoma production, monoclonal antibodies reached the commercial market in 1986, as hybridomas produced the first fully licensed therapeutic monoclonal antibody: Orthoclone OKT3 (muromonab-CD3, a *murine* mouse monoclonal antibody) for preventing kidney transplant rejection (*46*). Near the turn of the century in 1998 Matsuda et al published the complete nucleotide sequence of the human immunoglobulin heavy chain variable region gene locus detailing key genetic antibody features that can be engineered for targeted functionalities (*47*). This targeted genetic profiling information coupled parallel technological advancements in DNA sequencing starting at the 21st century generated new growth opportunities for human monoclonal antibody immune engineering. Starting in 2001 the first complete draft of the human genome spurred innovation in DNA sequencing technologies, leading to the development and widespread deployment of next generation sequencing (*48-50*). These technologies allow scientists a scalable sequencing method to study antibody phenotypes from individual clonal lineages as well as the full set of proteins defining the human antibody repertoire.

As history has demonstrated, a greater understanding of antibody-mediated immunity can accelerate new intervention strategies for clinically relevant infectious disease to achieve transformative global health outcomes. Using cutting-edge recombinant antibody engineering with high-throughput omics technologies

has allowed investigations into holistic antibody-mediated immunity at a level of molecular insight never achieved before. These techniques have been further supplemented and refined by structural, computational, and biophysical characterization analyses to define individual sequence-structure-function relationships for antibody recognition. These detailed molecular findings serve to translate modern monoclonal antibody discovery research into applications capable of to addressing the ongoing clinical suffering from a host of disease-causing agents (Fig 2) (*51, 52*). The impact of modern pandemics and epidemics related to HIV-1, Malaria, and COVID-19 represent ripe opportunities for translational immune engineering efforts to aid in the development of new global health interventions.

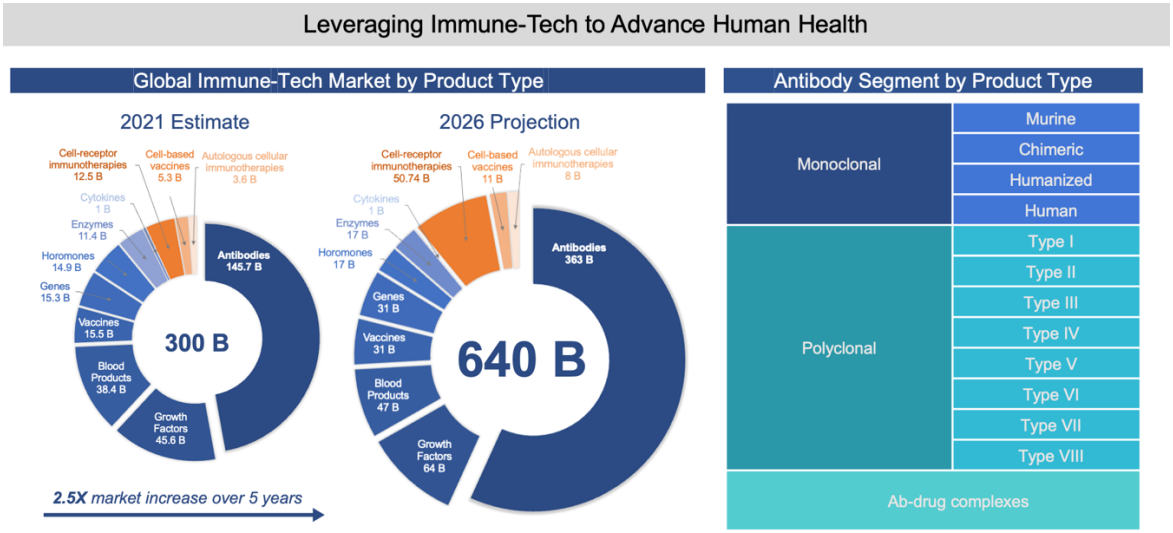


Figure 2. Leveraging immunology insights to create new biotechnologies that can advance human health.

1.1.2 Antibodies and Humoral Immunity Development

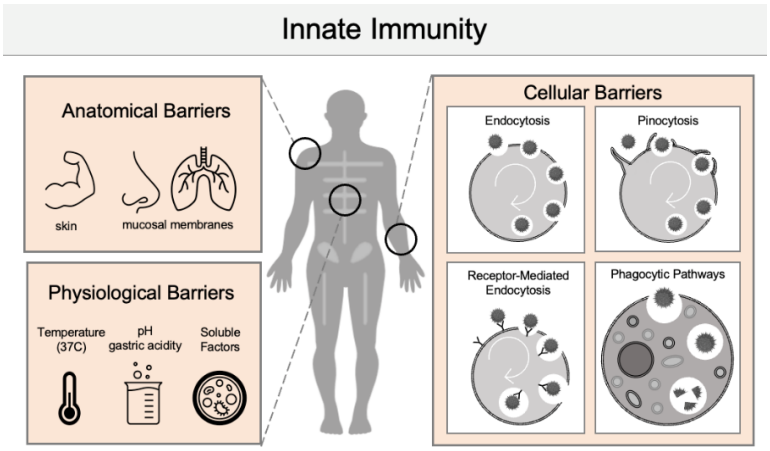


Figure 3. Overview of innate immunity.

Innate vs adaptive immunity - Animals have developed complex immune systems to distinguish self from non-self to protect themselves against diseases caused by invading microorganisms and cancer. Some of these systems are nonspecific defenses, i.e., innate (Fig. 3) while others are more specific i.e. adaptive (Fig. 4). Innate anatomical barriers such as skin and mucous membranes serve as a block to most microorganisms, and physiological barriers such as temperature, pH and soluble factors aid in the internal clearance of foreign bodies. Cellular components of innate immunity include endocytic and phagocytic cells that take in foreign substances to prevent further extra-cellular interactions with the body. Innate immunity also mediates inflammation, a general state of heightened immune response that can provide robust immune defense but can also lead to lasting tissue damage if left un-regulated.

Adaptive, or acquired, immunity has the capacity to generate diversified, long-term memory responses by forming highly specific, regulated cellular and molecular interactions leading to precise clearance of foreign substances known as *antigens* (Fig 4). There are three main cell types that recognize antigen in the adaptive immune response: B cells, T cells, and Antigen Presenting Cells (APCs). T cell and B cell are the lymphocytes primarily responsible for the specificity, memory, and diversity of immune recognition while APCs aid in the development of this acquired function.

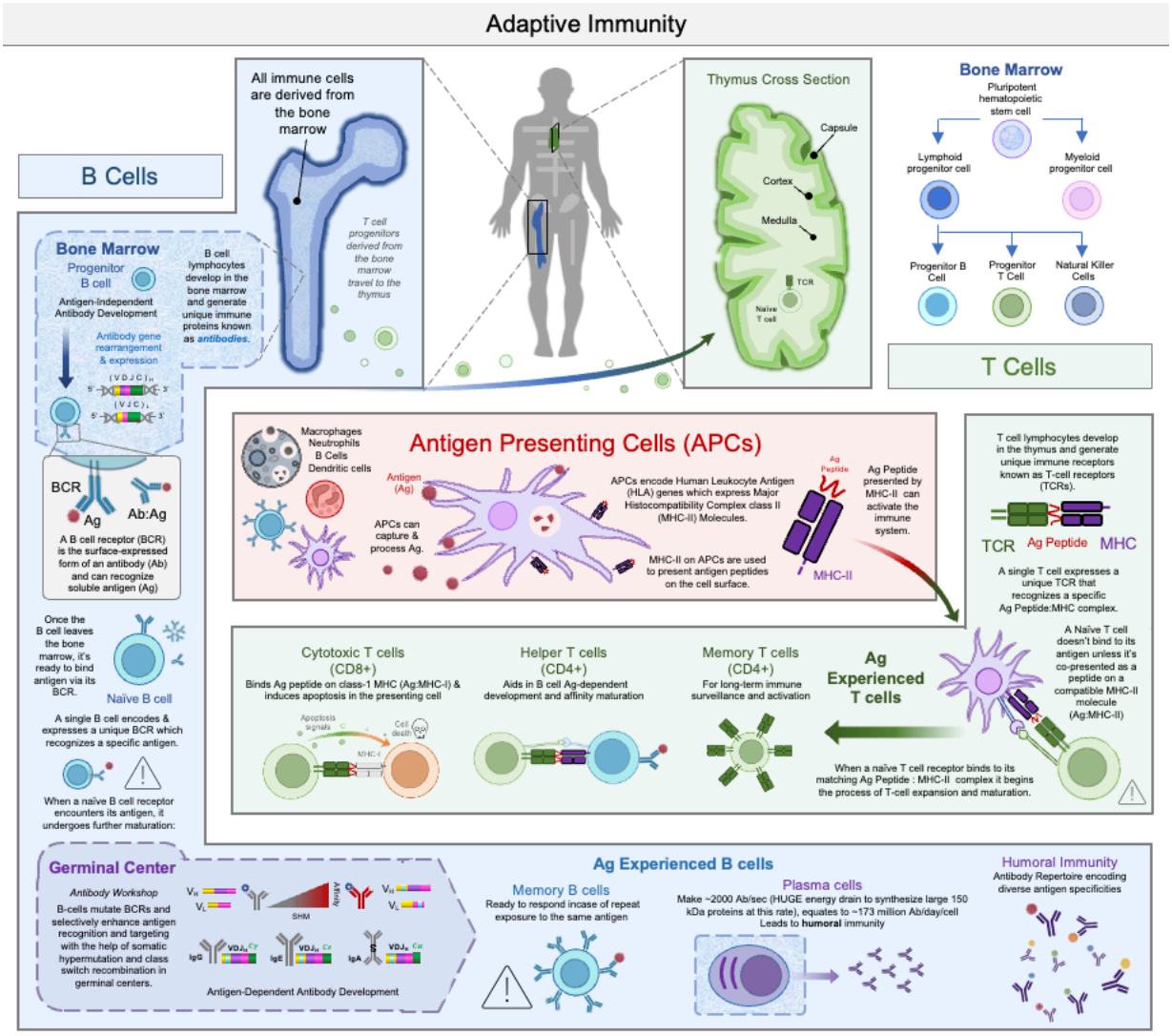


Figure 4. Overview of adaptive immunity.

B cell development starts in the bone marrow with the genetic recombination and expression of unique antigen receptors (i.e., B-cell receptors, BCRs) (Fig. 5). When a naïve B cell encounters its cognate antigen, it undergoes programmed cell division known as clonal expansion occurring in germinal development centers. With the help of APCs, somatic hypermutation and class switching is mediated in B cells by activation induced cytidine deaminase (AID), causing B cell somatic clonal expansion, leading to the

production of memory cells expressing genetically modified surface-bound antibodies and plasma cells secreting large quantities of tissue-specific antibodies for long-term humoral immune surveillance. T cell development starts with precursors in the bone marrow migrating to the thymus to mature, followed by development and expression of surface bound naïve T cell antigen receptors capable of recognizing antigen presented as a peptide fragment associated on another antigen-binding immune molecule: major histocompatibility complex (MHC). A T cell will not bind to its cognate antigen unless it's presented by an antigen presenting cell (APC) that has internalized, degraded, and processed said antigen protein into peptide fragments assembled with MHC on its surface. Mature T cells can be divided into memory/helper (CD4+) and cytotoxic (CD8+) subsets, each with distinct regulatory immune functions.

Development of long-lived memory B and T cells enables the body to respond more efficiently to repeat pathogen encounters. This fundamental immune "memory" phenomenon forms the basis of immunization intervention strategies by enabling the body to clear pathogens faster than would occur under naïve exposure conditions. Immune memory can be artificially induced using controlled immunization formats to deliver vulnerable antigens from target pathogens to the body without inciting disease. These vulnerable target antigens, known as *immunogens*, can initiate a primary immune response in recipients who have not yet encountered the pathogen. If an immunized individual later encounters a pathogen displaying the same immunogenic components presented by the vaccine, the immune system repeatedly recognizes this substance by recalling its immunization-induced immunological memory. Immunization-induced, or vaccine-induced, antibody responses have become an important line of defense in modern efforts to alleviate human suffering from infection and curb disease transmission rates (*53*).

In vivo B cell development & antibody maturation. Development of B cell receptors and antibodies, derived from *immunoglobulin* (Ig) proteins, occurs in tandem stages: antigen independent development and antigen dependent maturation, to produce homo-dimer protein macrostructures (~150 kDa) comprised of four amino acid chains: two "heavy" (~50 kDa) and two "light" (~25kDa) (Fig. 4, 5, and 6). Through inter-chain molecular interactions, antibody heavy and light chains assemble their variable regions into the

correct paired conformation structure to confer antigen recognition and biding. Antibody heavy chains comprise a constant region defined by five unique families of genes located on the *same* loci: Gamma (IgG), Alpha, (IgA), Mu (IgM), Epsilon (IgE), Delta (IgD). Expression of these constant genes produces distinct antibody *isotype* functions and access to different tissue compartments (*54*). Light chains also have two unique families of constant region genes but located on *different* loci: Kappa (IgK) and Lambda (Igλ) to aid in antibody assembly and function. Opposite to these constant regions on heavy and light chains are the antibody variable regions with large sequence diversity resulting in highly specific antigen binding.

The antigen binding sites on each immunoglobulin are identical, so a standard IgG antibody is considered bi-valent: each antibody molecule can bind up to two antigen molecules. Most antigens have multiple sites for antibody recognition, and it is not uncommon for multiple antibodies to form larger complexes with antigens to stimulate rapid immune clearance. The variable regions of antibody proteins can be further denoted into sub-regions by their functional involvement with antigen recognition: regions of the antibody that contact the antigen and form the surface of the interaction are labeled as complementary determining regions (CDRs) whereas the regions of the antibody structure that manipulate the display of the CDRs are labeled as framework regions (FRs). The CDRs form the antibody paratope surface which in turn recognizes and binds to the antigen epitope surface (Fig. 6)

The starting specificity of each B cell receptor is randomly determined by progenitor B cell gene rearrangement prior to encounter with antigen (antigen-independent development, Fig. 5), whereas the antigen-dependent specificity and tissue-specific antibody behavior is determined by affinity maturation and class switch recombination in germinal centers following B cell antigen-experience and activation. To make a naïve B cell receptor, the genes encoding the regions of the protein responsible for antigen recognition are rearranged to join sections of DNA into two distinct *DNA cassettes* encoding separate contiguous sequences for a heavy chain and a light chain, separated spatially in the human genome sequence by unique gene loci (Fig. 5). When productively rearranged, translated, and surface expressed as a homo-

dimer macro protein structure, these heavy:light paired DNA cassettes produce the distinct antigen-recognition function defining immunoglobulin proteins (Fig. 5 and 6).

In the variable region of the heavy chain protein, antigen recognition is defined by 3 families of subdomain genes: Variable (V_H) , Diversity (D_H), and Joining (J_H), whereas antigen recognition in the light chain originates from only 2 families of genes: Variable (V_L) and Joining (J_L) and does not include a Diversity segment like the heavy chain. With multiple genes in each family a massive combinatorial breadth of theoretically possible recombined heavy:light (VDJ_H:VJ_L) paired sequences generates diverse antigen recognition for humoral function. Progenitor B cells make these expression cassettes by joining one randomly selected heavy chain Diversity (D) gene to one random Joining gene (J) via *recombination activating genes 1 & 2* (RAG1 and RAG2) enzymes to generate pro-B cell recombined DJ_H DNA, followed by Variable (V) to Diversity-Joining ($V\text{-}DJ_H$) gene recombination using the same molecular machinery. The resulting pre-B cell variable heavy chain cassette (VDJ_H) has inherent diversity owed to the random V-D-J combinatorial gene selection process, and may be diversified further depending on how the regions were re-joined; if the regions join in frame (no gaps i.e. imprecise joining, deletions i.e. exonuclease trimming, and/or insertions i.e. TdT nucleotide addition that result in out-of-frame protein translation) the genes are said to be productively recombined. Should the recombined B cell receptor DNA result in a non-productive rearrangement that cannot be recovered through allelic inclusion, the immature B cell undergoes programmed cell death, or apoptosis, and is cleared from the bone marrow.

Productive rearrangement of the heavy chain variable genes induces expression of a surrogate light chain and rearrangement of the light chain variable genes, which proceeds in a similar manner as the heavy chain; productive joining of Pre-B cell light chain variable genes ($V_L \rightarrow VJ_L$) produces an immature B cell committed to expressing that VDJ_H : VJ_L paired heavy:light antibody sequence and resulting antigen specificity (*55*). Recall the constant region of the light chain is encoded by two classes: Kappa (IgK) or Lambda (Igλ) encoded on separate gene loci. Through a process of allelic inclusion, a Pre-B cell can attempt to produce a light chain cassette with in-frame V-J-C($_{\text{Kappa or Lambda}}$) DNA to pair with its heavy chain variable

(VDJ_H) DNA cassette during antigen independent development. Heavy chain constant region DNA rearrangement occurs in the periphery, outside the bone marrow, to produce tissue-specific, isotype defined antibody functions.

Following productive recombination, translation, and expression of immature B cell DNA cassettes VDJ_H : VJ_L as a surface-bound IgM : IgKorλ antigen-receptor, the immature B cell is assessed for non-self-reactivity at immune tolerance checkpoints before it is permitted to exit the bone marrow and search for cognate antigen as a naïve B cell. Since naïve B cells have not rearranged their heavy chain constant region DNA and there is no stop signal between Mu and Gamma heavy chain constant region genes, these cells produce both IgM and IgG class antibodies depending on splicing signals determined at the RNA level (Fig. 5). Upon recognition of cognate antigen via its BCR, a naïve B cell becomes "antigen-experienced" and migrates to a germinal center to begin antigen-dependent development (Fig. 4 and 5).

Using activation induced cytidine deaminase (AID) enzyme and the help of antigen presenting cells, antigen-experienced B cells begin a process known as clonal expansion in which cells divide and produce generations of somatic hypermutation-induced mutated BCR genes that selectively enhance recognition and affinity of target antigen. In this process, these random AID-induced point mutations accumulate and enhance the resulting protein's binding affinity to its cognate antigen. Mutations which increase affinity are preferentially selected and expanded over mutations which result in lower affinity antigen:receptor interactions, resulting in an expanded cell population with increased antigen recognition capabilities. In addition to optimization of binding affinity, antibody isotype class switching can occur via rearrangements in the heavy chain DNA mediated by AID to position the variable VDJ_H segment adjacent to a specific constant region downstream the DNA region encoding IgM and IgG. This allows expression of mature antibodies with isotypes distinct from the previous naïve B cell pre-cursor DNA.

Antigen experienced B cells comprised of memory B cells, short-lived plasmablasts, and long-lived plasma cells help mediate humoral immune responses. Memory B cells developed in germinal centers patrol

the peripheral body long-term for the next time antigen exposure occurs to help rapidly initiate an antibody-mediated immune response faster than naïve exposure alone. Plasmablasts and plasma cells are antibody secreting cells forming humoral immunity by producing large quantities of IgA/D/E or G depending on the constant region recombined and expressed downstream from the variable region genes. Isotype expression determines the tissue-specificity of antibody-mediated humoral immunity allowing the body to direct an immune response to afflicted compartments. Humoral immunity against a given pathogen is typically polyclonal i.e., generated by the activation of many different B cells and resulting in a complex mixture of distinct antibodies recognizing multiple antigenic epitopes to aid in pathogen elimination. While these polyclonal features of natural antibody repertoires produce robust immune recognition and clearance of a wide variety of antigens, the mixed display of phenotypes make it difficult to determine unique contributions from single antibody proteins. Evaluation of individual antibody proteins requires single B cell analysis to deconvolute individual phenotypes.

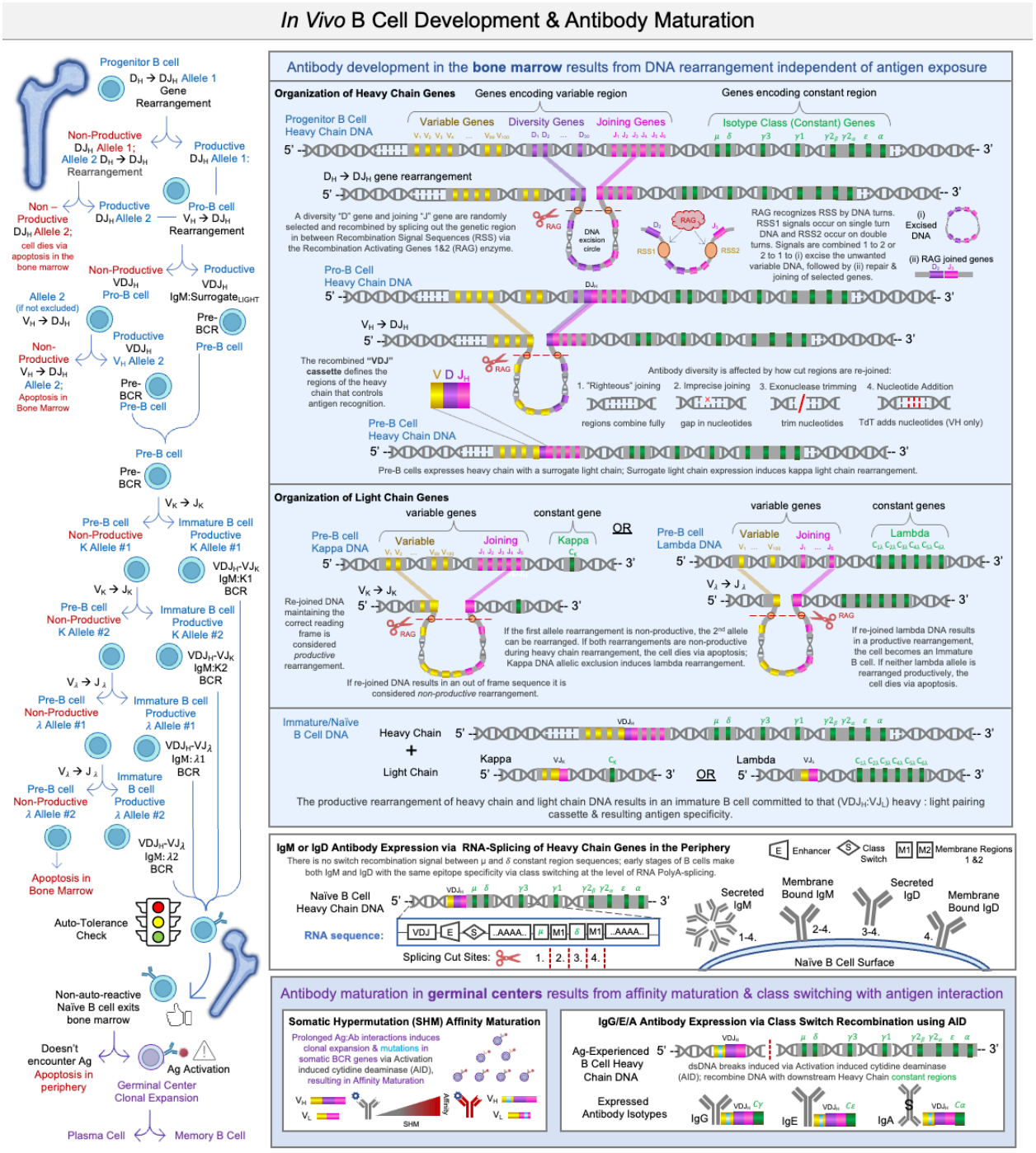


Figure 5. Detailed overview of antibody gene organization and expression resulting from B cell development and maturation.

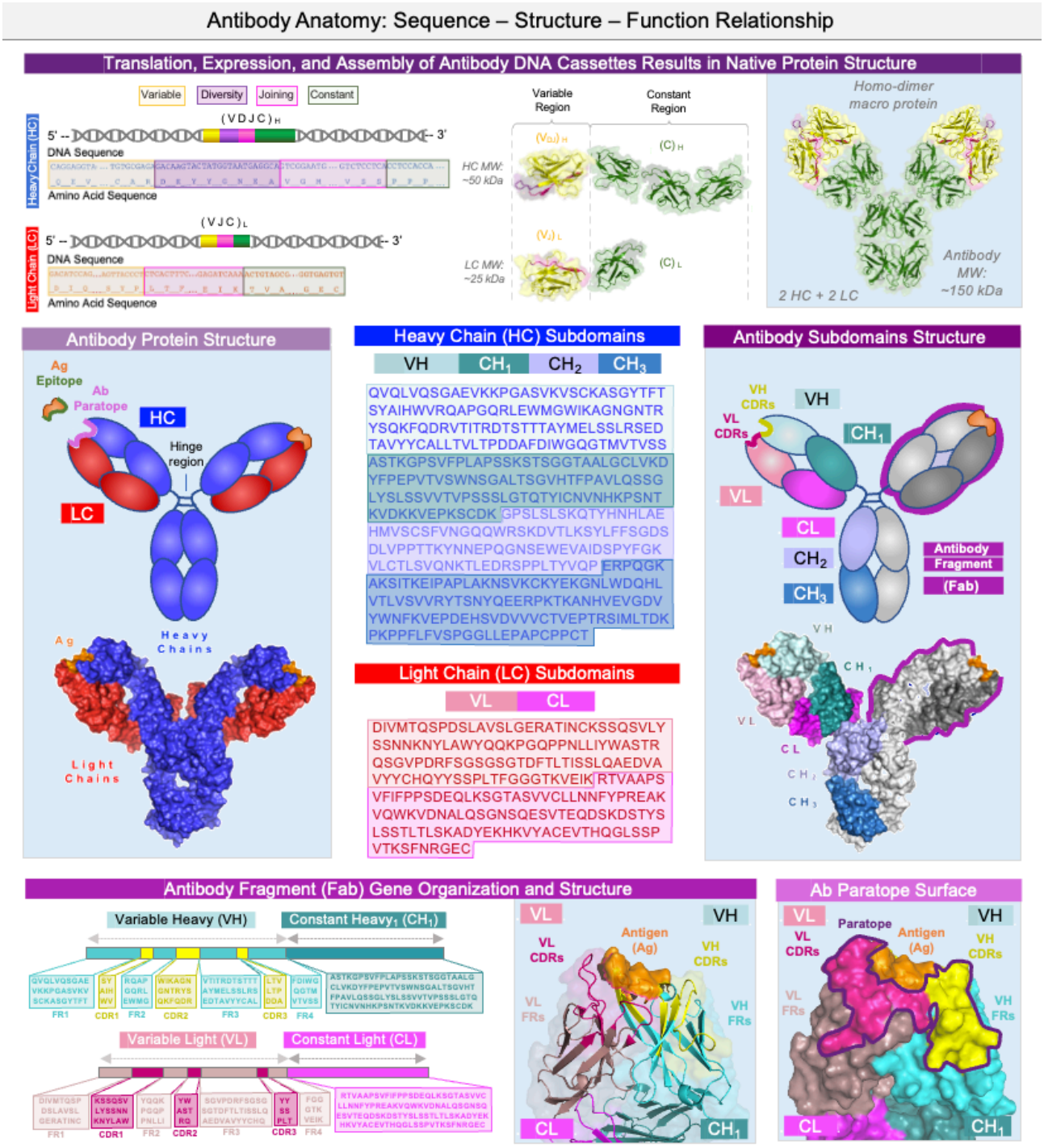


Figure 6. Overview of the relationship between antibody DNA sequence and protein structure leading to antigen recognition defining this unique class of immune protein function.

1.1.3 Molecular Biotechnologies for Monoclonal Antibody Engineering

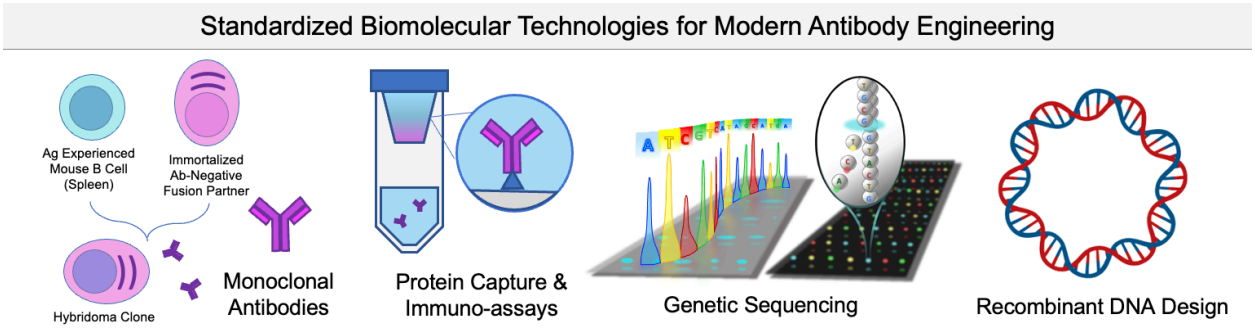


Figure 7. Monoclonal antibodies have been used to generate new insights into adaptive immunity through the application of standardized technologies such as immunoassays, genetic sequencing, and recombinant DNA design.

Hybridomas & monoclonal antibodies. Prior to the invention of monoclonal antibodies, most antibody-based translational products were derived from polyclonal, serum populations with multi-specificity resulting in unwanted secondary side effects. Reproducible methods to interrogate single antibody proteins were limited; individual B cells expressing antibodies of interest could be sequestered from biological samples, however these cell cultures expired quickly and could not be repeatedly interrogated. To address these issues, Georges Kóhler and Cesar Milstein invented a new technique in the 1970s involving the hybridization of murine (mouse) B cells and human tumor cells, creating *hybridomas* (-oma, from tumor, awarded a Nobel prize in 1984) (*6*). The invention of hybridoma technologies allowed for the immortalized isolation of antibodies arising from a single murine B-cell clone in culture, i.e., *monoclonal antibodies*. To form a hybridoma, a mouse would be injected with an antigen of interest and antigen-activate B cells would be isolated from harvested spleens, and then fused with an artificial myeloma tumor cell with immortalized function. The artificial myeloma cell had been previously modified via tissue culture so that it no longer produced its own cognate antibody, making suitable *fusion partners* for B cells encoding an original antibody of interest. Cell membrane fusion using polyethylene glycol (PEG) and selective culturing practices to remove non-hybrid cells produced hybridoma cells carrying the

immortalized function of the myeloma, as well as the selection-media survival and antibody production capabilities of the B cell. Individual hybridomas could be isolated and selected for further molecular characterization of antibody phenotypes.

Antibody protein isolation. Hybridomas mark a significant breakthrough in the development of antibody biologics, allowing for powerful single cell molecular interrogation of the sequence-structure-function relationship defining antibody recognition. Hybridoma libraries containing hundreds, or thousands of individual antibodies of interest could be screened *en masse* to quickly identify positive clones of interest using multi-well plates or microfluidic formats. Selected clones could then be amplified and grown in small volume formats or large-scale fermentation vats for multiple liters of protein culture. Large quantities of biologically produced antibody protein material can be isolated from liquid cell culture media using selective antibody-capture and/or liquid chromatography methods (*56*). Commercial and therapeutic application of monoclonal antibody (mAbs) biomaterials has led to the development of a variety of immune-based products with tightly regulated pharmacokinetic and pharmacodynamic profiles leading to their efficient production, function, and recycling in the body. All together these molecular characteristics result in a powerful immune-technology with high target specificity & controlled drug-ability. Since their invention, mAbs have been used in immunoassays to detect trace antigen, as drugs and drug carriers for targeted treatments, and for passive clinical immunization against infectious diseases.

Antibody sequencing and recombinant DNA design: Modern antibody engineering was further expanded by the advent of high-throughput genetic sequencing and bioinformatic analysis techniques (*9-11*). Preliminary genetic analysis of antibody sequences could be accomplished in limited batches using capillary electrophoresis methods, such as Sanger Sequencing, to decode antibody sequences at a throughput of 10^2-10^3 individual cells at a time (*16*). However, there was an enormous cost of time and human capital required to extract this information, and the low number of cells analyzed with these methods failed to capture repertoire-level phenomenon, creating a bottleneck for advanced molecular antibody research. Methods to decode antibody sequences at greater throughput were enabled by novel *sequencing*

by synthesis technologies resulting in *next generation sequencing* (NGS) methods (*23, 24*). Using NGS, antibody researchers can now determine B cell mediated pathology and molecular correlates of humoral immunity for millions of cells at an unprecedented depth for a fraction of the time and costs (*16, 18-20*).

Despite the launch of antibody research through hybridoma enabled insights, there were distinct disadvantages to therapeutic application of hybridoma-derived mAbs, as their murine B cell origins resulted in antibody drug products that were considered foreign and immunogenic in human applications; repeat injections of these drugs would eventually lead to an immune response and rejection, limiting therapeutic efficacy. Human antibody hybridoma production was restrained by a handful of factors: the same range of antigens that can be experimentally exposed to a mouse to generate diverse B cell responses cannot be tested in humans, there are no safe and adequate methods for isolation of activated antibody producing B cells from human spleens, and reversed human B cell → mouse myeloma hybrid fusion productions produced unstable derivatives.

Moving away from the idea of isolating and producing human antibodies through fusion with human B cells, modern antibody engineers have invented *in vitro* methods using *recombinant DNA* technologies to synthesize native monoclonal antibody sequences with enhanced drug-ability functions for translational protein design (*7*). A circular DNA vector that has been molecularly modified to include a "foreign" recombinant DNA sequence can be used to store genetic material for various applications. Circular recombinant antibody DNA vectors comprise a contiguous artificial DNA sequence formed by constituents from different organisms; DNA vectors from plasmid or virus sources suitable for recombinant antibody design must have (1) a cloning insertion locus defined by restriction sites for cloning in recombinant antibody genes (2) the ability to replicate independently, (3) selectable markers, and (4) a small size (10^3 bp magnitude).

Such vectors are used in molecular antibody cloning as a vehicle to artificially carry antibody genetic material into another cell, where it can then be replicated and/or expressed in a controlled manner.

The ability to create genetically modified cells expressing antibody proteins by cloning recombinant antibody DNA – isolating the antibody DNA of interest from any organism, ligating it to pre-selected vector DNA of choice, introducing that recombinant vector into host cells, and propagating those cells to promote recombinant antibody protein expression – represents an alternative bench-top-enabled approach to hybridoma technologies allowing researchers detailed exploration into human antibody genetics, molecular physio biology, and B cell functional development. Next generation sequencing paired with these recombinant monoclonal antibody generation methods has become an increasingly sophisticated approach used in antibody engineering research to study antibody repertoires with precise molecular detail (*57*).

1.2 Translational Antibody Engineering for Clinical Application

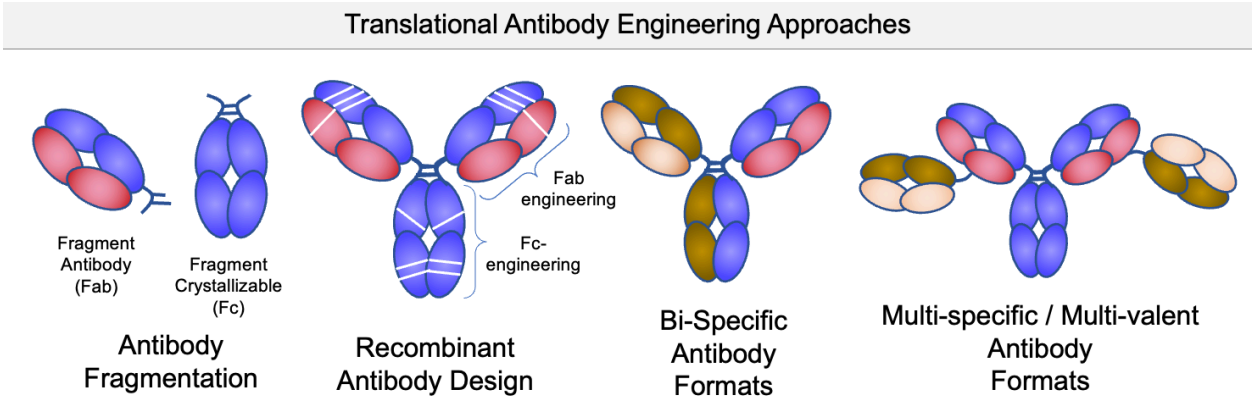


Figure 8. Engineered adaptations to natural antibody structure for advanced translational insights and applications.

There are many ways in which antibodies may be researched, characterized, and functionally applied in clinical applications (*58*). With the advent of standardized antibody protein isolation techniques came proteolysis experiments to dissect antibody function according to its structural components. Antibody fragments and the functional applications of these proteins have been studied for decades, using a variety of proteolysis and genetic engineering to produce mono or multivalent antigen / receptor binding fragments, such as the antibody fragment antigen binding (Fab region) and the fragment crystallizable / constant, or Fc, region (*59*). Through these studies it was determined the antibody variable region holds the key to antibody specificity, yet when expressed independently without the rest of the antibody structure, the variable region alone tends to have deficits in stability. In the late 1980s a single-chain fragment variable (scFv) format was invented by artificially introducing a peptide linker between the individual VH-VL domains (*60, 61*). The tunable drug features of scFv format, such as controlled multimerization and small molecular weight, quickly made these fragments critical building blocks for antibody research and potential molecules for targeted medicines. The smaller scFv format is particularly favored for applications requiring high targeting specificity and a low molecular weight to target hard-to-reach, cryptic epitope features (*62*).

Despite the many advantages of antibody fragment-based proteins, one of the main drawbacks for therapeutic applications is the lack of an Fc domain connected to antigen specificity. While much attention

is given to the variable region for its role in antigen recognition, the antibody Fc plays a critical role in how the body interacts with the antibody protein via binding to numerous tissue specific-receptors in the body (*54, 59, 63*). For example, the Fc region can bind to FcRn receptors that mediate antibody recycling in the blood stream, extending the overall half-life of the antibody protein in circulation up to weeks as opposed to days. In addition, the Fc region plays an important role binding to a C1q to recruit and activate immune compliment proteins. Mutant evaluation of the Fc region has revealed the impact of specific residues on various effector functions while detailing the impact of Fc glycosylation on *in vivo* antibody function (*64*).

A key antibody growth area that has developed in the 21st century includes bispecific, multi-specific, and multivalent antibodies (*65-70*). These antibodies combine two or more antigen specificities into a single molecule or achieve a higher degree of avidity through the additional of more binding sites to a single antigen. Bispecific and multi-specific antibodies do not occur in nature, thus studying these types of proteins require artificial methods for antibody synthesis. With the advent of monoclonal antibody production and characterization methods in the 1970s came a rising interest in multi-specific antibodies as potential therapeutics. For example, in the 1980s a method for producing bispecific antibodies was produced, described as a "hybrid" hybridoma *quadroma*, but due to the random pairing of heavy and light chains a high frequency of mis-matched non-native antibody pairings resulted in only 12.5% yields of the desired antibody product in the total theoretical population produced (*68, 70*). These low production yields created a bottleneck in translation for advanced applications requiring large quantities of protein. Recombinant DNA technologies (Fig. 7) allowed for the direct invention of single-sequence antibody molecules encoding multiple antigen specificities, such as bi-specific and multi-specific/multi-valent formats, while also allowing unique molecular effector functions through the inclusion of an Fc region. These recombinantly-designed formats offer translational insights into dual-antigen targeting applications; however, their complex pharmacokinetics and expression formats limit clinical translation (*65-67*). Non-native VH:VL antibody pairings tend to have stark decreases in drug-ability as compared to natively paired antibody sequences derived from *in vivo* functional development, indicating a translational advantage of natively paired antibody heavy:light chain antibodies (*71*).

1.3 Translational Antibody Engineering for Precision Medicine Design

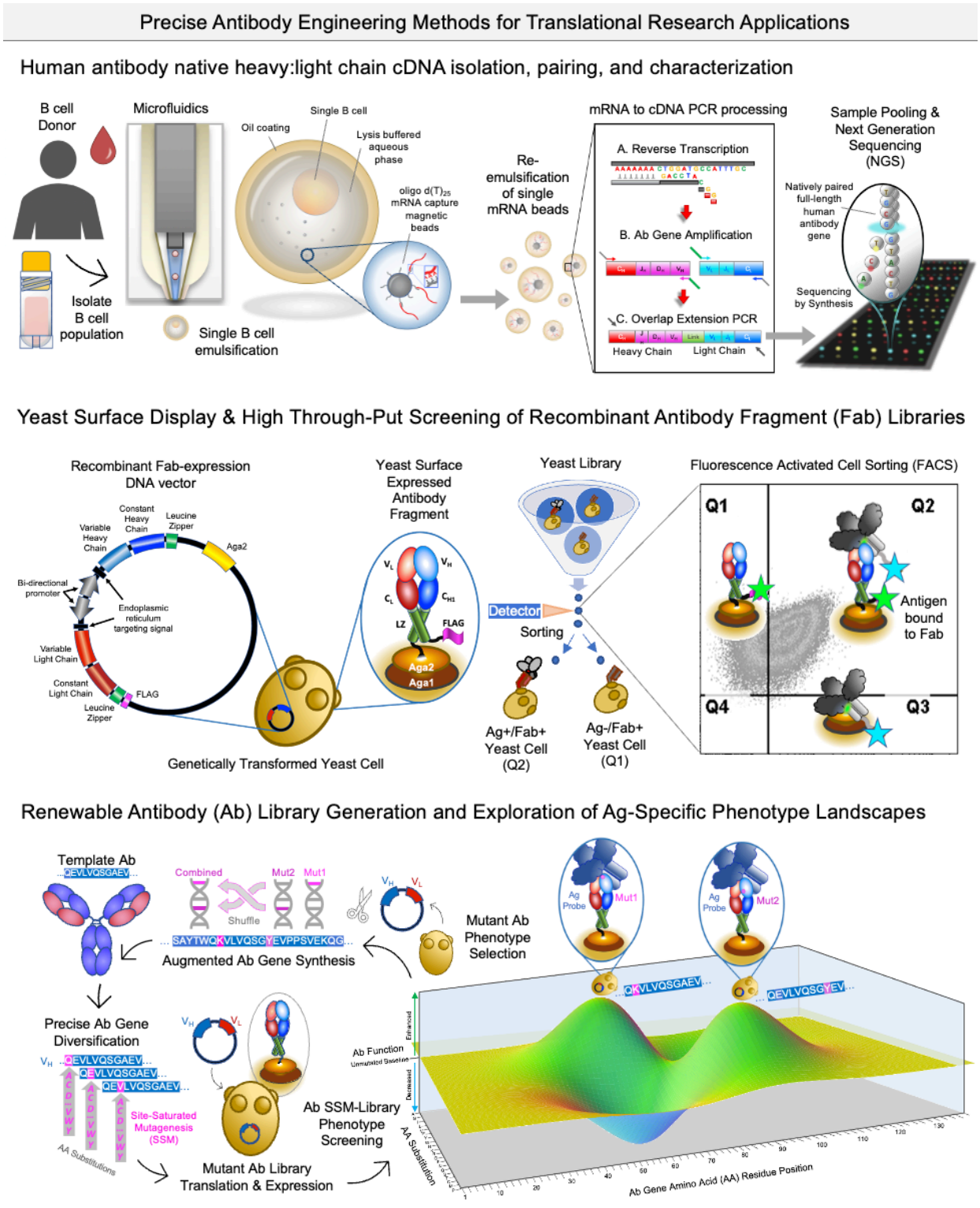


Figure 9. Advanced 21st century biotechnologies for antibody engineering.

As modern immune engineering practices built from decades of research have become standardized to develop new medicines, diagnostics, agricultural products, and forensic tools from antibody molecules, there have also been new bottlenecks identified along the way highlighting the limitations of antibody development using these standard practices alone (*72-75*). Limiting factors for antibody molecules used for therapeutic applications include risk of immune rejection and suboptimal drug-ability associated with non-natively paired or non-human molecules, lowering the number of antibody insights that can translate into clinical application; in addition, a lack of molecular understanding related to disease pathology define complex challenges for eliciting protective antibody-mediated solutions that can effectively mitigate disease. Recently emerged immune engineering technologies, including natively-paired antibody heavy:light chain complementary DNA libraries, *in vitro* functional antibody fragment antigen binding (Fab)-library display, high-throughput library screening, and targeted antibody library design, can be used to overcome these bottlenecks and promote new antibody research for precision medicine design (Fig. 9) (*18-20, 57, 76-81*).

Generating natively paired antibody sequences requires nontrivial molecular processing; the functional antibody repertoire is encoded by co-expressed heavy and light chain genes derived from separate genetic loci which presents physical barriers for paired heavy:light genetic analyses. Using microfluidic emulsification of single B cells in a lysis-buffered aqueous solution, actively transcribed B cell mRNA encoding natively paired antibody heavy:light variable genes (individual VDJ_H and VJ_L polyadenylated mRNA) can be captured onto magnetic beads and amplified using antibody-specific DNA primers and overlap extension RT-PCR to link the separately expressed native genes in a single cDNA amplicon. These techniques can be applied for in-depth functional interrogation of natively paired human antibody libraries on the order of millions (10^6-10^7) of cells per sample. This large scale natively paired antibody analysis can shed new light on *in vivo* antibody functions, define new biomolecular disease intervention strategies, and increase the likelihood of finding biocompatible drug candidates with potential for clinical applications.

Human immunoglobulin expression platforms can be used to generate gene libraries encoding diverse antibody sequences for *de novo* functional antibody discovery. Libraries of natively paired antibody sequences can be cloned into recombinant surface-expression vectors. Transformation of these vector libraries into controlled cell expression systems provides a platform to interrogate the genetic sequence and antibody performance on a library scale *en masse*. One method for *in vitro* protein library screening includes yeast surface display, in which genetically modified yeast cells express antibody fragment antigen binding (Fab) proteins on their surface (*76*). Yeast cells offer many functional advantages for experimental screening, as they have a rapid cell division cycle (2-8 hours), are easy to grow, and have been used to clone many eukaryotic genes with standardized molecular biology methods. By co-staining Fab-expressing yeast cells with fluorescent markers to measure Fab surface expression and fluorescently labeled antigen probes to measure antigen-binding of surface expressed Fab, advanced high-throughput fluorescence activated cell sorting (FACS) techniques can fractionate antibody library populations according to antigen binding avidity phenotypes (*20, 77, 78*).

These *in vitro* library generation and screening methods enable researchers to generate immortalized antibody libraries that can be repeatedly screened for recognition of diverse antigen targets. Banks of antibody sequences can be repeatedly probed to enhance antibody discovery rates and further study functional phenotype using DNA mutagenesis and directed evolution methods that measure how variations in the antibody amino acid sequences influence antibody binding affinity. Characterization of library screens yields paired antibody heavy:light sequences that can be used to produce novel biologics for the next generation of clinical research and precision medicines.

1.4 Chapter 1 Synopsis

A history of infectious diseases afflicting human health has motivated the development of new biotechnologies to help slow and prevent the spread of microbial agents. Over time, biotechnologies have become more precise as our understanding of disease mechanisms has increased. More recently, advanced immune technologies and antibody-based precision medicines have led to enhanced clinical health outcomes. Antibodies are a unique class of immune proteins capable of recognizing antigen foreign bodies with high precision, resulting in highly regulated, specific, humoral immune protection from disease. Characterization of natural human responses derived from *in vivo* infections can provide molecular insights and new antibody molecules related to disease pathogenesis. Antibodies derived from these studies can be adapted for targeted immune functions using recombinant engineering approaches, translating these molecules into new precision medicines. However, clinical application of this new knowledge is limited by outstanding scientific questions surrounding fundamental antibody-mediated immune mechanisms. Insights gained from precise antibody engineering experiments can be leveraged to answer questions surrounding fundamental antibody mechanisms, generating biomolecular blueprints for novel antibody-based solutions to clinical needs.

To achieve outstanding global health goals related to the prevention of infectious diseases, contemporary antibody engineering uses fundamental knowledge of microbial pathogenesis and human immunity to design new strategies which can combat, prevent, and/or treat infections with high efficacy. This discipline has developed in parallel to major growth in the studies of immunology, virology, pathology, structural biology, pharmacy, and bioengineering, bolstered by the development of new high-throughput -omics biotechnologies. Translational precision antibody engineering methods can be used to determine protective mechanisms of action at an unprecedented level of detail to aid in the design of next generation precision medicines.

2. Paired heavy and light chains contribute to potent SARS-CoV-2 neutralization in public antibody responses[1]

2.1 Introduction and Rationale: Antibody Discovery

Directly identify protective antibodies from convalescent COVID-19 patients and define epitope-specific molecular features associated with disease protection by rapidly mapping antibody responses to diverse viral antigens using renewable yeast-surface display technologies. As with other human viral infections, potent neutralizing mAbs can be mined from convalescent SARS-CoV-2 patients (*82-85*). Understanding mechanisms of protective antibody recognition can inform vaccine and therapeutic strategies against SARS-CoV-2. Here we hypothesize that neutralizing mAbs against both known and hypothetical protective epitopes can be mined from COVID-19 convalescent patients with high serum neutralization titers against SARS-CoV-2 to reveal molecular features associated with SARS-CoV-2 protection. Our study produced the discovery of a monoclonal antibody, 910-30, targeting the SARS-CoV-2 receptor-binding site for ACE2 belonging to a public antibody response encoded by IGHV3-53/IGHV3-66 genes. Sequence and structural analyses of 910-30 and related antibodies explore how class recognition features correlate with SARS-CoV-2 neutralization. Cryo-EM structures of 910-30 bound to the SARS-CoV-2 spike trimer reveal binding interactions and its ability to disassemble spike. Despite heavy-chain sequence similarity, biophysical analyses of IGHV3-53/3-66-encoded antibodies highlight the importance of native heavy:light pairings for ACE2-binding competition and SARS-CoV-2 neutralization. We develop paired heavy:light class sequence signatures and determine antibody precursor prevalence to be ~1 in 44,000 human B cells, consistent with public antibody identification in several convalescent COVID-19

[1] Bailey B. Banach, *et al*. Paired heavy- and light-chain signatures contribute to potent SARS-CoV-2 neutralization in public antibody responses, *Cell Reports*. Volume 37, Issue 1, 2021, https://doi.org/10.1016/j.celrep.2021.109771. B.B.B., G.C., A.S.F., C-H.S., S.N.L-A., K-T.Y, T.A.W., D.D.H., P.D.K., L.S., and B.J.D., designed the experiments; B.B.B., G.C., A.S.F., C-H.S., M.O., P.K., Y.T., P.W., M.S.N., Y.H., I.F., P.J.S., L.L., S.N.L-A., A.N., J.R.W., Y.L., X.P., B.M., A.D.L. and R.M. performed the experiments; A.S.O., I-T.T., J.Y., T.Z., E.R., and J.B. provided reagents for experiments, B.B.B., G.C., A.S.F., C-H.S., P.K., I.F., P.J.S., M.G-G., B.M., S.N.L-A., X.P, and B.J.D. analyzed the data; and B.B.B., G.C. A.S.F., C-H.S., P.D.K. L.S, and B.J.D. wrote the manuscript with feedback from all authors.

patients. These class signatures reveal genetic, structural, and functional immune features that are helpful in accelerating antibody-based medical interventions for SARS-CoV-2.

Background

The highly infectious nature of SARS-CoV-2 and significant prevalence of severe disease has caused immense global, social, and economic disruption (*86, 87*). SARS-CoV-2 marks the third known emergence of a novel beta-coronavirus in the past two decades, following its closest documented human pathogen severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002, and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 (*88-93*). Both SARS and SARS-CoV-2 infect human cells by binding the angiotensin convertase II receptor (ACE2) via the trimeric spike (S) class I fusion protein (*94, 95*). S protein comprises two subunits, S1 and S2. The S1 subunit contains a receptor-binding domain (RBD), which binds to ACE2. To enter cells, S undergoes a protease cleavage event that allows S1 to shed and expose the hydrophobic fusion peptide of the S2 subunit. SARS coronavirus predominantly enters cells via endosomes, assisted by cathepsin cleavage in the low pH (5.5-4.5) endosomal environment. SARS-CoV-2 acquired a new protease cleavage site that enables entry either at the cell surface after cleavage with TMPRSS2, or inside endosomes via protease cleavage similar to SARS, and the route of SARS-CoV-2 entry is likely dependent on the protease expression profile in target cells (*96, 97*). ACE2 interactions appear to play a role in pre-fusion S1 shedding (*98, 99*).

A detailed understanding of SARS-CoV-2 antibody neutralization should help to accelerate progress in medical interventions. Antibodies from several COVID-19 patients revealed public antibody responses against SARS-CoV-2 via shared genetic and structural elements in IGHV3-53 and IGHV3-66 heavy chain V-genes. This public antibody class targets a conserved RBD epitope in the S1 subunit that overlaps with the ACE2 binding site (*82, 84, 100-110*). IGHV3-53/3-66 public class antibodies share common genetic features including IGHV-gene-encoded motifs NY in the CDR-H1, SGGS in the CDR-H2, a relatively short CDR-H3 length, and comparatively low levels of antibody somatic hypermutation (*100, 103, 110, 111*). Both kappa and lambda light chains are represented in antibodies of this class (*103, 109, 110, 112-114*).

Despite heavy chain gene similarities, IGHV3-53/3-66 anti-RBD antibodies show a broad range of neutralization potencies (IC_{50}'s from 0.003 to 2.547 μg/mL), (*82-84, 101, 106, 108-110, 115, 116*). Given the low somatic hypermutation observed and the importance of germline-encoded recognition motifs, it remains unclear what unique molecular features lead to the diverse range of SARS-CoV-2 neutralization potencies among class members.

SARS-CoV-2 S displays a pH-dependent conformational switch that causes the 'up' position of the RBD to rotate to a 'down' position (*117, 118*). The RBD 'up' position is required for ACE2 engagement and for IGHV3-53/3-66 class antibody binding (*103, 114, 117*). A mutational variant, D614G, also emerged that influences the RBD "up" vs. "down" state and has constituted >97% of isolates world-wide (*119-123*). D614G is proximal to the RBD, and D614G appears to favor more RBD 'up' at both serological and endosomal pH (*118, 122, 124*). The D614G substitution enhances viral infectivity, competitive fitness, and transmission (*122, 123, 125, 126*); further information related to the effects of D614G on IGHV3-53/3-66 class recognition and neutralization is needed (*124*).

Here we discover a new member of the IGHV3-53/3-66 antibody class, mAb 910-30, with moderate neutralization capacity. To understand the features of potent IGHV3-53/3-66 class neutralization, we explore molecular and genetic features of 910-30 and other related antibodies, including heavy and light chain structural recognition motifs, biophysical correlates of neutralization, and the influence of D614 vs. D614G variants on IGHV3-53/3-66 class member interactions. Our study provides a detailed molecular understanding of how the public IGHV3-53/3-66 class leverages native heavy and light chain binding interactions, providing important data to accelerate medical interventions against the SARS-CoV-2 RBD epitope.

2.2 Results

Isolation and structural characterization of a novel neutralizing class member

We screened the immune repertoire of a COVID-19 convalescent patient, Donor 910 (*127*), to identify a new member of the public IGHV3-53/3-66 antibody class. Donor 910 serum ELISA showed potent S trimer recognition, and pseudovirus neutralization assays confirmed potent SARS-CoV-2 serum neutralization (**Appendix A. Fig. 14**) (*128*). Donor 910 cryopreserved peripheral blood mononuclear cells (PBMCs) were selected for analysis using a recently described method to clone natively paired heavy and light chain antibody variable regions into yeast display for functional screening (*20*). Yeast antibody display libraries were screened for binding against SARS-CoV-2 S protein probes by fluorescence-activated cell sorting (FACS) (**Appendix A. Fig. 14**). Bioinformatic interrogation of yeast display NGS data revealed monoclonal antibody 910-30 enriched 90-fold in the Round 2 sorted library against biotinylated S trimer, and 2,296-fold in Round 3 sorted libraries against His-labeled S trimer. Based on its strong enrichment, mAb 910-30 was expressed as IgG in HEK293 cells for neutralization assays. 910-30 showed a half-maximal inhibitory concentrations (IC_{50}) of 0.071 μg/mL against a VSV SARS-CoV-2 pseudovirus (*84*), and 0.142 μg/mL against authentic SARS-CoV-2 (**Fig. 10A**) (*110, 129, 130*).

910-30 structural recognition was next characterized by cryo-electron microscopy (cryo-EM). Negative-staining electron microscopy revealed particles of 910-30 fragment antigen-binding (Fab) bound to SARS-CoV-2 S2P at pH 5.5 (**Fig. 10B**) (*118*), and across a broader pH range of 4.0-7.4 (**Appendix A. Fig. 14**). Subsequent cryo-EM mapping and molecular modeling of 910-30 Fab in complex with SARS-CoV-2 S2P protein at pH 5.5 showed 1 Fab bound to 1 RBD in the up position when mixing Fab and spike in a 1:1 molar ratio (**Fig. 10C**, **Appendix B. Fig. 15, Appendix E. Supplemental Table 1**), whereas a 9:1 Fab:spike molar ratio revealed mostly disordered spike (**Fig. 10D**, **Appendix B. Fig. 15, Appendix E. Supplemental Table 1**), with an RBD that still fit the cryo-EM map consistent with **Figure 10C**. Structural modeling of ACE2 (PDB entry 6M0J) and 910-30 (PDB entry 7KS9) in complex with SARS-CoV-2 RBD confirmed ACE2 binding site recognition (**Fig. 10E**). Binding analysis against yeast-displayed aglycosylated N343Q RBD(333-537) confirmed that 910-30 recognized a glycan-independent epitope

(**Appendix C. Fig. 16A**) (*131*). Antibody titrations showed a 910-30 IgG K_D to RBD of 230 pM (191 - 268 pM 95% confidence interval) (**Appendix C. Fig. 16B**), and that 910-30 competes with human ACE2 (hACE2) for binding RBD, like other IGHV3-53/3-66 class members (**Appendix C. Fig. 16C**).

Potent antibodies of the IGHV3-53/3-66 class compete strongly with ACE2 for binding to Spike

Structural analysis of IGHV3-53/3-66 antibody recognition shows substantial overlap between the ACE2 binding site and the shared class epitope (**Fig. 11A**) (*82-84, 100, 104, 108, 111, 117*). Despite low reported somatic hypermutation and common epitopes, IGHV3-53/3-66 antibody class members still show a broad range of neutralization potencies (**Fig. 11B**, **Appendix F. Supplemental Table 2**). To better understand molecular features of potent antibody neutralization in this class, we assessed biophysical performance of a panel of weak, moderate, and potent IGHV3-53/3-66 class members. We selected: 1-20 (a potent neutralizer), 910-30 (a moderate neutralizer), and B38 (a weak neutralizer), along with a VH-gene matched control (mAb 4-3) that neutralizes poorly and likely targets a different site on RBD (**Appendix G. Supplemental Table 3**) (*109, 118*). Preliminary IgG ELISA revealed that the more potent class members (1-20, 910-30) bound more tightly to full-length spike but similar binding to RBD (**Fig. 11C**). Pseudovirus and authentic virus neutralization assays showed two orders of magnitude across IC_{50} neutralization potencies for the panel (**Fig. 11D**), confirming that antibody neutralization within the class is driven by more than simple recognition of the ACE2 binding site on RBD.

We next assessed the ability of potent and weak neutralizers to compete with dimerized human ACE2 (dhACE2) for binding to spike (S2P). In a competition ELISA with pre-mixed antibody and dhACE2, we found that more potently neutralizing class members competed more strongly with dhACE2 than less potent Abs (**Fig. 11E**). The most potent mAb (1-20) required 73 dhACE2 molecules for 50% binding inhibition of 1 IgG molecule. 1-20 was 6-fold more competitive with dhACE2 than the moderate neutralizer 910-30 (dhACE2 molar excess IC_{50} = 12), and 150-fold more competitive than B38 (dhACE2 molar excess IC_{50} = 0.48). Sequence analysis confirmed high sequence similarity and low levels of somatic hypermutation for the panel, as previously reported for the IGHV3-53/3-66 antibody class (**Fig. 11F**) (*105, 110*).

Unique heavy and light chain interactions drive potent neutralization for the IGHV3-53/3-66 neutralizing antibody class

Given the broad variations in potency despite high sequence similarity, we next sought to understand the features associated with potent SARS-CoV-2 neutralization. IGHV3-53/IGHV3-66 anti-SARS-CoV-2 antibodies show diverse light chains, with the two defining heavy chain genes (IGHV3-53 and IGHV3-66) pairing with at least 14 light chain genes (**Appendix F. Supplemental Table 2**). We constructed a panel of 12 non-native heavy-light swapped antibody variants to explore light chain pairing features using four IGHV3-53/3-66-encoded mAbs (1-20, 910-30, B38, and 4-3 included as an IGHV gene control). 11 out of 12 non-native antibodies expressed successfully and were assayed for SARS-CoV-2 pseudovirus neutralization (B38 heavy:1-20 light failed to express). Heavy-light swaps revealed substantially reduced neutralization for nearly all non-native heavy-light combinations, with only the most potent antibody heavy chain (1-20) achieving significant neutralization with a non-native light chain (**Fig. 12A**). More potent neutralization was correlated with strong dhACE2 competition (**Fig. 12B**), consistent with natively paired antibody performance (**Fig. 11E**). As all four heavy chain genes have low somatic hypermutation & high sequence similarity (**Fig. 11F**, **Appendix G. Supplemental Table 3**). **Figure 12A** shows that native light chains make substantial contributions to neutralization performance, and that non-native pairings reduced antibody performance, despite conserved heavy and light sequence signatures. To explore structural features, we generated *in silico* models of 910-30/B38 heavy:light swaps by superimposing the light chains and replacing the native light chain. Structural alignment of the 910-30/B38 swaps on RBD showed significant CDR-L3 displacement, especially at the heavy:light interface. Changes in the heavy:light chain contact region can affect the orientation of all CDRs, resulting in sub-optimal binding geometry for non-natively paired heavy:light swaps. These *in silico* data suggested that native heavy:light chain assembly was important for proper CDR orientation and protective RBD recognition in this antibody class.

To better understand determinants of antibody recognition, we next performed a structure-based alignment to analyze interactions and identify light chain signatures of class membership (*24, 132*). We followed numerous anti-RBD IGHV3-53/3-66 antibody lineages natively paired with different light chain

V-genes, including KV1-33, KV1-9, KV1-39, KV3-20, and LV2-8 (*100, 105, 108, 110, 111*). Structural analyses of antibody contact sites revealed that conserved residues in VH and VL genes contributed 56-75% of binding surface area (BSA) (**Fig. 12C and Appendix C Fig. S16D**). **Appendix C. Supplemental Figure 16D** shows the IGHV3-53/3-66 heavy chain projected surface with germline-encoded amino acids, including interaction residues ^{31}SNY on CDR-H1, and ^{52}YSGxSxY on CDR-H2 (x indicates any residue) that provide multiple hydrogen bonds interactions with Thr415, Gly416, Lys417, Asp420, Tyr421, Leu455, Tyr473, Ala475, and Asn487 in RBD. We verified previous reports that class members have shorter 6-11 amino acid CDR-H3s, and that key CDR-H1 and CDR-H2 motifs are only present in IGHV3-53/3-66 genes. Sequence-structure alignments revealed that kappa chain class members use a conserved [DGS]xSx(1,2)[FY] motif of 11 or 12 amino acids starting at residue 27a or 28 in the CDR-L1 to form hydrogen bonds with RBD residues Gln498 and Asn501 (**Fig. 12C**). In contrast, lambda chains in the class use a different ^{29}GY[KN] motif with 14 amino acids in CDR-L1 that interacts with RBD residues Gly502 and Tyr505 (**Fig. 12C**). We defined the $^{27a/28}$[DGS]xSx(1,2)[FY] motif of 11 or 12 amino acids in the CDR-L1 as a signature of kappa chain class members, and the ^{29}GY[KN] motif on 14 amino acids CDR-L1 as a signature of lambda class members. We did not observe conserved residues in the CDR-L3 targeting RBD, and we also note that the CDR-L1 motifs defined here can be encoded by multiple light chain V-genes (**Fig. 12C**, **Appendix H Table S4**).

Structural comparison of eight natively-paired IGHV3-53/3-66 class members showed that variable domains bound RBD with the same orientation, reflecting conserved heavy chain recognition and defining conserved light chain residues (**Fig. 12D**). Structural alignment revealed a conserved hydrogen bond network responsible for RBD recognition by CDR-H1. The backbone carbonyl of Gly26$_{HC}$ interacts with the amide group of Asn487; the backbone CO of Ser31$_{HC}$ contacts the hydroxyl of Tyr473; the side chain amide of Asn32$_{HC}$ contacts the carbonyl of Ala475; and the hydroxyl of Tyr33$_{HC}$ acts as hydrogen bond donor to Leu455 backbone CO within CDR-H1. For the CDR-H2, the hydroxyl group of Ser53$_{HC}$ targets both the backbone CO of R457 and the hydroxyl of Tyr421, the latter being involved in a hydrogen bond with the NH of Gly54$_{HC}$ as well; and the hydroxyl of Ser56$_{HC}$ interacts with the carboxyl of Asp420. Overall,

CDR-H2 interactions are less conserved among IGHV3-53/3-66 members compared to CDR-H1: the hydrogen bond between the $Tyr52_{HC}$ hydroxyl and the Lys417 amine is observed only in B38 and CV30, while the hydroxyl of Thr415 and the NH of Gly416 are targeted by the hydroxyl of $Tyr/Phe58_{HC}$ only in B38, CV30 and 910-30. Structural comparison of light chain residues shows a strongly conserved tyrosine ($Tyr32_{LC}$) in CDR-L1 at the heavy:light interface, which provides a stabilizing hydrophobic environment to the aromatic ring of Tyr505, together with $Val28/29_{LC}$ (in CV30, CC12.3, COVA2-04), $Ile29_{LC}$ (in 910-30, B38, CB6, CC12.1) or $Tyr30_{LC}$ (in C105). For CC12.1, B38 and 910-30, $Ser30_{LC}$ interacts with the side chains of Gln498 and Asn501.

We used published antibody repertoire data and these new defined sequence signatures to estimate the prevalence of antibody class precursors in healthy human immune repertoires (*133-135*). Antibody lineages with anti-SARS-CoV-2 IGHV3-53/3-66 signatures were identified in approximately 1 in 44,000 reported human antibody sequences (**Fig. 12E**) (*135*), which was high compared to previously studied anti-HIV-1 VRC01-class antibody precursors that occur in approximately 1 in 1-4 million human antibodies (*136*). This frequency estimate reflects the prevalence of precursor B cells with the *potential* to develop into protective mAbs of this antibody class, however, we emphasize that not all 1:44,000 potential precursor B cells are anticipated to develop into protective mAbs. The comparatively high frequency of IGHV3-53/3-66 anti-RBD precursors in human B cells supports the recovery of antibodies from this class in multiple convalescent COVID-19 patients.

RBD up/down conformation influences S protein recognition for the IGHV3-53/3-66 antibody class

The RBD 'up' position is required for ACE2 engagement, as well as IGHV3-53/3-66 class antibody binding (*95, 103, 117*). Cryo-EM analysis at endosomal pH has revealed a pH-mediated conformational switch that rotates RBD domains down at pH 5.5-4.5 (*118*), and the D614G mutation also alters RBD 'up' vs. 'down' dynamics. To understand how pH-based alteration of 'up' vs. 'down' RBD prevalence influence IGHV3-53/3-66 class recognition, we investigated antibody binding at three pH values related to known RBD 'up' versus 'down' states (**Fig. 13A**) (*95, 98, 99, 117, 118, 122, 137*). dhACE2 competition ELISA at pH 5.5 and 4.5 showed that IGHV3-53/3-66 class members compete in a concentration-dependent

manner with dimeric human ACE2 for binding to SARS-CoV-2 S2P spike, and to D614G S2P spike (**Fig. 13B**, **Appendix C Fig. 16**). Single-cycle surface plasmon resonance showed that the extremely potent mAb 1-20 recognized S protein and RBD with no loss in affinity at endosomal pH, whereas less potent antibodies 910-30 and B38 showed reduced affinity in the endosomal pH range (**Fig. 13C**, **Appendix D Fig. 17**). We compared authentic virus neutralization IC_{50} potencies (from **Fig. 11D**) to the ratio of mAb-Spike affinity (**Appendix D Fig. 17**) divided by reported dhACE2-Spike affinity (*137*), which suggested that potent mAb neutralization was correlated with mAb affinity across all tested pH values (**Fig. 13D**). Finally, a qualitative Octet pH series analysis using D614 S2P spike showed that as pH reduces (and RBDs preferentially rotate down), the potent neutralizer mAb 1-20 exhibited strong recognition of D614 S2P spike for pH 6.0, whereas 910-30 showed reduced binding below pH=6.5, and the least potent B38 binding showed reduced binding below pH=7.0 (**Fig. 11E***,* left panel). In contrast, all class members maintained strong binding to mutant D614G S2P spike in the endosomal pH range (where one RBD likely remains up), and the potent mAb 1-20 recognized D614G spike down to pH 4.0 (**Fig. 11E**, right panel). Thus, the most potent antibodies maintained a bound state at endosomal pH conditions for D614 S2P spike. We also observed that antibody class members effectively recognized the native RBD-up conformation for D614G across a broad pH range. Our data support ACE2 competition as a functional signature of IGHV3-53/3-66 public antibody class neutralization, and we show that the RBD-up vs. RBD-down conformation substantially influenced the ability of IGHV3-53/3-66 class antibodies to recognize spike trimer.

2.3 Discussion

Importantly, these findings were made during the midst of the COVID-19 pandemic, underscoring the utility of these methods to rapidly generate relevant scientific insights related to major public health concerns. Enhanced understanding of IGHV3-53/3-66 class-based spike recognition can provide insight into immune monitoring, antibody discovery, and vaccine design against SARS-CoV-2. Structural analysis of a novel class member mAb 910-30 revealed previously undescribed spike disassembly at high occupancy, and our antibody class comparative studies showed that native heavy:light pairing was essential for potent neutralization, despite high similarities in heavy chain sequences. We note that non-native heavy:light swaps still encoded the general structural signatures features required for effective SARS-CoV-2 neutralization by antibodies of this class, and we anticipate that directed evolution could restore effective neutralization of non-native heavy-light antibody pairings.

Comparative sequence-structure analyses enabled identification of conserved light chain class signatures, defined as $^{27a/28}$[GDS]xSx(1,2)[FY] (kappa) and ^{29}GY[KN] (lambda) residues in CDR-L1 that made important contributions to RBD recognition. We also note that class member light chains use common aromatic/hydrophobic residues ^{28}Val, ^{29}Ile/Val, or $^{30/32}$Tyr30/32 to achieve similar interactions with ^{505}Tyr in the RBD, which is part of the shared ACE2 and IGHV3-53/3-66 class binding epitope. These shared light chain features illuminate the structural rationale for broader light chain diversity among IGHV3-53/3-66 class members. Interestingly, conserved residues in both the heavy and light chain of IGHV3-53/3-66 antibodies targeted RBD residues that have been mutated in emerging SARS-CoV-2 variants. The effect of single mutations observed in B.1.1.7 and B.1.351 on 910-30 neutralization potency has been recently reported (*138*). A decrease in 910-30 efficacy is observed for the K417N and the N501Y mutants, while only a minor change is observed for the E484K mutant. These results are consistent with the structural data, as Lys417 and Asn501 are part of the epitope of 910-30, while Glu484 is not.

The frequency of anti-SARS-CoV-2 IGHV3-53/3-66 precursor antibodies in healthy donors (around 1 in 44,000) was more common than the previously studied anti-HIV-1 VRC01-class antibody precursors observed in 1 per 1-4 million antibodies (*136*). In addition, it has been shown that anti-HIV-1 VRC01-class

antibodies require much higher levels of somatic hypermutation (SHM) to achieve potent neutralization (*136*). The comparably limited SHM required for anti-SARS-CoV-2 IGHV3-53/3-66 class antibodies appears to be a feature of IGHV germline gene neutralizing interactions and the need to recognize more similar viral variants, as compared to HIV-1 broadly neutralizing antibodies that must recognize broadly diverse viral variants and show limited germline gene neutralization. These findings help explain the observed reproducibility of public IGHV3-53/3-66 anti-RBD antibodies in convalescent COVID-19 patients.

D614G S2P spike shows a greater prevalence of RBD-up than D614G, which may enhance spike and the ACE2 host receptor recognition to confer higher D614G viral infectivity (*122, 123, 125, 126*). Conversely, a sustained RBD 'up' also could make the virus more sensitive to neutralization, as the exposed 'up' RBD enhances exposure of vulnerable epitopes (*118, 126*). We outlined differences in RBD display caused by the D614G mutation that enhance antibody class recognition of spike across a broad pH range, and we show that D614G had no detrimental impact on IGHV3-53/3-66 antibody class neutralization, which agrees with prior reports (*108, 139, 140*). Interestingly, only the most potent antibodies could bind to the D614 variant at endosomal pH, demonstrating that high-affinity antibody recognition can prevent D614 RBD from rotating down at pH 5.5-4.5. These data imply that screening for antibody recognition of D614 S2P at endosomal pH could be an effective method to identify potent anti-SARS-CoV-2 antibodies, and other studies have reported potent antibodies recognizing S trimer even in the context of RBD down conformations (*141*). We also found that ACE2 competition at pH 7.4 was correlated with potent antibody protection, consistent with cell surface attachment via ACE2 at serological pH.

In summary, here we report the discovery of a new public IGHV3-53/3-66 antibody class member and outline the unique heavy and light chain interactions that lead to potent immune recognition of both D614G and D614G spike variants. These data enhance our understanding of the public IGHV3-53/3-66 antibody class and highlight its convergent neutralization features to accelerate anti-SARS-CoV-2 antibody mapping and inform future efforts to identify and elicit neutralizing antibody responses against COVID-19.

2.4 Methods

Human Sample Collection

Informed consent was obtained for all study participants under IRB-AAAS9010 (Hong Kong University). Donor 910 (*127*) was a 73 year-old male. Patient serum was collected for ELISA and neutralization assays, and PBMCs were cryopreserved for subsequent B cell receptor gene capture and antibody screening.

Expression and Purification of SARS-CoV-2 Antigens

The antigen probes used for sorting yeast surface displayed libraries were prepared as previously described (*84*). Briefly, expression vectors encoding the ectodomain of the SARS-CoV-2 S protein was transiently transfected into Expi293 cells and then purified five days post transfection using on-column purification methods.

Production of SARS-CoV-2 Pseudovirus

SARS-CoV-2 pseudovirus was generated using recombinant Indiana vesicular stomatitis virus (rVSV) as previously described (*84, 142-144*). HEK293T cells were grown to 80% confluency then used for transfection of pCMV3-SARS-CoV-2-spike (kindly provided by Dr. Peihui Wang, Shandong University, China) using FuGENE 6 (Promega). Cells were cultured to grow overnight at 37 °C with 5% CO2. Then medium was removed and VSV-G pseudotyped ΔG-luciferase (G*ΔG-luciferase, Kerafast) was harvested to infect the cells in DMEM at a MOI of 3 for 1 h. Then cells were washed three times with 1× DPBS. DMEM supplemented with anti-VSV-G antibody (I1, mouse hybridoma supernatant from CRL-2700; ATCC) and was added to the inoculated cells. The cells were then cultured overnight. The supernatant was removed the following day and clarified by centrifugation at 300 g for 10 mins before storing at −80 °C.

Emulsion Overlap Extension RT-PCR and Yeast Display Library Generation

B cells were isolated from Donor 910 cryopreserved PBMCs. Non-B cells were depleted by magnetic bead separation, and $CD27^+$ antigen-experienced B cells were isolated by positive magnetic bead separation (EasySep Human B cell enrichment kit w/o CD43 depletion, STEMCELL Technologies, Vancouver,

Canada, and CD27 Human Microbeads, Miltenyi Biotec, Auburn, CA, USA). Antigen-experienced B cells (memory B cells) were stimulated *in vitro* for 5 days to enhance antibody gene transcription. For stimulation, cells were incubated 5 days in the presence of Iscove's Modified Dulbecco's Medium (IMDM) (Thermo Fisher Scientific) supplemented with 10% FBS, 1x GlutaMAX, 1x non-essential amino acids, 1x sodium pyruvate and 1x penicillin/streptomycin (Life Technologies) along with 100 units/mL IL-2 and 50 ng/mL IL-21 (PeproTech, Rocky Hill, NJ, USA). B cells were co-cultured with irradiated 3T3-CD40L fibroblast cells that secrete CD40L (kind gift of John Mascola, Vaccine Research Center, NIAID) to aid B cell expansion. Single B cells were captured in emulsion droplets via a flow focusing device with concentric nozzles flowing suspended cells, lysis buffer with mRNA capture magnetic oligo (dT)-coated magnetic beads, and a viscous oil solution to form stable droplets compartmentalizing single B cells with lysis buffer and the mRNA capture beads (*19*). Captured beads loaded with single-cell mRNA were re-emulsified and the captured RNA product was reverse transcribed using a SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase (Thermo Fisher Scientific). The specific immunoglobulin VH and VL genes were then processed with an overlap-extension RT-PCR to link native heavy and light chains into a single amplicon, introducing two restriction sites: NheI and NcoI between the VH and VL genes for downstream yeast library generation (*20*). Natively paired antibody heavy and light chain sequencing and yeast surface display library generation were performed as described previously (*18-20, 145, 146*).

For yeast library generation, cDNA libraries were amplified with primers containing the yeast display vector restriction sites: AscI and NotI, used for subcloning into the yeast display vector. PCR amplified products were purified by agarose gel extraction and digested with AscI and NotI restriction enzymes followed by subsequent ligation into the yeast display vector backbone. This step was performed in duplicate for each library with separate Kappa- or Lambda- gene-specific primers and a corresponding Kappa or Lambda display vector to generate Kappa and Lambda libraries. Ligated plasmid libraries were transformed into high-efficiency electrocompetent E. coli, expanded overnight, and maxiprepped to isolate the plasmid library DNA product. Maxiprepped plasmid libraries were digested with NheI and NcoI restriction enzymes to remove the native linker from VH:VL pairing. Digested product was purified by

agarose gel extraction, and then ligated with a pre-digested DNA gene encoding a bidirectional Gal1/Gal10 promoter inserted between the VH and VL sequences. The resulting ligated product was again transformed into high-efficiency electrocompetent E. coli, expanded overnight, and maxiprepped to isolate the plasmid library DNA product now containing the bidirectional promoter. A final PCR amplification was performed to amplify the VH:bidirectional promoter:VL amplicon with overhanging homologous ends to the pCT backbone for high-efficiency yeast transformation into AWY101 using an homologous recombination method previously described (*76*). Transformed libraries were passaged twice in SD-CAA to ensure a 1:1 ratio of plasmid DNA to yeast colony (*76*).

FACS Screening of Yeast Libraries

To induce Fab surface expression yeast libraries were incubated in SGD-CAA media at 20 °C, 225 rpm for 36 hrs. For the first round of sorting, $3x10^7$ presorted cells were washed twice with staining buffer (1x PBS with 0.5% BSA and 2 mM EDTA). Washed yeast display libraries were stained with 20 nM of trimer antigen and a monoclonal anti-FLAG-FITC marker to measure Fab expression (Monoclonal ANTI-FLAG M2-FITC antibody, Sigma-Aldrich, St. Louis, MO, USA). For staining with the NHS-Biotin S-Trimer Protein probe, cells were mixed with 20 nM un-labeled antigen and a monoclonal anti-FLAG-FITC marker (Monoclonal ANTI-FLAG M2-FITC antibody, Sigma-Aldrich, St. Louis, MO, USA) used to measure VL surface expression. This mix was incubated for 15 minutes at 4 °C with gentle agitation on a platform shaker. Following incubation, a Streptavidin PE conjugate (Streptavidin, R-Phycoerythrin Conjugate Premium Grade, Thermo Fisher Scientific, Waltham, MA, USA) was added to the re-suspended mix to fluorescently label the biotinylated antigen protein and the sample was again incubated for 15 minutes at 4 °C with gentle agitation on a platform shaker. These NHS-Biotin S-Trimer Protein samples were then washed 3x and re-suspended in a final volume of 1 mL in staining buffer before being filtered through a 35 micron-filter cap FACS tube. For staining the with His-Strep-II S-Trimer Protein probe, cells were incubated with un-labeled antigen for 15 minutes at 4 °C with gentle agitation on a platform shaker. Samples were then washed 3x with staining buffer, and resuspended in a common mix containing the monoclonal anti-FLAG-FITC marker and a monoclonal anti-His-PE antibody (PE anti-His Tag Antibody,

BioLegend, San Diego, CA, USA) to label surface expressed, antigen bound Fab. These cells were again incubated for 15 minutes in the fluorophore mix at 4 °C with gentle agitation on a platform shaker. The fluorescently labeled samples were then washed 3x and resuspended in a final volume of 1 mL staining buffer before being filtered through a 35 micron-filter cap FACS tube. Samples were kept in the dark on ice until sorting. Subsequent rounds of enrichment sorting were performed using the same staining procedure, but for only $5x10^6$ input cells and 250 μL final resuspension volume.

A SONY Multi-Application 900 cell sorter running SONY LE-MA900FP Cell Sorter Software was used to detect all FITC+/PE+ cells from each sample and sort them into low pH SD-CAA media. The gating strategy used was previously described (*20*). Sorted yeast were expanded for 24-48 hrs at 30 °C, 225 rpm and then passaged into SGD-CAA media to induce Fab expression for the next round of sorting. This process was repeated for 3-4 rounds of sorting to enrich for Fab-expressing antigen-binding library populations. In addition to the antigen-positive sorts, an aliquot of each yeast library was washed and stained with only the anti-FLAG-FITC marker, and all FITC+ (i.e., VL+) cells were sorted and sequenced for use as a reference database for NGS enrichment ratio calculations. Analysis of flow cytometry data was conducted using Flowjo10.4 (Flowjo, LLC, Oregon, USA).

NGS Analysis of Sorted Yeast Libraries

After each round of FACS enrichment, yeast libraries were expanded via incubation at 30 °C for 24-48 hrs. An aliquot of this culture was used for high-efficiency yeast plasmid DNA extraction (*147*). A high-fidelity polymerase (Kapa Hifi HotStart Mastermix, Kapa Biosystems, Massachusetts, USA) and primers targeting the yeast display vector backbone were used to amplify HC and LC genes from each library (*20*). A second round of primer-extension PCR with barcoded primers added a unique identifier to all HC and LC from a particular library (*19*). Sorted libraries were sequenced on the Illumina 2x300 MiSeq platform and sequencing was performed for each library after each round of FACS enrichment. Data processing of Illumina Raw FASTQ data was performed as reported previously (*19, 20*). Briefly, Illumina sequences were quality-filtered to improve read quality, followed by V(D)J gene identification and annotation of CDR3 regions using IgBLAST (*148*). Antibody clonal lineages were tracked across yeast sort rounds by their

CDR-H3 amino acid sequence and enrichment ratio. Enrichment ratios were calculated by comparing sequence prevalence in each sorted libraries to the unsorted, Fab-expressing (VL+) antibody library.

Antibody Production and Purification

The 910-30 antibody was codon-optimized, cloned into mammalian expression plasmid, and expressed as full human antibody IgG1s by co-transfection into Expi293 cells. Heavy and light chain plasmids were co-transfected into Expi293F (ThermoFisher) mammalian cells using the ExpiFectamineTM 293 Transfection Kit (Thermo Fisher Scientific, Massachusetts, USA) and culture in 37 °C shaker at 125 rpm and 8% CO_2. On day 6 post transfection, the supernatant from transient transfection culture were purified with Protein G or A resin (GenScript, New Jersey, USA) and concentrated using an Amicon Ultra-4 Centrifugal 30K Filter Unit (MilliporeSigma, Maryland, USA), then stored at 4 °C.

ELISA Binding Assays to S trimer and RBD

S trimer and RBD enzyme-linked immunosorbent assays (ELISAs) (Fig. 2C) were performed in triplicate. 175 ng of antigen per well was coated onto 96-well ELISA plates at 4 °C overnight. Plates were washed and then blocked with 100 µL of blocking buffer at 37 °C for 2 hrs. Purified antibodies were serial diluted using dilution buffer, added to the antigen-coated blocked plates, and then incubated at 4 °C for 1 hr. Plates were washed and 50uL of a secondary anti-human kappa light chain detection antibody (A18853, Invitrogen, Carlsbad, CA) was added to each well and incubated at room temperature for 1 hr. After the final wash, 50 µL TMB substrate (00-4203-56, ThermoFisher Scientific, Waltham, MA) was used to detect antibody binding to antigen measuring absorbance at 405 nm.

Pseudovirus SARS-CoV-2 Viral Neutralization Assay

SARS-CoV-2 pseudovirus neutralization assays were performed as previously described (*84*). Briefly, pseudovirus particles were generated from recombinant Indiana VSV (rVSV) expressing SARS-CoV-2 S protein. Neutralization was assessed by incubating pseudoviruses with serial dilutions of purified antibody, and scored by the reduction in luciferase gene expression.

Authentic SARS-CoV-2 Viral Neutralization Assay

Authentic virus neutralization assays were performed as previously described (*84*). Briefly, to measure the neutralizing activity of purified mAbs an end-point dilution assay in a 96-well plate format was performed. Each mAb was 5-fold serially diluted starting at 20 μg/mL in triplicate. Dilutions were incubated with live SARS-CoV-2 for 1 hr at 37 °C, and post-incubation the virus-antibody mixture was transferred onto a monolayer of Vero-E6 cells and incubated for 70 hrs. CPE from the resulting cell incubations were visually scored for each well in a blinded fashion by two independent observers.

dhACE2 Competition ELISA

Antibodies were assayed for dhACE2 competition by enzyme-linked immunosorbent assays (ELISAs) (Fig. 2E, 3C, S3I) in triplicate. ELISA experiments were performed in parallel at three pH values 7.4, 5.5, and 4.5. ELISA 96-well plates were coated with 175 ng per well of antigen in pH-adjusted PBS and incubated at room temperature for 1 hr. Ag-coated ELISA plates were washed and blocked with 100 μL of blocking buffer and incubated at room temperature for 1 hr. Purified antibodies were serial diluted and pre-mixed with dhACE2 using pH-adjusted dilution buffer. Ab:dhACE2 premixes were added to the pre-blocked, antigen-coated plates and incubated at room temperature for 2 hrs. Plates were washed and 50 μL of 1:2000 diluted, pH-adjusted secondary anti-human kappa light chain detection antibody (A18853, Invitrogen, Carlsbad, CA) solution was added to each well and incubated at room temperature for 1 hr. After the final wash, 50 μL Super AquaBlue substrate was used to detect antibody binding to antigen measuring absorbance at 405 nm.

RBD Glycan Recognition Analysis via Yeast Display

For plasmid construction, pJS699 (S-RBD (333-537)-N343Q for fusion to the C-terminus of AGA2) was synthesized by PCR amplifying pUC19-S-ecto with primers PJS-P2196/PJS-P2197 (2.9kb) and PJS-P2198/PJS-P2199 (0.65kb). The resulting products were fractionated by agarose gel electrophoresis and the bands corresponding to the desired products were excised from the gel and purified using a Monarch DNA Gel Extraction Kit (NEB). The fragments were assembled using NEBuilder HiFi DNA assembly master mix (NEB) according to the manufacturer's instructions and 5 μL of the reaction was transformed

into chemically competent *E. coli* Mach1 (Invitrogen) and selected on LB agar supplemented with 50 μg/mL kanamycin.

To create the display construct of S-RBD (333-537)-N343Q fused to the C-terminus of Aga2p, pJS697 was digested with BsaI-HFv2 (NEB) and purified using a Monarch PCR & DNA Cleanup Kit (NEB). pJS699 was digested with NotI-HF (NEB), the reaction fractionated by agarose gel electrophoresis, and the band corresponding to S-RBD (0.83kb) excised and purified using a Monarch DNA Gel Extraction Kit (NEB). The two fragments were co-transformed (in a 2.4:1 molar ratio of S-RBD to backbone) into chemically competent *S. cerevisiae* EBY100 (*149*) and selected on M19D agar. M19D contained 5 g/L casamino acids, 40 g/L dextrose, 80 mM 2-(*N*-morpholino) ethanesulfonic acid (MES free acid), 50 mM citric acid, 50 mM phosphoric acid, 6.7 g/L Yeast Nitrogen Base (Sigma), and was adjusted to pH 7 with 9M NaOH, 1M KOH.

Recombinant human ACE2-Fc and CR3022 were received as a gift from Neil King and David Veesler at the University of Washington. Human ACE2-Fc was produced and purified as described (Walls et al., 2020). CR3022 (*130*) was expressed by transient transfection in Expi293F cells and purified by protein A affinity chromatography and SEC using a Superdex 200 10/300 GL. Specificity was verified by measuring binding to SARS-CoV-2 RBD and irrelevant antigen.

For yeast display screening, EBY100 harboring the RBD display plasmid was grown in 1 mL M19D overnight at 30 ^{o}C. Expression was induced by resuspending the M19D culture to OD_{600}=1 in M19G (5 g/L casamino acids, 40 g/L galactose, 80 mM MES free acid, 50 mM citric acid, 50 mM phosphoric acid, 6.7 g/L yeast nitrogen base, adjusted to pH 7 with 9 M NaOH, 1 M KOH) and growing 22 hrs at 22 °C with shaking at 300 rpm. Yeast surface display titrations were performed as described (*150*) with an incubation time for 910-30 of 4 hrs and using secondary labels anti-c-myc-FITC (Miltenyi Biotec) and Goat anti-Human IgG Fc PE conjugate (Invitrogen Cat. No. 12-4998-82). Titrations were performed in biological replicate.

Glycosylation-Independent Binding for Antibody 910-30

EBY100 harboring the RBD display plasmid was grown in 1 mL M19D overnight at 30 ºC. Expression was induced by resuspending the M19D culture to OD_{600}=1 in M19G (5 g/L casamino acids, 40 g/L galactose, 80 mM MES free acid, 50 mM citric acid, 50 mM phosphoric acid, 6.7 g/L yeast nitrogen base, adjusted to pH 7 with 9 M NaOH, 1 M KOH) and growing 22 h at 22 °C with shaking at 300 rpm. Yeast surface display titrations were performed as described (*150*) with an incubation time for 910-30 of 4 hrs at room temperature and the secondary labels anti-c-myc-FITC (Miltenyi Biotec) and Goat anti-Human IgG Fc PE conjugate (Invitrogen Catalog # 12-4998-82). Titrations were performed in biological replicate (n = 2) with three technical replicates.

910-30 IgG was chemically biotinylated using NHS-Ester biotin (ThermoFisher EZ-Link Biotin Cat. No. 20217) at a 20:1 molar ratio of biotin:IgG according to manufacturer's instructions. $1x10^5$ yeast cells were labelled with no protein or 100 nM non-biotinylated CR3022, hACE2 or 910-30 for 30 min at room temperature in PBSF (PBS containing 1 g/L BSA). The same cells were then labelled with 1 nM chemically biotinylated 910-30, in the same tube without washing, for 30 min at room temperature in PBSF. The cells were centrifuged and washed with 200 µL PBSF. They were labeled with 0.6 µL FITC, 0.25 µL SAPE and 49.15 µL PBSF for 10 min at 4 °C. Cells were then centrifuged, washed with PBSF, and analyzed on a flow cytometer. Experiments were performed with three technical replicates and two biological replicates.

Delineation of IGHV3-53/3-66 Sequence Signatures

A structure-based method was applied to define sequence signatures for the HV3-53/3-66 class COVID neutralizing antibody (*24*). Briefly, protein structures of IGHV3-53/3-66 antibodies complexed with RBD or spike were selected for analysis, and the buried surface area (BSA) between antibody and RBD was calculated by the PDBePISA server (https://www.ebi.ac.uk/pdbe/pisa/). We examined the BSA larger than 20 $Å^2$, and residues making contacting with the RBD projected surface that were encoded by the conserved germline sequence were selected as initial class sequence signatures, and amino acids from somatic hypermutations were used to refine the signature of the class antibody. For germline sequence alignments, heavy and light chain germline sequences were downloaded from IMGT (*132*) and the sequences of CDR1 and CDR2 were extracted and aligned based on Kabat numbering. ANARCI server

was used to number amino acid sequences of antibody (http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/anarci).

Antibody Class Frequency Estimation

The frequency of antibody class was estimated using OLGA software based on defined motif (*134*). NGS samples of three healthy donors (NCBI Short Read Archive accession code: PRJNA511481) were used to analyze heavy and light chain lineage precursor frequencies (*135*). The ratio of human kappa and light chain (60:40) was obtained from Bräuninger et. al., 2001 (*133*). Antibody class precursor frequency was calculated as:

(Frequency of heavy chain × Frequency of kappa chain × kappa chain ratio) + (Frequency of heavy chain × Frequency of lambda chain × lambda chain ratio)

Negative Stain Cryo-EM

Samples were diluted to a spike concentration of about 20 µg/ml. A 4.7-µl drop of the diluted sample was applied to a glow-discharged carbon-coated copper grid. The grid was washed with a buffer with the same pH as the sample buffer (10 mM HEPES with 150 mM NaCl for pH 7.4; 10 mM acetate with 150 mM NaCl for the lower pH values). Protein molecules adsorbed to the carbon were negatively stained with 0.75% uranyl formate. Datasets were collected using a ThermoFisher Talos F200C electron microscope equipped with a Ceta CCD camera. The microscope was operated at 200 kV, the pixel size was 2.53 Å (nominal magnification: 57,000), and the defocus was set at -1.2 µm. Particles were picked and extracted automatically using in-house written software (YT, unpublished). 2D classification was performed using Relion 1.4 (*151*).

Cryo-EM Sample Preparation

SARS-CoV-2 S2P spike was expressed and purified as described in Wrapp et al., 2020 (*95*). 910-30 Fab was prepared by incubating the full 910-30 IgG with immobilized papain for 3 hrs at 37 °C in 50 mM phosphate buffer, 120 mM NaCl, 30 mM cysteine, 1 mM EDTA, pH 7. Purified SARS-CoV-2 spike was diluted to a final trimer concentration of 0.33 mg/mL and mixed with 910-30 Fab in a 1:1 molar ratio (sample 1) or 1:9 molar ratio (sample 2). The final buffer for both samples was 10 mM sodium acetate, 150

mM NaCl, pH 5.5; 0.005% w/v n-Dodecyl β-D-maltoside (DDM) was added to the mixture to prevent aggregation during vitrification. After incubation for 1 hour on ice, a volume of 2 μL was applied to a glow-discharged carbon-coated copper grid (CF 1.2/1.3 300 mesh) and vitrified using a Vitrobot Mark IV with a wait time of 30 s and a blot time of 3 s.

Cryo-EM Data Collection, Processing and Model Fitting

Cryo-EM data were collected using the Leginon software (*152*) installed on a Titan Krios electron microscope operating at 300 kV, equipped with a Gatan K3-BioQuantum direct detection device. The total dose was fractionated for 3 s over 60 raw frames. Data processing including motion correction, CTF estimation, particle picking and extraction, 2D classification, ab initio model generation, 3D refinements and local resolution estimation for both sample 1 and sample 2 datasets were carried out in cryoSPARC 2.15 (*153*). The coordinates of SARS CoV-2 spike with 1 RBD up, PDB entry 6VYB (*117*), were employed as initial template to model the cryo-EM map of 910-30 Fab in complex with SARS-CoV-2 spike (sample 1). The RBDs were modeled using the crystallographic structure of RBD in complex with B38 Fab (PDB entry 7BZ5) (*109*) as a template. The variable region of 910-30 Fab was initially modeled using PDB models 7BZ5 and 5SX4 (*154*) for the heavy and light chain respectively. The residues at the Fab:RBD interface were modeled by structural comparison of 910-30 Fab with 7 different antibodies belonging to the IGHV3-53/3-66 class. Automated and manual model building were iteratively performed using real space refinement in Phenix (*155*) and Coot (*156*) respectively. EMRinger (*157*) and Molprobity (*158*) were used to validate geometry and check structure quality at each iteration step. UCSF Chimera (*159*) and Chimera X (*160*) were used to calculate map-fitting cross correlation (Fit-in-Map tool) and to prepare figures.

Octet Binding Experiments

Binding of mAbs 4-3, B38, 910-30, and 1-20 to SAR-CoV-2 S2P D614 and D614G variants was assessed an a FortéBio Octet HTX instrument (FortéBio). Experiments were run in tilted black 384-well plates (Geiger Bio-One) at 30 °C and 1,000 rpm agitation. Running buffer was comprised of 10mM of the corresponding pH buffer plus 150mM NaCl, 0.02% Tween20, 0.1% BSA and 0.05% sodium azide. The

following buffers were used to achieve the range of pH: pH 9 (borate), pH 8.5 (Tris), pH 8 (Tris), pH 7.4 (PBS), pH 7 (HEPES), pH 6.5 (MES), pH 6 (MES), pH 5.5 (NaAc), pH 5 (NaAc), pH 4.5 (NaAc), pH 4.2 (NaAc), pH 4.0 (NaAc). 300 nM IgG solution was used for immobilization at pH 7.4 on anti-human IgG Fc capture biosensors (FortéBio) that were pre-hydrated for 30 minutes. Sensors were then equilibrated in pH 7.4 buffer for 30 s followed by 180 s in the altered pH buffer. Binding was assessed at 200nM S2P D614 or D614G and response recorded for 180 s. Dissociation in the respective buffer was measured for 300 s. The Data Analysis Software HT v12.0 (Fortebio) was used to subtract reference well signal from loaded sensor dipped into buffer without spike protein. The maximum association response (nm) is reported at each pH.

SPR Binding Experiments

SPR binding experiments were performed using a Biacore T200 biosensor, equipped with a Series S SA chip. The running buffer varied depending on the pH of the binding reaction; experiments at pH 7.4 were performed in a running buffer of 10 mM HEPES pH 7.4, 150 mM NaCl, 0.1% (v/v) Tween-20; at pH 5.5 experiments were performed in 10mM sodium acetate pH 5.5, 150 mM NaCl, 0.1% (v/v) Tween-20; and at pH 4.5 in 10 mM sodium acetate pH 4.5, 150 mM NaCl, 0.1% (v/v) Tween-20. All measurements were performed at 25 °C.

Biotinylated S2P was captured over independent flow cells at 750-900 RU. 910-30 and 1-20 IgGs were tested over the biotinylated S2P surfaces at four concentrations ranging from 1-27nM, while B38 and 4-3 were tested at four concentrations ranging from 3-81 nM, to account for higher binding KDs. Biotinylated RBD was captured over independent flow cells at 250-500 RU and B38 was tested at four concentrations ranging from 3-81 nM, 910-30 and 4-3 were tested at four concentrations of 1-27 nM and 1-20 at four concentrations ranging from 0.333-9 nM, to account for differences in their binding affinities. To avoid the need for surface regeneration that arises with the slowly dissociating interactions, we used single-cycle kinetics binding experiments. The four concentrations for each IgG were prepared in running buffers at each of pH, using a three-fold dilution series.

Binding of 910-30, 4-3 and B38 over the S2P or RBD surface as well as over a streptavidin reference surface was monitored for 120 s, followed by a dissociation phase of 120-1080 s depending on the interaction at 50 μL/min. For the interaction of 1-20 with the RBD, which showed an unusually slow dissociation rate, an extended dissociation phase of 4500 s was necessary to extrapolate accurate apparent dissociation constants. Four blank buffer single cycles were performed by injecting running buffer instead of Fab to remove systematic noise from the binding signal. The data was processed and fit to 1:1 single cycle model using the Scrubber 2.0 (BioLogic Software). The results from these assays, are reported in terms of apparent kinetic parameters and K_Ds to account for potential avidity effects arising from the binding of bivalent IgGs to trivalent S2P.

QUANTIFICATION AND STATISTICAL ANALYSIS

IC50 calculations were reported using GraphPad Prism software (version 8.4.3). Briefly, experimental data was imported and modeled using a least squares regression method to fit the data to a variable slope (four parameter) inhibitor vs. response curve with bottom parameters constrained to zero. Flow cytometry analysis was carried out using FlowJo software (version 10.4). The Spearman rank order correlation was calculated using cor.test function in base R. Spearman ρ and the p-values for the test were used to determine the strength of the correlation between tested variables.

KEY RESOURCES TABLE

Unique link for KRT: https://star-methods.com/?rid=KRT5f99969f70be3

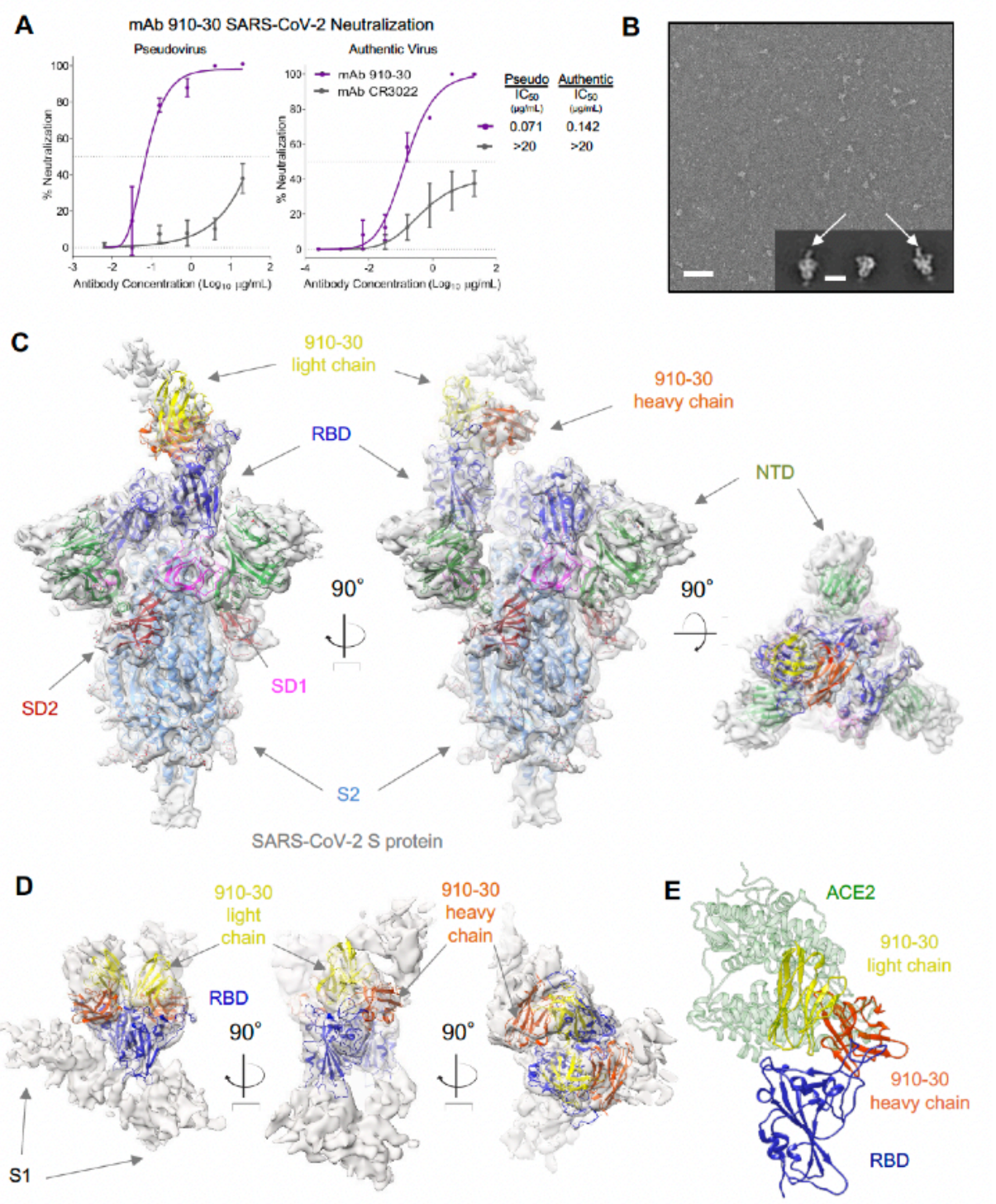


Figure 10. A novel SARS-CoV-2 neutralizer in the reproducible IGHV3-53/3-66 antibody class targets the ACE2 binding site of both ordered and disassembled spike.

Fig. 10. A novel SARS-CoV-2 neutralizer in the reproducible IGHV3-53/3-66 antibody class targets the ACE2 binding site of both ordered and disassembled spike. (A) The novel SARS-CoV-2 neutralizing antibody 910-30 shows moderately potent neutralization capacity compared to the control mAb CR3022 in both VSV-pseudo-type virus and authentic virus assays. Data are represented as mean ± SD. **(B)** Negative-staining electron microscopy at pH 5.5 revealed 910-30 Fab bound to SARS-CoV-2 S2P protein. A representative micrograph is shown. Inset shows representative 2D class averages; arrows point to bound Fab fragments. Scale bars: 50 nm (micrographs) 20 nm (2D class averages). **(C)** Cryo-EM map and molecular model of 910-30 Fab in complex with SARS-CoV-2 spike at 4.75 Å resolution. Only one conformation, with 1 Fab bound to 1 RBD up, is observed when mixing Fab and spike in a 1:1 molar ratio at pH 5.5. NTD is colored in green, RBD in blue, SD1 in magenta, SD2 in red, S2 in light blue, 910-30 antibody heavy chain in orange, 910-30 light chain in yellow. **(D)** Cryo-EM map obtained when 910-30 Fab and spike are mixed in a 9:1 molar ratio at pH 5.5. The only observed species is a mostly disordered spike in which RBD and 910-30 Fab still fit the map consistently with the properly folded spike:910-30 complex shown in (C). RBD is shown in blue, 910-30 heavy chain in orange, 910-30 light chain in yellow. **(E)** The structural superposition of ACE2 (PDB entry 6M0J) and 910-30 (PDB entry 7KS9) in complex with SARS-CoV-2 RBD shows a representative ACE2 competition mechanism defining IGHV3-53/3-66 class neutralization. ACE2 is colored in green, 910-30 heavy chain in orange, 910-30 light chain in yellow, RBD in blue. See also Figure S1, Figure S2, and Table S1.

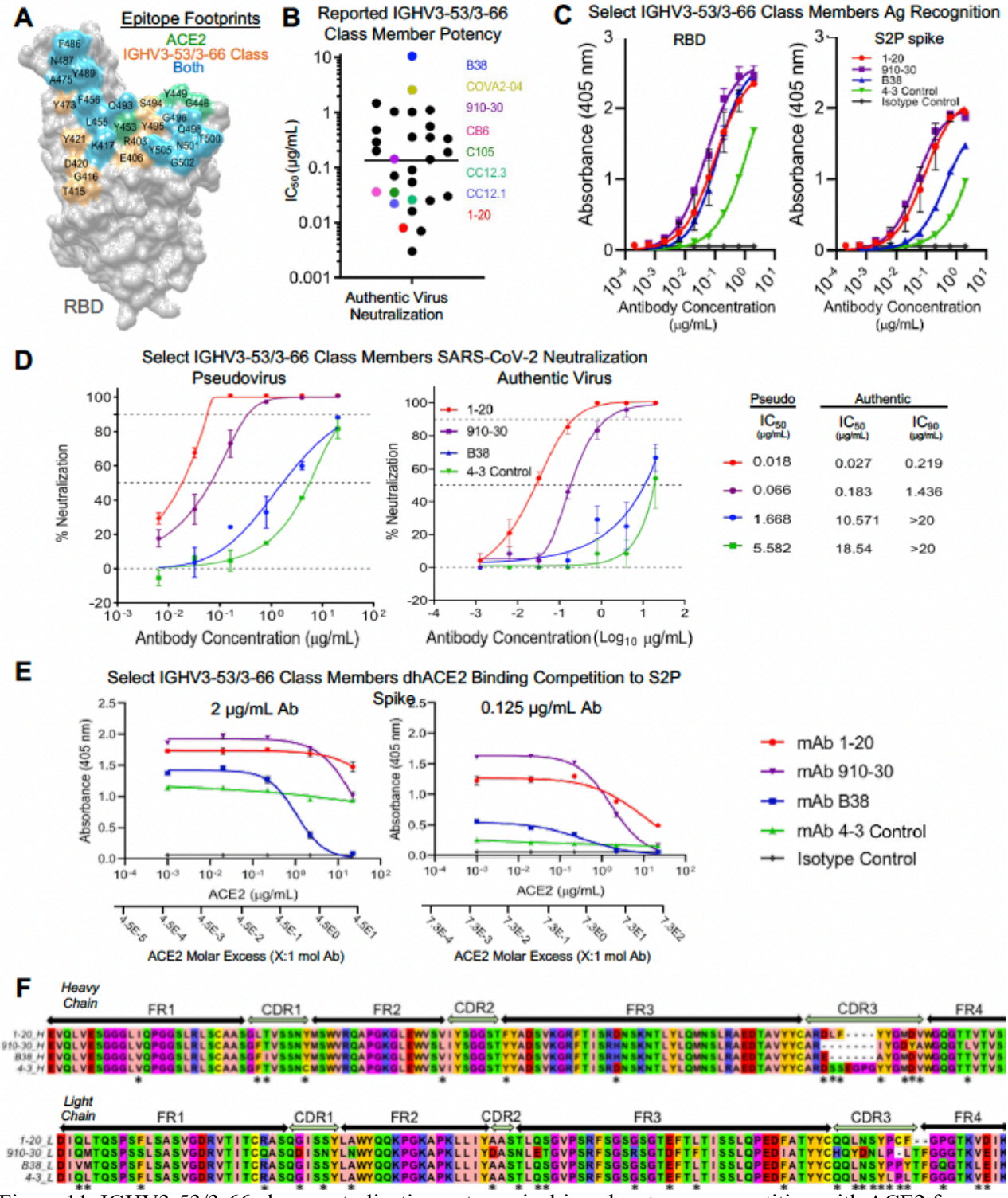
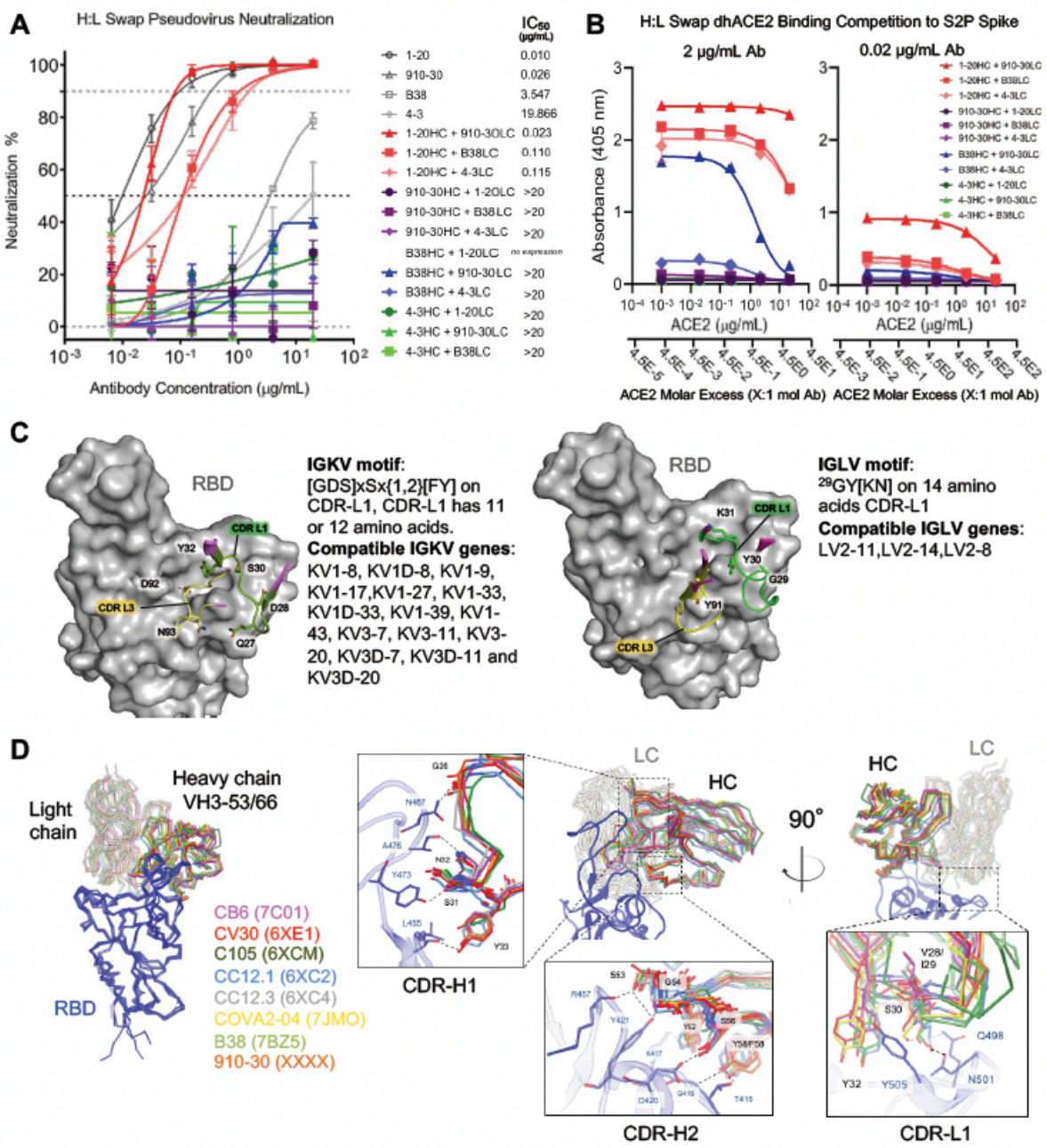


Figure 11. IGHV3-53/3-66 class neutralization potency is driven by strong competition with ACE2 for spike S2P recognition.

Fig. 11. IGHV3-53/3-66 class neutralization potency is driven by strong competition with ACE2 for spike S2P recognition. (A) Epitope footprint comparison between ACE2 binding site and the IGHV3-53/3-66 class epitope on RBD show substantial overlap. Residues interacting with ACE2 only are shown in green, residues targeted by the IGHV3-53/3-66 class only are shown in orange, residues that overlap are shown in light blue. Other sites on RBD are represented in gray. **(B)** Dot-chart of reported wild-type authentic virus neutralization IC_{50} titers for previously published IGHV3-53/3-66 anti-RBD antibodies. Data were plotted without correcting for any differences in neutralization assay protocols. Line indicates the mean of IC_{50} values. A list of antibodies, IC_{50} values, and citations are provided in Supplemental Table 2. **(C)** IgG ELISA binding titrations for select IGHV3-53/3-66 class members against S2P spike and RBD antigens, with an IGHV gene-matched control (mAb 4-3) and an isotype control. Data are represented as mean ± SD. **(D)** Pseudovirus & authentic virus neutralization show that IC_{50} neutralization potency ranges two orders of magnitude between the selected IGHV3-53/3-66 class members, along with an IGHV gene-matched control (mAb 4-3). Data are represented as mean ± SD. **(E)** dhACE2 competition ELISA against SARS-CoV-2 S2P spike showing constant IgG concentrations with increasing dhACE2 (ACE2) concentrations. dhACE2 concentration is provided as both μg/mL and as ACE2 molar excess units. Data are represented as mean ± SD. **(F)** Sequence alignment of heavy chain (upper) and light chain (lower) genes selected for detailed investigation. 4-3 is an anti-RBD antibody encoded by IGHV3-66 that does not compete with ACE2, and serves as a heavy chain gene-matched control. See also Figure S3, Table S2, and Table S3.



E

Healthy Donor	*HV3-53/HV3-66 Antibody Precursor Frequency*			
	IGH	IGK	IGL	VH:VL
HIP1	5.06E-04	6.89E-02	5.92E-03	2.21E-05
HIP2	4.31E-04	6.32E-02	4.70E-03	1.72E-05
HIP3	7.92E-04	5.89E-02	3.53E-03	2.91E-05

Average Precursor Frequency

2.28E-05, or 1 in 43,800

Figure 12. Heavy and light chain analyses reveal critical contributions of both VH and VL for potent antibody neutralization in the IGHV3-53/3-66 antibody class.

Fig. 12. Heavy and light chain analyses reveal critical contributions of both VH and VL for potent antibody neutralization in the IGHV3-53/3-66 antibody class. (A) Heavy and light chain swap neutralization panel produced from four mAbs (1-20, 910-30, B38, and 4-3 included as an IGHV gene control) assayed for SARS-CoV-2 pseudovirus neutralization. Data are represented as mean ± SD. **(B)** dhACE2 competition ELISA against SARS-CoV-2 S2P protein. Constant concentrations of heavy-light-swapped IgG were titrated with varying dhACE2 (ACE2) concentrations. dhACE2 concentration is provided in μg/mL, and also as ACE2 molar excess units. Data are represented as mean ± SD. **(C)** Combined structure and sequences analyses reveal IGHV3-53/3-66 class light chain kappa (left panel) and lambda (right panel) genetic elements associated with the RBD contact interface. CDR-L1 residues are not specific to IGKV1-33 (910-30) and IGLV2-8 (C105) (Table S4). **(D)** Left panel: structural superposition of IGHV3-53/3-66 Fab variable domains in complex with RBD shows the same binding orientation for 8 different class antibodies aligned on RBD. RBD is shown in blue, Fab heavy chain in solid color, Fab light chain in transparent color, according to antibody name and the PDB code (shown in parentheses). Right panel: close-up views of the Fab:RBD interface for the eight IGHV3-53/3-66 antibodies superimposed on RBD. Conserved interactions of CDR-H1, -H2 and -L1 define the structural signatures responsible for viral neutralization by the IGHV3-53/3-66 antibody class. **(E)** Estimated probability of IGHV3-53/3-66 class pre-cursor antibodies derived from healthy donor (HIP1, HIP2, HIP3) immune repertories. See also Figure S3 and Table S4.

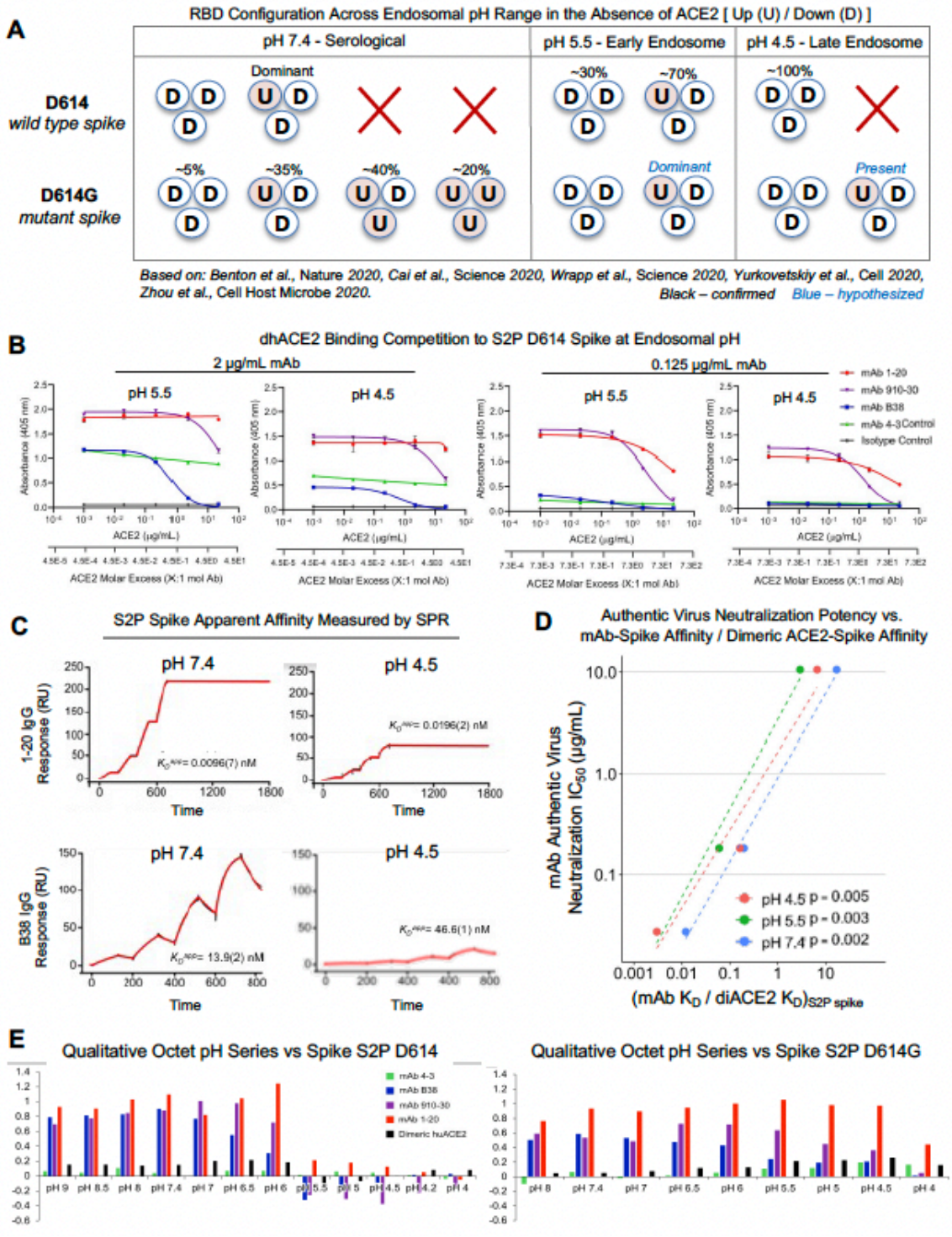


Figure 13. Up/down conformational changes of RBD influence IGHV3-53/3-56 antibody class recognition of spike protein across the serological to endosomal pH range.

Fig. 13. Up/down conformational changes of RBD influence IGHV3-53/3-56 antibody class recognition of spike protein across the serological to endosomal pH range. (A) Schematic of RBD conformational states inferred by Cryo-EM and experimental analysis for un-ligated D614 and D614G spike. U and D denote 'up' and 'down' RBD configurations, respectively. Percentages denote observed particle populations. **(B)** dhACE2 competition ELISA at endosomal pH against SARS-CoV-2 S2P protein. Constant IgG concentrations were used with varying dhACE2 (ACE2) concentrations at pH 5.5 and 4.5. dhACE2 concentration is provided as both μg/mL and as ACE2 molar excess units. Data are represented as mean ± SD. **(C)** Single-cycle SPR kinetic assays for 1-20 and B38 IgG at serological and endosomal pH against biotinylated spike. Black traces represent experimental data and red traces represent the fit to a 1:1 interaction model. The number in parentheses represents the error of the fit in the last digit. **(D)** Correlations between authentic virus neutralization potency (from Fig. 2D) versus the ratio of antibody IgG affinity to spike S2P (from Fig. S4A) divided by dhACE2 affinity to spike S2P (reported from Zhou et al, 2020b). **(E)** Qualitative octet pH series for wild type S2P spike and escape mutant D614G S2P spike across a range of pH values. More potent IGHV3-53/3-66 class members retained binding at low pH against spike as compared to less potent class members. See also Figures S3 and S4.

3. Highly protective anti-malarial antibodies via precision yeast display library generation and screening[2]

3.1 Introduction and Rationale: Antibody Optimization

Antibody Optimization – Evaluate the mutational fitness landscape of protective monoclonal antibody variable regions to screen for mutations enhancing antibody potencies while revealing molecular features correlated to disease protection. Molecular-level understanding of acquired pathogen immunity can be enhanced by studying protective mAbs (*161-163*). Here we hypothesize that a limited amount of amino acid substitutions can further improve a neutralizing mAb potency and/or breadth. We test this hypothesis using a protective anti-malaria mAb: CIS43 (*164*). The monoclonal antibody CIS43 targets the *Plasmodium falciparum* circumsporozoite protein (PfCSP) and prevents malaria infection in humans following single intravenous administration for up to 9 months. To enhance the protective potency and clinical utility of CIS43, we used iterative site-saturation mutagenesis and DNA shuffling to generate precise gene-variant yeast display libraries to screen for improved recognition of PfCSP antigens. We identified several mutations – predominately in framework regions – that improved recognition and combined these to produce a panel of antibody variants. The most potent antibody, CIS43_Var10, had three mutations and showed ~10-fold enhanced *in vivo* protection compared to CIS43. Co-crystal and cryo-electron microscopy structures of CIS43_Var10 with the peptide epitope or PfCSP, respectively, revealed functional roles for each of these mutations. The unbiased site-directed mutagenesis and screening pipeline combined with functional antibody analyses described here represents a powerful approach to enhance protective potency and enable broad clinical use of anti-malarial antibodies.

[2] Bailey B. Banach (BBB), *et al*. Highly protective anti-malarial antibodies via precision yeast display library generation and screening. Designed the experiments: BBB, Prabhanshu Tripathi, Thuy Duong Nguyen, Marie Pancera, Azza H. Idris, Robert A. Seder, Peter D. Kwong, and Brandon J. DeKosky; Performed the experiments: BBB, Prabhanshu Tripathi, Lais Da Silva Pereira, Jason Gorman, Marlon Dillon, Patience K. Kiyuka, Thuy Duong Nguyen, Barbara Flynn, Joseph R. Francica, Nicholas K. Hurlburt, Tracy Liu, Li Ou, Reda Rawi, Arne Schön, Chen-Hsiang Shen, I-Ting Teng, and Baoshan Zhang; Analysed the data: BBB, Prabhanshu Tripathi, Jason Gorman, Ahmed S. Fahad, Thuy Duong Nguyen, Joseph R. Francica, Reda Rawi, Arne Schön, Chen-Hsiang Shen, Azza H. Idris, Robert A. Seder, Peter D. Kwong, and Brandon J. DeKosky; Writing: BBB and Brandon J. DeKosky. Reviewing & Editing: all authors. *In submission.*

Background

Malaria affects 200 to 400 million people each year, causing nearly 400,000 annual deaths (*165*). *Plasmodium falciparum* is the primary parasitic strain that causes malaria morbidity and mortality in Africa. Malaria infection is initiated following a mosquito bite in which sporozoites, the infectious form of *P. falciparum,* are injected into the host and quickly migrate to the liver to infect hepatocytes. A range of treatment and control solutions currently exist, yet global control of malaria remains a persistent challenge and new preventive solutions are urgently necessary to curb continued transmission and disease. Solutions that neutralize initial sporozoite infection have the capacity to fully interrupt the infection cycle, preventing progression to liver stage disease and providing a promising approach for sterilizing immune protection.

The *P. falciparum* circumsporozoite protein (PfCSP) is the most prevalent antigen on the surface of sporozoites and a major target of antibody responses that confer protection against malaria infection, though generally at very high antibody titers (*166*). PfCSP contains numerous amino acid repeat regions; a dominant 'NANP' repeat motif comprises much of the PfCSP central region, a minor 'NANPNVDP' repeating sequence occurs four times in the central region, and a unique junctional epitope containing one 'NPDP' motif bridges the N-terminal region with the central region. Recently, a number of naturally- or vaccine-elicited human monoclonal antibodies (mAbs) were revealed to target key regions in PfCSP to provide sterilizing protection against malaria infection following passive transfer in mice (*28, 166-169*). Moreover, in a recent human clinical trial, 40 mg/kg of intravenously delivered anti-$PfCSP_{junction}$ mAb: CIS43LS, modified with the half-life extending "LS" Fc-mutation, provided protection for up to 9 months against controlled malaria challenge in a small number of participants (*170*).

These data support the utility of CIS43 and other PfCSP-targeting mAbs for long-acting and effective malaria prophylaxis. Production costs of protein drugs, however, remain substantially greater than the production cost for small molecules, creating economic barriers to large-scale clinical use of anti-malarial antibodies. Improvements in antibody potency should reduce therapeutic dose, thereby providing flexibility in route of administration and potential cost savings as well as increasing the global availability of antibody-based anti-malarial solutions.

Prior reports by our group (*171*) and others (*172-175*) have demonstrated that screening precision site-saturation mutagenesis (SSM) antibody libraries can lead to efficient drug potency improvements, especially when combining multiple mutations for synergistic effects. Importantly, these precision libraries can sample rare variants that are not frequently observed in natural *in vivo* antibody evolution (*171, 172*), including variants that require multiple nucleotide substitutions within a single codon, and potentially rare combinations of mutations distributed throughout the entire antibody variable region. Iterative precision library generation and *in vitro* functional screening allows stepwise introduction of potentially beneficial mutations to the parent sequence for isolation of improved variants with fewer modifications than alternative error-prone PCR methods and mouse models. This efficient approach maintains high sequence homology with native human antibody genes and reduces the risk of immune rejection associated with long-term clinical use of highly mutated genes (*176-178*). However, most protective anti-PfCSP antibodies recognize non-overlapping repeat-peptides with high affinity, confounding the native human immune response, and it has been unclear how to optimize the complex protein interactions to discover anti-malarial antibodies with sufficient potency for broad clinical use.

In this study, we report the application of *in vitro* precision antibody engineering techniques to improve the protective potency of antibody CIS43 by enhancing its binding interactions to *P. falciparum* CSP-derived antigens. We found that ~10-fold improvement of the already potent CIS43 monoclonal antibody could be achieved with three engineered mutations, and that variants with enhanced protection displayed higher affinity to full-length PfCSP and the CIS43 junctional epitope. To provide insight into the mechanism by which identified mutations improved protective function, we determined co-crystal and cryo-electron microscopy (cryo-EM) structures of improved antibodies with junctional peptide epitope and PfCSP, respectively. These data describe a pipeline for efficient antibody engineering with an overall aim of providing substantially improved antibodies that can be used widely for cost-effective prevention of infectious diseases such as malaria.

3.2 Results

Anti-malarial antibody library screening enhances recognition of CSP antigens

We sought to improve the potency of CIS43 while introducing a minimal number of mutations to the parent sequence. Accordingly, we applied precision antibody library generation and screening strategies for efficient improvement of CIS43 anti-malarial activity (**Fig. 18**). We used sequential gene library generation and screening with functional yeast display and Next-Generation Sequencing (NGS) to scan the CIS43 heavy and light chain variable regions for potentially beneficial mutations (Fig. 1A). Gene libraries containing all possible single amino acid (AA) substitutions were first generated by site-saturation mutagenesis (SSM) over the heavy chain variable gene (VH) and light chain variable region gene (VL) separately (*171, 179*). The resulting single mutation DNA libraries, (VH-SSM):VL and VH:(VL-SSM), were cloned into plasmids for yeast surface expression as fragment antigen binding (Fab) libraries while ensuring >100-fold SSM library coverage (**Appendix I Figure 23A, Appendix O Table 5**). Plasmid libraries were used to transform genetically engineered yeast to express Fabs via a galactose-inducible bidirectional promoter followed by screening for antigen recognition with PfCSP-derived antigen probes (**Fig. 18B, Appendix I Figure 23B-C, Appendix P Table 6**) (*171, 180*).

We hypothesized that enhanced affinity to CSP antigens would improve antibody potency. Accordingly, Fab libraries were stained with the CSP-derived antigen probes (**Fig. 18B, Appendix P Table 6**) and fractionated *en masse* into high-, medium-, or low-affinity binding populations against full length PfCSP (fl_PfCSP) and Peptide 21 (Pep21) via fluorescence-activated cell sorting (FACS) (**Fig. 18C, Appendix I Figure 23B, D-E**). Affinity phenotypes were determined for surface-expressed Fab by comparing the relative intensities of Fab expression and antigen binding (*171, 180*). Affinity-fractionated single-mutation libraries were enriched for a total of three rounds of FACS. The third round high-affinity sorted libraries showed minimal observable differences in binding phenotypes compared to the monoclonal CIS43 control (**Fig. 18C, Appendix I Figure 23E**), suggesting that single mutation variants had a limited capacity to improve function. Bioinformatic analysis of the Round 3 high-affinity libraries revealed enriched single mutants predicted to be beneficial for PfCSP-antigen recognition (**Appendix J Figure 24A-**

C), however, ELISA characterization of selected single mutation soluble antibodies suggested that those single mutants contained only limited improvements to antigen recognition (**Appendix J Figure 24D, Appendix Q Table 7**).

To identify additional synergistic mutation combinations that improved antigen recognition, we next generated and screened multi-mutation gene libraries (**Fig. 18A** and **18D, Appendix K Figure 25**). We hypothesized that enriched mutations from high-affinity libraries represented some beneficial progress toward improved antigen recognition and an elimination of deleterious single mutations, and therefore we used the single mutation Round 3 high-affinity libraries as template to generate multi-mutation libraries (**Fig. 18A** and **18C**). Multi-mutation Library 1 was generated by extracting the enriched VH and VL genes from template yeast libraries and subcloning into a shared plasmid backbone to create combinatorial heavy and light chain variants (**Appendix K Figure 25A**). To increase library diversity, DNA shuffling was also performed on the single mutation template VH and VL DNA pools, separately, to generate combinations of enriched mutations within the VH and VL genes (*181*). Shuffled-VH and shuffled-VL genes were then subcloned into a single plasmid backbone to generate a combinatorial & shuffled-VH:VL multi-mutation (Library 2). Finally, an additional degree of diversity was included using the shuffled-VH:VL genes as template to repeat SSM ("re-SSM"), and we subcloned the re-SSM genes into another plasmid backbone to generate additional paired VH:VL genes with multiple mutations (Library 3).

These multi-mutation plasmid libraries were used to transform yeast, maintaining at least 1 x 10^7 library transformants at each step (**Appendix O Table 5**). Expression was induced in the three multi-mutation yeast libraries, and surface-expressed Fabs were screened against five CSP-derived antigens (**Fig. 18B, Appendix P Table 6**). We enriched high-affinity variants across four rounds of FACS to generate over 60 sorted multi-mutation samples (**Fig. 18D**). Multi-mutation libraries from the fourth round of high-affinity screening showed significant improvements to antigen binding (**Fig. 18D**) and NGS data was mined to identify enriched sequences predicted to enhance anti-malaria binding (**Fig. 19B**, **Appendix K Figure 25B, Appendix Q Table 7**).

Single mutation variants confer limited improvements in antigen recognition

NGS analysis of yeast-transformed single-mutation libraries revealed robust representation of amino acid substitutions prior to sorting (**Appendix I Figure 23A**). Bioinformatic analysis of the yeast-surface expressed libraries (measured as the VL+ expressing population) (**Appendix I Figure 23B**) revealed that nearly all 7,488 theoretical single mutation variants expressed at detectable levels in the Fab-expressing pre-sort libraries (**Appendix I Figure 23C**). Prior to antigen screening, single mutation libraries showed generally comparable binding to monoclonal CIS43 (**Appendix I Figure 23D**). Post-screening, high-affinity single mutation libraries still showed limited observable antigen recognition differences from monoclonal CIS43 (**Fig. 18D**), suggesting that the enriched single mutants provided limited functional improvements. In contrast, the low-affinity single-mutation libraries showed clear phenotypic binding differences from monoclonal CIS43 (**Appendix I Figure 23E**), with a sharp loss of antigen recognition after three rounds of low-affinity sorting. Medium-affinity sorted libraries maintained consistent binding characteristics with monoclonal CIS43.

The phenotypic differences between libraries that we observed by flow cytometry suggested that single mutation libraries had been successfully fractionated according to their binding affinity phenotypes. Bioinformatic analysis of NGS data was next used to quantitatively track single amino acid (AA) mutations across each sort condition to identify enriched variants of interest (**Fig. 19A, Appendix J Figure 24A-B**). We enumerated the prevalence and enrichment ratio (ER) of each single AA mutation to quantify variant performance across sorted populations, binning each single mutation variant into affinity bins based on ER (**Fig. 19, Appendix J Figure 24A-B**). Affinity-binned groups revealed consistencies in sort data against malaria antigens (**Fig. 19A**) and most sequences that strongly enriched in high-affinity sorts were depleted in low-affinity sorts and not strongly enriched in medium affinity sorts (**Appendix J Figure 24B**). These data agreed with the phenotypic observations by flow cytometry that our screening strategy had fractionated the single mutation gene variant populations according to binding phenotypes.

Single mutation variants with ER >10 in high-affinity sorts and <1 in low affinity sorts were predicted to have improved antigen binding. Based on bioinformatic analysis of NGS data, 26 single mutation variants and 4 rationally combined multi-mutation variants were selected for expression and

biophysical characterization as soluble antibodies (**Fig. 19A, Appendix J Figure 24B, Appendix Q Table 7**). Notably, multiple substitutions at the second residue in framework region 1 (FR1) of the VH gene showed strong enrichment (**Fig. 2A**, *top*), where a rare valine to arginine substitution (VH_V2R) was highly enriched against Pep21 while the unmutated template CIS43 sequence was simultaneously depleted in this same library (**Appendix J Figure 24A,** *2nd graph*), suggesting that the VH_V2R was able to outcompete the template CIS43 protein via FACS. Top additional VL single mutations of interest included (if different, *Kabat numbering* in italics*)*: VL_A19V (FR1), VL_N22H (FR1), VL_N51R/*N45R* (FR2), VL_S82D/*S76D* (FR3), VL_L84F/*L78F* (FR3), and VL_V91D/*V85D* (FR3), each having enrichment ratios >10 in the fl_PfCSP and Pep21 high affinity sorts and <1 in low affinity sorts (**Fig. 19A** *bottom*, **Appendix J Figure 24B, Appendix Q Table 7**). However, we also observed that the unmutated CIS43 template gene was present and/or remained dominant in several of the single-mutation high-affinity screens (**Appendix J Figure 24A**), suggesting that in many cases the enriched single mutations did not substantially enhance antigen affinity over unmutated CIS43.

Structural mapping of FACS-predicted high-affinity single mutation variants showed that a majority of enriched residues were structurally distal to the antigen-binding paratope in the framework regions, and only a select few mutations were enriched in the complementary determining regions (CDRs) (**Fig. 19A, Appendix J Figure 24C,** PDB: 6B5M). ELISA analysis of the top FACS-predicted high-affinity single mutation variants showed comparable antigen recognition for individually expressed IgG single mutation variants versus parent CIS43 (**Appendix J Figure 24D**), consistent with the limited phenotypic improvements observed in high-affinity FACS populations (**Fig. 18C**). These limited improvements to antigen recognition suggested that single mutations would also confer limited improvements in protection, and that multi-mutation combinations would be required to enhance antigen recognition.

Multi-mutation variants show improvements in antigen recognition

The three multi-mutation VH:VL Fab libraries were next screened for four rounds of high-affinity binding against five CSP-derived antigens (**Fig. 18A-B, Fig. 19B, Appendix J Figure 24**), and Round 4 sorted samples showed substantial phenotypic improvements to antigen recognition via flow cytometry

(**Fig. 18D**). A comparison of antigen stains between high-affinity groups with single mutations (**Fig. 18C**) and multi-mutations (**Fig. 18D**) revealed that multiple mutations dramatically improved binding over single mutants alone. NGS data was mined to quantitatively track multi-mutation variants across sort conditions (**Fig. 19B**, **Appendix K Figure 25**). High-affinity sorted multi-mutation libraries showed a higher frequency of single and double mutations on the VH and VL in all sorted libraries when compared to the pre-sort population. (**Appendix K Figure 25A, Appendix R Table 8**), and the unmutated CIS43 template gene was depleted from the sorted population in all multi-mutation sorts (**Appendix K Figure 25B**). Certain multi-mutation variants were enriched across several sorting conditions, suggesting that our iterative library screening converged on variants with strongly improved malaria antigen recognition. Evaluation of multi-mutation library NGS data identified several of these enriched antibody sequences of high interest (**Fig. 19B**, **Appendix K Figure 25B, Appendix Q Table 7**), and the top 24 multi-mutation variants with ER >10 against multiple antigens and high prevalence values in Round 4 sorted libraries were selected for expression as soluble antibodies for functional & biophysical characterization (**Appendix Q Table 7**).

Multi-mutation variants provided robust protection against malaria challenge *in vivo*

We next sought to evaluate the level of protection provided by CIS43 variants with improved affinity. As there are no reliable *in vitro* functional assays for anti-malaria antibody potency, we used an *in vivo* murine challenge model to quantify passively administered antibody protection (*169*) (**Fig. 20A**). These *in vivo* challenge studies are limited by the number of experimental animal groups evaluated at one time, so we prioritized the variants in our 24 multi-mutation panel based on affinity to the NPDP19 junctional peptide. Variant affinity was determined based on the AlphaLISA bead proximity luminescent signal, and the top ten multi-mutation variants demonstrating enhanced NPDP19 recognition by AlphaLISA were selected for challenge evaluation (**Fig. 20B**).

To evaluate the capacity of mutant antibody variants to prevent malaria infection, mice were injected with anti-malaria antibodies intravenously two hours prior to intravenous injection with 2,000 malaria sporozoite particles that had been genetically modified to encode GFP and luciferase markers. At two days post-challenge, mice were injected with d-luciferin and analyzed with IVIS Spectrum *in vivo*

imaging to quantify luminescence in the liver associated with liver stage infection from sporozoites. We similarly analyzed luminescence at 6 days post-challenge for mice with substantial protection at Day 2, which provided a highly sensitive evaluation of the level of parasitemia in challenged mice.

Challenge studies showed that single mutation variants provided some protection (**Fig. 20C**), but this protection was not enhanced over the original CIS43, which was consistent with the limited affinity enhancements observed by flow cytometry and ELISA (**Fig. 18C**, **Appendix J Figure 24D**). In contrast, several multi-mutation variants provided robust protection against malaria infection (**Fig. 20D**, **Appendix L Figure 26**) consistent with improvements in antigen recognition observed by flow cytometry (**Fig. 18D**). Variants 2, 10, 11, and 18 were significantly more potent than CIS43 in protecting against liver burden at Day 2 after challenge, with CIS43_Var10 showing over 10-fold improvement versus the parent CIS43 molecule. Variants 2, 10, and 11 also showed statistically significantly lower parasitemia at 6 days post-challenge than CIS43 (**Fig. 20E-F**).

Based on the promising functional data reported for CIS43_Var10 we sought to compare its efficacy to other protective mAbs. A recently isolated CSP-targeting human mAb, L9 (*169*), is 2-3 fold more potent than CIS43 and is the current "best-in-class" benchmark for protective capacity of anti-malaria antibodies. In our *in vivo* functional challenge study CIS43_Var10 was ~3-fold more potent than the antibody L9 (*169*) when comparing parasitemia levels at 6 days post-challenge (**Fig. 20D, 20F, Appendix L Figure 26**). These *P. falciparum* challenge studies demonstrated that our antibody engineering pipeline could improve CIS43 monoclonal antibody potency by ~10-fold with only 3 introduced mutations (VH_V2R, VL_L30F/*L27$_c$F*, and VL_V91S/*V85S*), and confirmed that several other multi-mutation combinations could substantially improve antibody performance with very few mutations introduced to the CIS43 template antibody.

Structural basis of improved antibody-based protection

Having identified four top CIS43-variant antibodies (Variants 2, 10, 11, and 18), we next sought to determine the functional and structural mechanisms leading to their improved protective efficacy (**Fig. 21**). Variants 2, 10, 11, and 18 contained combinations of six unique substitutions: VH_V2R, VH_A97V/*A93V*,

VL_L30F/*L27$_c$F*, VL_N51R/*N45R*, VL_Q61A/*Q55A*, VL_V91S/*V85S* (**Appendix Q Table 7,** template numbering/*Kabat numbering*). We co-crystallized CIS43_Var2 and CIS43_Var10 with Peptide 21 containing the PfCSP major antigenic NPDP junctional epitope, and determined the structures of CIS43_Var2 and CIS43_Var10 in complex with Peptide 21 at 1.40 Å and 1.55 Å resolution, respectively (**Fig. 21A, Appendix S Table 9**). We analyzed these data, along with modelled structures of CIS43_Var11 and CIS43_Var18, to compare structural features of variant antigen recognition (**Fig. 21A**).

Notably, while two of the six mutations providing the basis for the top four improved variant function occurred in complementary determining regions (CDRs): VL_L30F/*L27$_c$F* (CDR-L1) VL_Q61A/*Q55A* (CDR-L2), none of the six mutations made direct contact with Peptide 21 (**Fig. 21B**). In addition, two of the six mutations (VH_V2R and VL_V91S/*V85S*) were rare mutations infrequently observed in natural antibody sequences (**Appendix M Figure 27**). A comparison of CIS43_Var2 and CIS43_Var10 structures over parent CIS43 showed few structural differences; in CIS43_Var2, the arginine in VH_V2R interacted with Asp100$_B$, which appeared to alter the position of the CDR-H3 by 2.2 Å (as measured at the β-carbons of Asp 100$_B$), with a resulting change in the orientation of the C-terminal region of Peptide 21 to accommodate this movement (**Fig. 21B**, *left*). Notably, in CIS43_Var10, VH_V2R induced a smaller alteration in the CDR-H3, perhaps influenced by the neighboring VL_V91S/*V85S*, and Peptide 21 assumed its canonical position (**Fig. 21B**, *right*).

Improved performance for the most protective antibody variants appeared to be associated with increased binding affinity against the junctional NPDP epitope. Overall, we observed little modification in stoichiometry for the second binding event against the minor repeat NANPNVDP contained in Peptide 21 and Peptide 25 (**Fig. 21C**, **Appendix N Figure 28**). Affinity assays quantified the degree of improved binding for top variants, with bio-layer interferometry (BLI) showing tighter affinity to fl_PfCSP, NTDS_5/3, Peptide NDPD19, Peptide 21, and Peptide 25, but not to the major repeat NANP sequence contained in Peptide 29 for the most improved variants (**Fig. 21C**, **Appendix N Figure 28A**). Isothermal calorimetry (ITC) analysis of CIS43 and engineered antibody variants binding to CSP showed the $K_{D,1}$ of the top 4 variants (relating to interaction with the junctional peptide) showed enhanced affinity compared

to CIS43 (**Appendix N Figure 28B**). Interestingly, CIS43_Var18, which contained both key mutations that occurred in the CDRs (VL_L30F/*L27$_c$F* in CDR-L1, and VL_Q61A/*Q55A* in CDR-L2), showed substantially improved affinity to Peptide 21, but only marginally improved affinity to NPDP19 (**Fig. 21C**), indicating differential improvement in affinity to highly similar peptides (**Appendix P Table 6**). Furthermore, we observed CIS43_Var2 (containing VH_V2R and VL_L30F/*L27$_c$F*,) and CIS43_Var10 (containing VH_V2R, VL_L30F/*L27$_c$F*, and VL_V91S/*V85S*) to be closely matched to one another in affinity for Peptide 21 and Peptide 25. However, the more functionally improved variant, CIS43_Var10, had stronger affinity to Peptide 29 than CIS43_Var2, suggesting synergistic contributions from the VL_V91S/*V85S* substitution were necessary for potent anti-malaria function related to recognition of the major NANP repeat.

While the crystal structure revealed the VH_V2R mutation to orient the position of the CDR-H3, the importance of the VL_L30F/*L27$_c$F* was not immediately clear as it faced away from the bound peptide. To further investigate possible interactions, we determined the cryo-EM structure of CIS43_Var10 Fab in complex with fl_PfCSP (**Fig. 22**). 2D-class averages showed clear Fab signal, with additional signal shown extending from the tip of the Fv region corresponding to bound, but disordered fl_PfCSP (**Fig. 22A**). 3D reconstructions for multiple classes were determined at 6.1-10.9 Å resolution (**Fig. 22B, Appendix T Table 10**). Two of the classes were used for rigid-body fitting with the Peptide 21-bound crystal structure of CIS43_Var10 and showed the VL_L30F/*L27$_c$F* facing away from the peptide but directly into a reconstruction density attributed to fl_PfCSP (**Fig. 22B**). Multiple classes were also obtained for two Fabs linked by their co-binding of the same fl_PfCSP in close proximity (**Fig. 22C**). These data support a focus on improving affinity to the junctional epitope of CIS43 as a site of potent immune protection and emphasize full length PfCSP as an important target for sterilizing antibody-based protection against malaria infection.

3.3 Discussion

Here we show that precision antibody engineering efforts can effectively enhance the potency of anti-malarial antibodies. CIS43_Var10 provided ~10-fold improved protective potency against malaria challenge compared to CIS43, and these substantial functional improvements were achieved with minimal mutations to the CIS43 template antibody. Our study outlines a new pathway for highly effective mAb engineering for antibody-based protection against malaria infection and defines several critical features of PfCSP as vaccine and antibody targets.

The complex features of anti-PfCSP antibody responses have made it historically challenging to discover antibodies with optimized molecular interactions sufficient for clinical use. Most protective anti-PfCSP antibodies recognize non-overlapping repeat peptides with high affinity and low SHM (*28, 166-169*); prior studies have suggested little correlation between SHM and immune protection for PfCSP-specific antibodies (*166, 182*). Structural diversity and sequence degeneracy associated with the long-range, extended spiral conformation of PfCSP's repeat regions have also been shown to induce homotypic Fab-Fab interactions in potent natural antibodies (*183-186*). In light of these studies, our antibody improvement campaigns demonstrated that enhanced binding affinity to the junctional epitope and full length PfCSP together can effectively improve protective potency with minimal mutations to the antibody sequence, even for naturally potent antibodies like CIS43.

We found that the introduction of precision mutations for functional screening, paired with combinatorial shuffling, can lead to strong potency improvements by eliciting rare combinations of mutations that target critically important antigen structural sites. The most improved variant, CIS43_Var10, comprised three mutations (VH_V2R, VL_L30F/*L27$_c$F*, and VL_V91S/*V85S*) and provided ~10-fold improved protective potency against malaria challenge compared to CIS43. Notably, while one of these mutations occurred in a CDR (VL_L30F/*L27$_c$F*, CDR-L1), the crystal structure revealed that none of the three mutations made direct contact with Peptide 21. CryoEM analysis of VL_L30F/*L27$_c$F* facing into density attributed to the PfCSP suggests that this mutation interacts beneficially with a region not observed in the peptide crystal structure; the density extended from the Fv region of the Fab and is indicative of

fl_PfCSP maintaining some structure in the areas close to or directly interacting with the Fab, but becoming structurally disordered as it extends outward. Furthermore, the observation of multiple classes with two Fabs co-binding closely to the same fl_PfCSP strongly suggest that the two arms of an IgG would display bivalent binding that would provide higher avidity. Interestingly, the other two CIS43_Var10 mutations (VH_V2R and VL_V91S/*V85S*) are both rare mutations that require multiple nucleotide substitutions within a single codon, and are infrequently observed in natural antibody sequences. These data highlight the advantages of performing screens with SSM to easily sample rare mutations throughout the entire variable region, and suggest natural anti-malaria antibody development may be limited by incomplete sampling of rare mutations *in vivo* (*31, 187, 188*) (**Appendix M Figure 27**).

Another advantage of performing first-round screens via SSM may be that the comparatively small library size for SSM (~10^4) can easily be screened with adequate coverage in yeast display, and any mutations that are deleterious to binding (up to 50% of all possible mutations as shown in our previous studies (*171*)) can be eliminated prior to combinatorial shuffling. Thus, in early rounds of single-mutation high-affinity screens, detrimental mutations are removed while promising mutations are enriched for subsequent recombination. By beginning recombination with a smaller, high-quality single mutation pool, the DNA shuffling and combinatorial mutagenesis steps have greater efficiency and an enhanced probability of identifying synergistic mutations. In contrast, other traditional and less precise library generation techniques like error-prone PCR can introduce too much mutational variation per round, increasing the likelihood of introducing deleterious or nonbeneficial mutations. Error-prone PCR also has difficulty sampling important rare mutations in isolation, some of which may require 3 nucleotide substitutions, but which are easily sampled via SSM (*171*). DNA shuffling of heavy and light beneficial single mutants provides great advantages over natural and *in vivo* B cell-based antibody discovery by combining mutations either across the paired heavy and light chain variable regions, as well as within each heavy/light variable region, and selecting for combined beneficial mutations with an enhanced probability of synergistic effects. One additional advantage of precise and efficient antibody engineering is that fewer mutations are introduced compared to alternative methods, which can reduce drug rejection risks (*176,*

178). Minimizing mutational divergence from human genes is a critical design parameter for antibody drugs (and generally for protein therapeutics), especially for those drugs with a long circulation half-life given to immunocompetent individuals, which is a major use case for highly potent anti-malarial antibodies. Subsequent rounds of precision mutational screening could yield even further benefits beyond the 10-fold enhanced potency reported here, and other monoclonal antibody clones have already been substantially enhanced using the same improvement pipeline in parallel efforts by our group and others.

Our efforts defined key epitopes on PfCSP as potential antigens for vaccine and antibody targets. In particular, the junctional epitope of PfCSP site is a promising target for additional potency improvements. Future antibody engineering efforts will lead to further improvements, both for CIS43 and other antibody templates such as L9 (*169*), particularly in combination with other advanced methods for antibody discovery and screening like engineered mouse models (*189-191*). Improved potency of anti-malarial mAbs should enable reducing dosage requirements, which may allow for subcutaneous (versus intravenous) administration and potential reductions in health care costs. In addition, the use of Fc mutations to enhance half-life could lead to protection with only a single subcutaneous injection over a period of 6-12 months (*170*), which would be optimal for seasonal and perennial control in infants and children. Transient antibody expression technologies, including mRNA-Based gene delivery (*192, 193*) and vectored immunoprophylaxis (*194, 195*), could provide additional methods of antibody delivery, and the improved efficacy obtain here with CIS43_Var10 may enable the assessment of lower dose antibody protection delivered by these alternative methods. We expect that precision antibody engineering strategies will be used increasingly to enhance the potency of antibody drugs targeting infectious diseases, and that combinations of improved molecules and drug delivery strategies will make broadly accessible antibody-based treatments a tangible reality to curb global disease transmission in the coming decades.

3.4 Methods

SSM Library Generation and Cloning into Yeast Display

Single-mutation plasmid libraries were constructed using a previously reported one-pot site saturation mutagenesis (SSM) protocol (*81*). CIS43 heavy chain variable region (VH) and light chain variable region (VL) gene libraries containing a single amino-acid substitution were separately generated using mutagenic primers that contained degenerate single codons to express all 20 amino acids at each residue of the VH and VL. A process of single-strand nicking, exonuclease digestion, and degenerate oligonucleotide-primed PCR was used to generate each library (VH and VL) containing the comprehensive set of single-residue substitutions across the CIS43 variable regions (*196*). The separate (VH-SSM):VL and VH:(VL-SSM) libraries were cloned into a yeast display vector containing a FLAG-marker to quantify Fab surface expression, as well as a leucine zipper and protein disulfide isomerase expression for enhanced expression of diverse antibody libraries (*20*). The plasmid DNA libraries were used to transform yeast populations, and library sizes of at least 4×10^5 were maintained in all cloning steps, as described previously (*78*).

Fluorescence-Activated Cell Sorting (FACS) of Single-Mutation Libraries

Transformed yeast libraries were cultured in SGDCAA medium to induce antibody fragment antigen binding (Fab) surface expression, (20 g/L galactose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4.H_2O$; SGCAA from TEKnova, Hollister, CA) supplemented with 2 g/L dextrose (SGDCAA) for 36 h at 20°C and 225 rpm. Yeast libraries were washed and stained with an anti-FLAG FITC monoclonal to quantify Fab expression (F4049, Clone M2, Sigma-Aldrich, Burlington, MA). Full length *plasmodium falciparum* circumsporozoite protein (fl_PfCSP) and Peptide 21 (NPDPNANPNVDPNAN) antigen probes were generated by fluorescently labeling the biotinylated constructs with an anti-biotin-PE label (12-9895-82, clone BK-1/39, Thermo Scientific, Waltham, MA). To screen single mutation Fab libraries for antigen-recognition, anti-FLAG-FITC-labeled libraries were co-stained with fluorescently conjugated antigen probes. In the first round of FACS, 3e7 yeast cells were

stained and sorted using three gates to sort and collect low, medium, and high affinity yeast library groups by the ratio of Fab surface-expression versus Ag binding for the cell population, as described previously (*77, 78*). For all sorting experiments, control libraries of FITC+ yeast were also collected to track the identity and initial prevalence of library variants with successful Fab expression. Collected cells were grown in low-pH SDCAA (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 10.4 g/L trisodium citrate, and 7.4 g/L citric acid monohydrate, pH 4.5) for 24-48 h in a 30°C incubator shaking at 225 rpm. After sorted yeast collection and culture, each sort was repeated (using either low, medium, or high affinity, to match the initial Round 1 sort) for an additional two rounds.

Multi-Mutation Library Design and Screening

Plasmid libraries were isolated from sorted yeast cells using previously described DNA extraction methods (*196*). VH and VL genes were amplified from library plasmids using Kapa Hifi HotStart ReadyMix (Kapa Biosystems, Roche. Wilmington, MA) (*20*). Three VH:VL paired combinatorial multi-mutation libraries were designed using template DNA from the enriched single-mutation high-affinity screens. Library 1 was generated by combinatorial pairing of enriched single-mutations VH and VL genes using restriction enzyme digest and sub-cloning. Library 2 was generated by performing DNA shuffling with enriched single-mutation VH and VL gene libraries, separately (*79*). Template DNA was fragmented with DNAseI, followed by homologous reassembly and reamplification of the shuffled genetic material_ENREF_7. The separate shuffled-VH and shuffled-VL genes were combined in a single shuffled-VH:VL library via restriction enzyme digest and sub-cloning. Library 3 was generated with an additional round of SSM performed on the separate shuffled-VH and shuffled-VL genes. The re-SSM-shuffled genes were then combined into a single VH:VL library via restriction enzyme digest and sub-cloning.

Yeast libraries were next generated using multi-mutation plasmids, as described above for the single mutation library screens, and library sizes exceeding 1×10^7 were maintained in all cloning steps. Expression was induced and stained as above using an anti-FLAG FITC monoclonal antibody (F4049,

Clone M2, Sigma-Aldrich. Burlington, MA) and a fluorescently labeled antigen probe. A truncated version of PfCSP (NTDS_5/3_CSP) displaying 5 of the major repeat peptides (NANP) with 3 of the minor repeat peptides (NVDP) was also used for screening. NTDS_5/3_CSP was based on the 3D7 strain and contains four N-terminal domain stabilizing (NTDS) mutations that prevent dimer formation by mutating an N-terminal free cysteine in the original sequence from "QEYQC" to "QEYQS" and increase expression by mutating the charged resides in the sequence: "KKNSRSLGE" mutated to "SSNSASLGE" immediately upstream of the "RSLGE" PEXEL II NTD sequence (*169*). Biotinylated Peptide 25 (NVDPNANPNVDPNAN) and Avi-tagged Peptide 29 (NANPNANPNANPNAN) antigen probes were also generated by fluorescently labeling the constructs with an anti-biotin-PE label (12-9895-82, clone BK-1/39, Thermo Scientific. Waltham, MA) and streptavidin-PE (SA-PE, Thermo Fischer Scientific. Waltham, MA), respectively. Multi-mutation libraries were enriched for high-affinity binding as described above for single mutation library screening; 3×10^7 yeast cells were stained and the 0.1% highest affinity variants in the library, as determined by the highest ratio of antigen binding to Fab expression (*77, 78*). Control libraries of FITC+ yeast were also collected to track the identity and initial prevalence of multi-mutation library variants with successful Fab expression. Multi-mutation libraries were enriched for high-affinity antigen binders for an additional three rounds using FACS. Collected cells were cultured in low-pH SDCAA (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 10.4 g/L trisodium citrate, and 7.4 g/L citric acid monohydrate, pH 4.5) for 24-48 h at 30 °C and 225 rpm.

NGS and Bioinformatic Analysis of Sorted Antibody Libraries

VH and VL genes were extracted and amplified from sorted yeast libraries and submitted for Next Generation Sequencing (NGS) on an 2x300bp Illumina MiSeq platform. First, plasmids were isolated from cultured yeast cells as described previously (*196*). VH and VL genes were amplified from library plasmids using Kapa Hifi HotStart ReadyMix (Kapa Biosystems, Roche. Wilmington, MA) as described previously (*20, 78, 197*). Amplified libraries were prepared for NGS with an addition round of PCR to incorporate barcodes and adapters for Illumina sequencing.

Raw Illumina fastq sequence reads were processed as described previously (*78, 197, 198*). NGS reads were quality-filtered for a score of 30 over 90% of the raw reads using Fastxtoolkit (v0.0.14 http://hannonlab.cshl.edu/fastx_toolkit/). Filtered reads were then processed using IgBlast software to reference the IMGT database and determine complete variable region gene alignments (*148, 199*). Once variable region alignments were obtained, mutant CIS43 sequences were aligned to the template CIS43 antibody sequence using Usearch software. Mutations were defined by determining the percent identity match to the template CIS43 gene and denoting the substituted amino acid residue according to the template CIS43 gene sequence, with Base 1 corresponding to the start of the variable framework region. Kabat numbering identifiers are also listed for key mutations and referred to in structural data (Figures 21-22, Appendix Q Table 7). The number of reads for unique sequences were enumerated in each library, and these data was used to determine the prevalence, or frequency, of each variant in the sorted libraries:

$$Prevalence_{Variant\ X\ in\ Library\ Y} = \frac{Number\ of\ Reads\ of\ Sequence\ X\ in\ Library\ Y}{Total\ number\ of\ Reads\ in\ Library\ Y}$$

We defined a variant's enrichment ratio (ER) as the change in sequence prevalence from the initial Fab expressing (VL-FITC+) sorted control library to the experimentally screened library:

$$ER_{Sequence\ X\ in\ Library\ Y} = \frac{Prevalence\ of\ Sequence\ X\ in\ Library\ Y}{Prevalence\ of\ Sequence\ X\ in\ Fab\ Expressing\ Control\ Library}$$

Each single mutation variant was binned into high, medium, and low affinity screened groups by comparing prevalence and enrichment ratio values across different screening conditions, as we reported previously (*77, 78*). Multi-mutation variants were analyzed by their ER in high-affinity sorted library screens, using the multi-mutation VL-FITC+ prevalence in the denominator.

Antibody Expression

Antibody variable heavy chain and light chain sequences were codon optimized, synthesized and cloned into a VRC8400 (CMV/R expression vector)-based IgG1 vector as previously described (*200*). The variants were expressed by transient transfection in Expi293 cells (ThermoFisher Scientific, Waltham, MA) using Turbo293 transfection reagent (SPEED BioSystems, Gaithersburg, MD) according to the manufacturer's

recommendation. 50 μg plasmid encoding heavy-chain and 50 μg plasmid encoding light-chain variant genes were mixed with the transfection reagents, added to 100 ml of cells at 2.5×10^6/ml, and incubated in a shaker incubator at 120 rpm, 37°C, 9% CO2. At 5 days post-transfection, cell culture supernatant was harvested and purified with a Protein A (GE Healthcare, Chicago, IL) column. The antibody was eluted using IgG Elution Buffer (ThermoFisher Scientific, Waltham, MA) and were brought to neutral pH with 1 M Tris-HCl, pH 8.0. Eluted antibodies were dialyzed against PBS overnight and were confirmed by SDS-PAGE before use.

AlphaLISA Characterization of CIS43 Variants

AlphaLISA® (Perkin-Elmer, Waltham, MA) is a bead-based proximity assay in which singlet oxygen molecules, generated by high energy irradiation of Donor beads, transfer to Acceptor beads, which are within a distance of approximately 200 nm. It is a sensitive high throughput screening assay that does not require washing steps. A cascading series of chemical reactions results in a chemiluminescent signal. Purified antibodies were diluted to 100 nM in AlphaLISA® buffer (PBS + 0.05% Tween-20 + 0.5 mg/mL BSA). Subsequently, 5 μL of the IgGs were transferred to an OptiPlate-384 assay plate (white opaque, PerkinElmer, Waltham, MA), mixed with 10μL (10 nM final conc.) of biotinylated peptide probe and 10 uL (10 μg/mL final conc.) of Anti-human IgG (Fc specific; Perkin-Elmer, Waltham, MA) acceptor beads. After an hour of incubation at RT, non-shaking, 25 uL (40 μg/mL final conc.) of streptavidin donor beads (Perkin-Elmer, Waltham, MA) were added. The plate was then incubated for 30 min at RT in the dark before the AlphaLISA signal was detected using a SpectraMax® i3x multi-mode microplate reader (Molecular Devices, San Jose, CA).

Evaluation of Protective Antibody Performance in Mouse Models of Malaria Challenge

Female 6- to 8-weeks old B6(Cg)-Tyrc-2J/J albino and Balb/c mice were obtained from The Jackson Laboratory (Bar Harbor, ME). All animals were maintained and cared for in accordance with the American

Association for Accreditation of Laboratory Animal Care Standards. All mouse procedures were performed according to protocols approved by the Institutional Animal Care and Use Ethics Committees of the National Institute of Allergy and Infectious Diseases (Animal Study Protocol VRC-20-0855).

To generate sporozoites, transgenic *P. berghei* (strain ANKA 676m1C11, MRA-868) expressing full-length *P. falciparum* CSP and a green fluorescent protein/luciferase fusion protein (Pb-PfCSP-GFP/Luc-SPZ or Pb-PfCSP-SPZ) were obtained from salivary glands of infected mosquitoes, as previously described (*201*). Briefly, *Anopheles stephensi* (Nijmegan) mosquitoes were obtained and reared from a colony maintained at the Laboratory of Malaria and Vector Research (NIAID, NIH). Female mosquitoes were allowed to feed on 6- to 8-week-old female Balb/c mice infected with blood-stage Pb-PfCSP- GFP/LUC parasites. After infection, mosquitoes were maintained in an incubator at 19-20°C and supplied with a sterile cotton pad soaked in 10% sucrose, changed every 48 hrs. Eighteen to 21 days following mosquito infections, salivary glands were dissected and ground in 400 µL of L-15 medium (Millipore-Sigma, Burlington, MA), and viable SPZs were counted in a Neubauer chamber.

To assess the protective efficacy of anti-PfCSP mAbs *in vivo*, varying concentrations ranging 50 to 300 µg of PfCSP mAbs were diluted in sterile filtered 1x PBS (pH 7.4; total volume 200 µL/mouse) and administered into the tail veins of female 6- to 8-week old B6(Cg)-Tyrc-2J/J albino mice (The Jackson Laboratory). After 2 hours, mice were intravenously challenged in the tail vein with 2,000 freshly harvested Pb-PfCSPGFP / Luc-SPZ in Leibovitz's L-15 medium (Thermo Fisher Scientific, Waltham, MA, USA).

To assess malaria infection, mice were injected intraperitoneally (i.p.) with 150 µL d-luciferin (30 mg/mL, PerkinElmer, Waltham, MA), anesthetized with isoflurane and imaged with an IVIS® Spectrum *in vivo* imaging system (PerkinElmer) 10 minutes after luciferin injection. Parasite liver load was assessed 40-42 hours after challenge, whereas parasitemia was measured 6 days following challenge. Parasite load was quantified by analyzing a region of interest (ROI) in the upper abdominal region for liver stage, or whole animal for parasitemia; bioluminescence or the total flux (photons/second; p/s) was measured using the manufacturer's software (Living Image 4.5, PerkinElmer).

CIS43 variants were compared to an equivalent dose of CIS43 or L9 as benchmark controls, using the two-tailed Mann-Whitney test. Statistical differences were estimated on the parasite load (liver burden or parasitemia) raw bioluminescence measurements (Appendix K Figure 25) or on the $\log_{10}$(normalized liver burden or parasitemia) values. To enable comparison of parasite infection between different studies, the liver burden (or parasitemia) of each group was normalized based on the geometric mean of the liver burden/parasitemia values from the untreated mice in the same experiment. Data were plotted and graphed using GraphPad Prism, unless otherwise stated. P values less than 0.05 were considered significant (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001) as indicated in the figures.

Isothermal Titration Calorimetry (ITC)

For ITC experiments, a N-terminal domain stabilized version of fl_PfCSP (based on 3D7 clone of the NF54 strain (PlasmoDB ID: PF3D7_0304600.1) with increased expression was used. Four amino acid mutations (as indicated in Appendix P Table 6) were introduced in the N-terminal domain of recombinant fl_PfCSP(*167*) to generate NTDS_fl_CSP (PfCSP_SAmut_C5S, ref. (*169*). These changes remove an enzymatic processing site and help prevent dimerization upon solubilization through the conversion of a cysteine to a serine resulting in increased yields and consistent analyses.

ITC was carried out using a MicroCal VP-ITC microcalorimeter from Malvern Panalytical, (Northampton, MA, USA). Prior to the experiments, both PfCSP and different antibodies were exhaustively dialyzed against PBS, pH 7.4. The antibody solution, prepared at a concentration of ~25 μM (expressed per antigen binding site), was injected in 7 μl aliquots into the calorimetric cell containing PfCSP at a concentration of ~0.3 μM. All titrations were performed at 25 °C. The exact concentrations of the reactants in each experiment were determined from the protein absorbance at 280 nm. The heat evolved upon each injection of antibody was obtained by integration of the calorimetric signal, *dQ/dt*. The heat associated with binding to PfCSP was obtained by subtracting the heat of dilution from the heat of reaction. The separate heats were plotted against the molar ratio, and values for the enthalpy change, ΔH, the association constant, K_a (the dissociation constant, $K_d = 1/K_a$) and the stoichiometry, N, were obtained by nonlinear regression of the data

to a binding polynomial accounting for the binding to two sets of sites with different binding affinities *(202)*. Gibbs energy, ΔG, was calculated from the relation $\Delta G = -RT\ln K_a$, where R is the is the universal gas constant, (1.987 cal/(K × mol)) and T the absolute temperature in kelvin. The entropy contribution to Gibbs energy, $-T\Delta S$, was calculated from the known relation $\Delta G = \Delta H - T\Delta S$. The results were expressed per mole of antigen binding sites and the stoichiometry, N, denotes the number of antigen binding sites per mole of PfCSP.

Crystallization and Structural Analysis

Antibody Fab and peptide 21 (PfCSP residues 101–115) complexes were prepared by mixing 1:2 molar ratio to a concentration of 15 mg/ml. Crystallization conditions were screened in Hampton Research screening kits, Wizard screening kits, Precipitant Synergy screening kits using a mosquito robot. Crystals initially observed from the wells were manually reproduced. The CIS43_Var2 : P21 complex crystal grew in 0.2 M ammonium sulfate and 28% w/v polyethylene glycol 4,000; the CIS43_Var10 : P21 complex crystal grew in 15% isopropanol, 0.2 M ammonium citrate pH 3.5 and 25% w/v polyethylene glycol 3,350; Crystals were cryoprotected in 20% glycerol and flash-frozen in liquid nitrogen. Data were collected at a temperature of 100 K and a wavelength of 1.00 Å at the SER-CAT beamline ID-22 (Advanced Photon Source, Argonne National Laboratory, Lemont, IL). Diffraction data were processed with the HKL2000 suite_ENREF_13. Structure solution was obtained by molecular replacement with Phaser using CIS43 Fab structures (PDB ID: 6B5M) as a search model. Model building was carried out with Coot (*156*). Refinement was carried out with Phenix (*203*). Ramachandran statistical analysis indicated that the final structures contained no disallowed residues or no more than 0.22% disallowed residues. Data collection and refinement statistics are shown in Appendix S Table 9.

Cryo-EM structures of CIS43_Var10 in complex with fl_PfCSP

CIS43_Var10 was mixed at a 1:1 molar stoichiometry with fl_PfCSP and the complex was deposited on a C-flat grid (protochip.com). The grid was vitrified in ethane with an FEI Vitrobot Mark IV set to a wait time of 30 seconds, blot time of 3 seconds and blot force of 1. Automated data collection on a Titan Krios electron microscope was performed with Leginon (*152*) with a Gatan K3 direct detection device. Exposures were collected in movie mode for a 2 s with the total dose of 51.15 e–/Å^2 fractionated over 50 raw frames. Images were pre-processed using Appion (*204, 205*) during collection; individual frames were aligned and dose-weighted using MotionCor2 (*206*) and CTFFind4 (*207, 208*) was used to estimate the CTF. CryoSPARC 2.15 (ref. (*153*)) was used for 2D classifications, ab initio 3D reconstruction, homogeneous refinement, and nonuniform 3D refinement. Initial 3D reconstruction and final refinements were performed using C1 symmetry. Data collection and refinement statistics are shown in Appendix T Table 10.

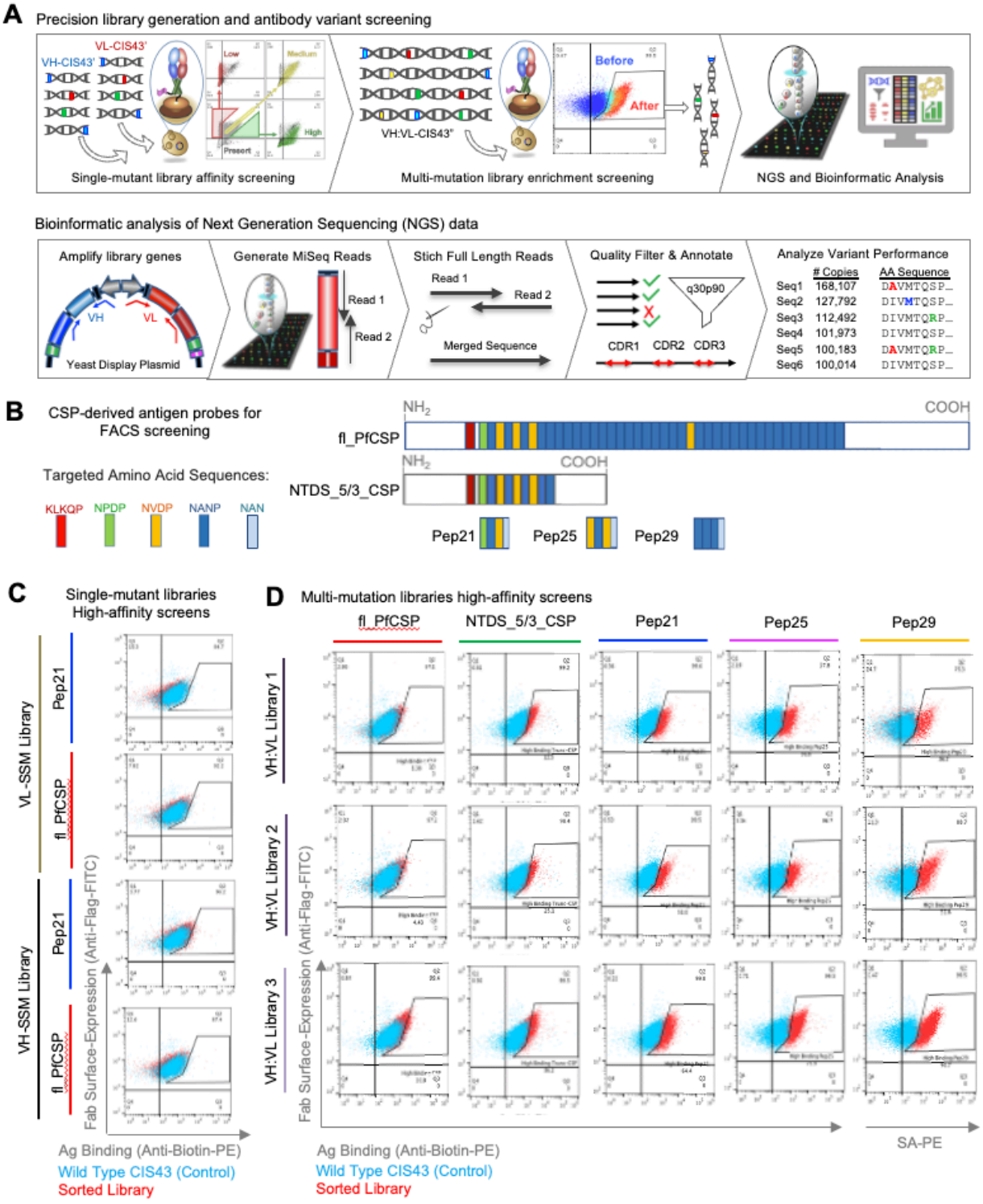


Figure 14. Precise directed evolution techniques were used to optimize the monoclonal antibody CIS43 for enhanced binding to P. falciparum circumsporozoite protein (PfCSP) epitopes.

Fig. 18. Precise directed evolution techniques were used to optimize the monoclonal antibody CIS43 for enhanced binding to *P. falciparum* circumsporozoite protein (PfCSP) epitopes.

(A) Site-saturation mutagenesis (SSM) libraries were generated separately for the heavy chain variable region (VH) and light chain variable region (VL) of the anti-malaria antibody CIS43 and cloned into yeast Fab display vectors for affinity-based FACS. Plasmid DNA was extracted from high-affinity sorted libraries, and additional combinatorial mutagenesis and screening was performed to select high-affinity multi-mutation variants by FACS. Sorted libraries were analyzed using next generation sequencing (NGS) to quantitatively track each variant across sorting rounds. NGS data were mined to identify improved CIS43 antibody gene variants. After Illumina 2x300bp sequencing, raw fastq sequence files were merged and quality-filtered to obtain high-quality in-frame antibody amino acid sequences. Unique amino acid sequences were compiled into library prevalence values, which were followed to track changes in variant populations across each round of sorting.

(B) Five CSP-derived antigen probes were used for yeast display screening. Probes include full length *P. falciparum* circumsporozoite protein (fl_PfCSP), an n-terminal domain stabilized truncated-CSP format (NTDS_5/3_CSP), and three peptides derived from PfCSP (Pep21, Pep25, and Pep29.)

(C) VH-SSM and VL-SSM single mutation libraries were sorted for high affinity against fl_PfCSP and Pep21 (red) compared to wild-type CIS43 (cyan), which revealed minimal enhancements in library-scale affinity in single mutation libraries.

(D) Flow cytometric analysis of combinatorial multi-mutation Libraries 1, 2, and 3 after three rounds of sorting showed substantially enhanced affinity against fl_PfCSP, NTDS_5/3_CSP, Pep21, Pep25, and Pep29 (red) compared to wild-type CIS43 (cyan).

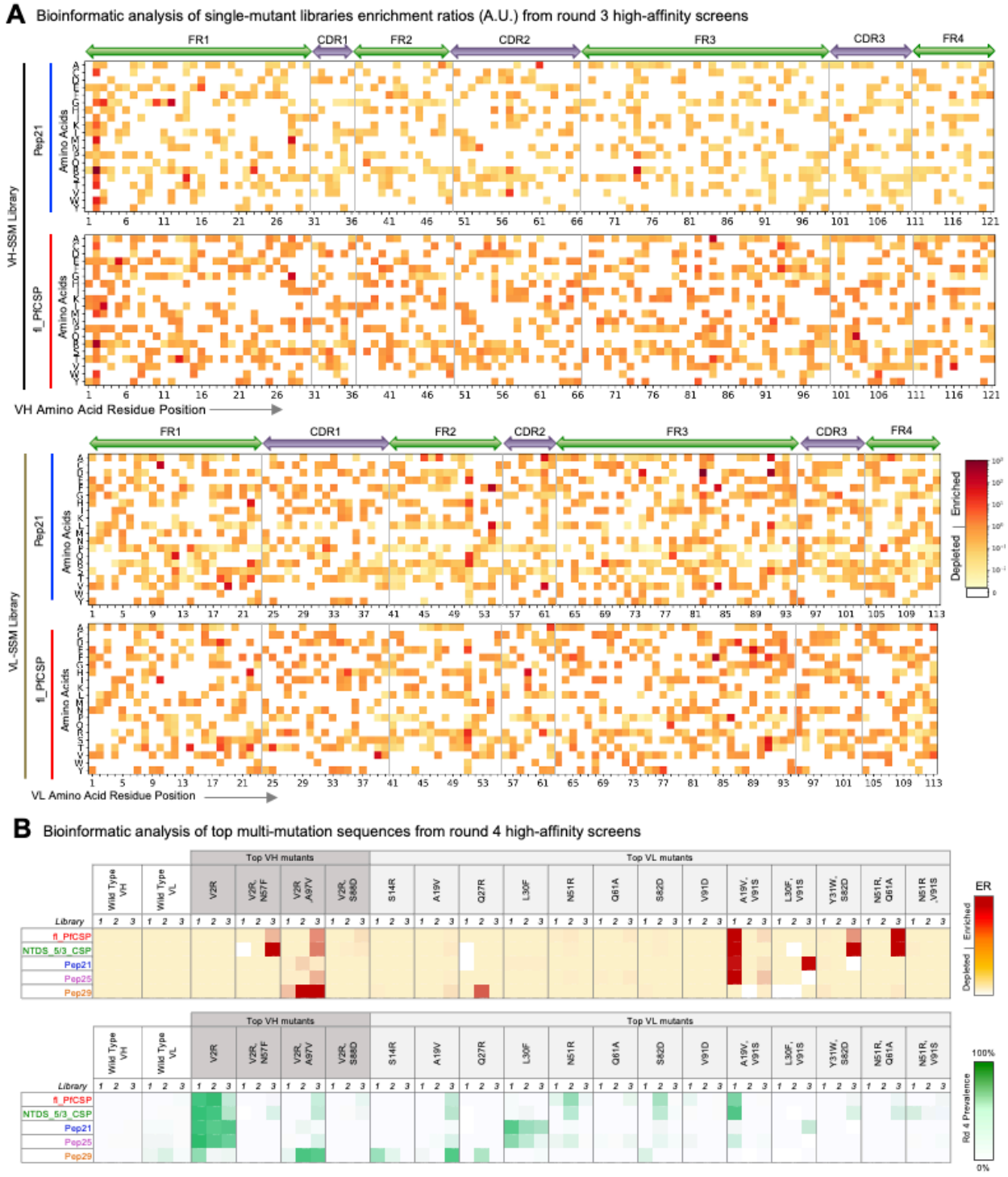


Figure 15. NGS analysis of FACS-screened libraries revealed single- and multi-mutation variants with enhanced PfCSP antigen recognition.

Fig. 19. NGS analysis of FACS-screened libraries revealed single- and multi-mutation variants with enhanced PfCSP antigen recognition.

(A) Heat maps revealed the functional impact of single mutations against fl_PfCSP and Pep21 antigens from high-affinity enriched libraries.

(B) Bioinformatic analyses of multi-mutation sequences enrichment ratios (ER) and prevalence values after four rounds of sorting showed that the dominant enriched mutant sequences out-competed unmutated CIS43 template genes for antigen recognition after screening for high-affinity binding against fl_PfCSP, NTDS_5/3_CSP, Pep21, Pep25, and Pep29 antigens.

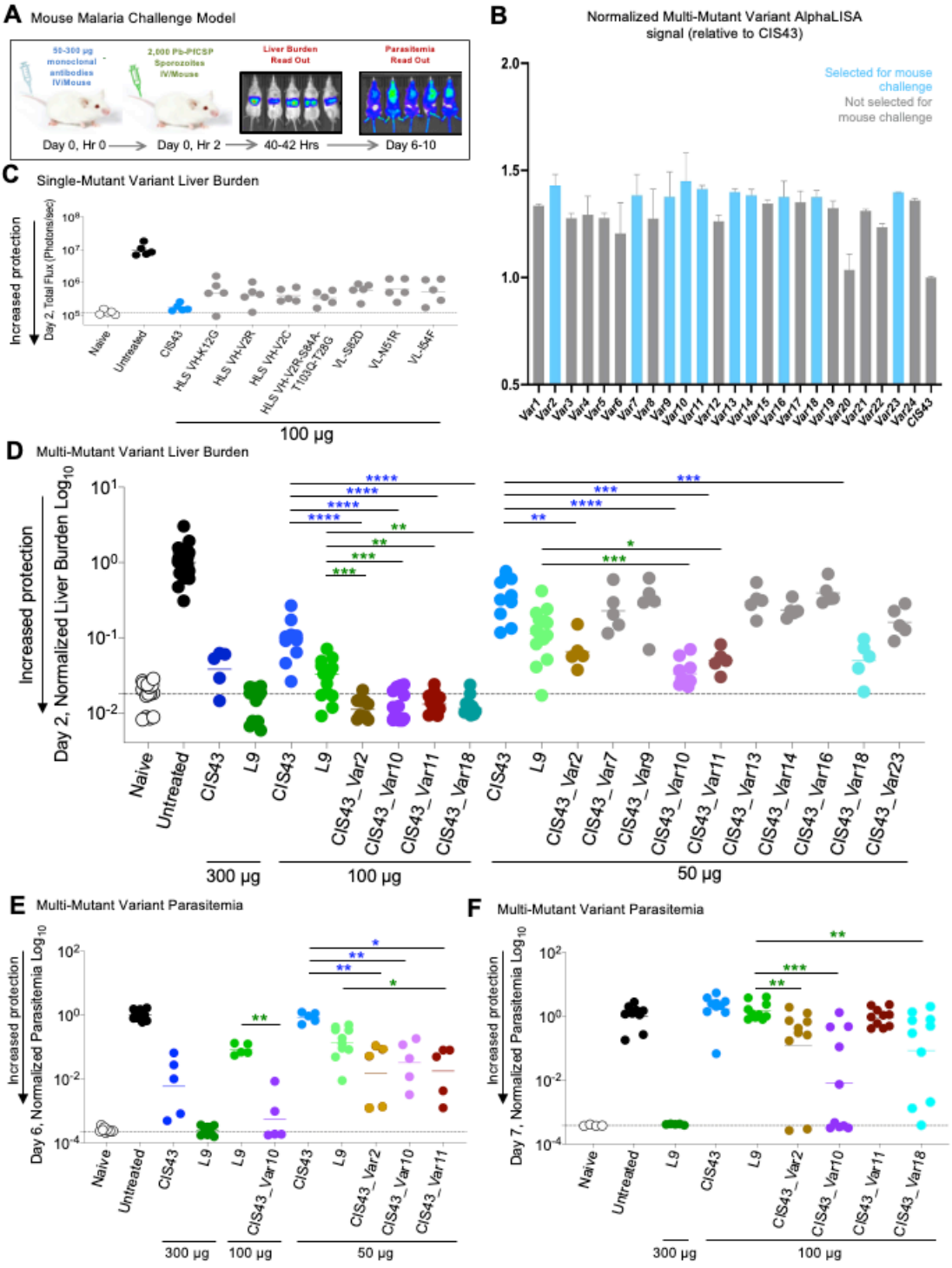


Figure 16. Malaria challenge studies revealed up to 10-fold improved potency for multi-mutation CIS43 variants.

Fig. 20. Malaria challenge studies revealed up to 10-fold improved potency for multi-mutation CIS43 variants.

(A) A mouse malaria challenge model was used to quantify protective potency of anti-PfCSP mAbs. Anti-PfCSP mAbs were passively transferred into mice 2 h prior to infection with transgenic Pb-PfCSP-GFP/Luc, a rodent malaria parasite (*P. berghei, Pb*) with the endogenous PbCSP replaced with PfCSP, and which also constitutively expressed luciferase. Luciferin substrate was injected for bioluminescent *in vivo* imaging by IVIS for an efficient, sensitive and non-invasive means to quantify liver stage (Day 2) and blood stage (Day 6) parasite infection.

(B) Multi-mutation variant antibodies were screened for functional recognition by apparent AlphaLISA affinity to NPDP19. The top 10 selected antibodies for *in vivo* protection analyses are shown in light blue.

(C) Single-mutation variant liver burden on Day 2. Following passive transfer of 100 μg CIS43 (blue) or CIS43 variants (grey), Albino-B6 mice were challenged i.v. with Pb-PfCSP-GFP/LUC SPZ before imaging by IVIS. Total flux indicated the extent of malaria infection in the liver, with naïve (background, white) and untreated (maximum burden, black).

(D) Multi-mutation variant liver burden on Day 2. Following passive transfer of 50 μg CIS43, or the indicated multi-mutation variants over several different experiments, the liver burden of each antibody group was normalized by the geometric mean of liver burden values using untreated mice in the same experiment. Statistical differences were assessed using Mann-Whitney test. P-values ranges are indicated by stars, with P-value < 0.05 (*), P-value < 0.01 (**), P-value < 0.001 (***), and P-value < 0.0001 (****). Raw data from this experiment is presented in **Supplementary Fig. 3**.

(E) Multi-mutation variant parasitemia on Day 6. Following passive transfer of CIS43, L9 and the indicated multi-mutation variants at specified doses, liver burden was assessed at Day 2 (captured in **Fig. 3D**, **Supplementary Fig. 3C** and **Supplementary Fig. 3D**). Mice treated with CIS43_Var2, CIS_Var10 or CIS43_Var11 showed a high degree of protection at Day 2 and were followed and re-imaged again on Day 6 to quantify blood stage infection and/or parasitemia. Statistical differences were assessed using Mann-Whitney test.

(F) Multi-mutation variant parasitemia on Day 7. Following passive transfer of indicated doses of CIS43, L9 and the indicated multi-mutation variants, liver burden was assessed at Day 2 (captured in **Fig. 3D**, or **Supplementary Fig. 3D**). Mice were followed and re-imaged again at Day 7 to assess parasitemia. Statistical differences were assessed using Mann-Whitney test.

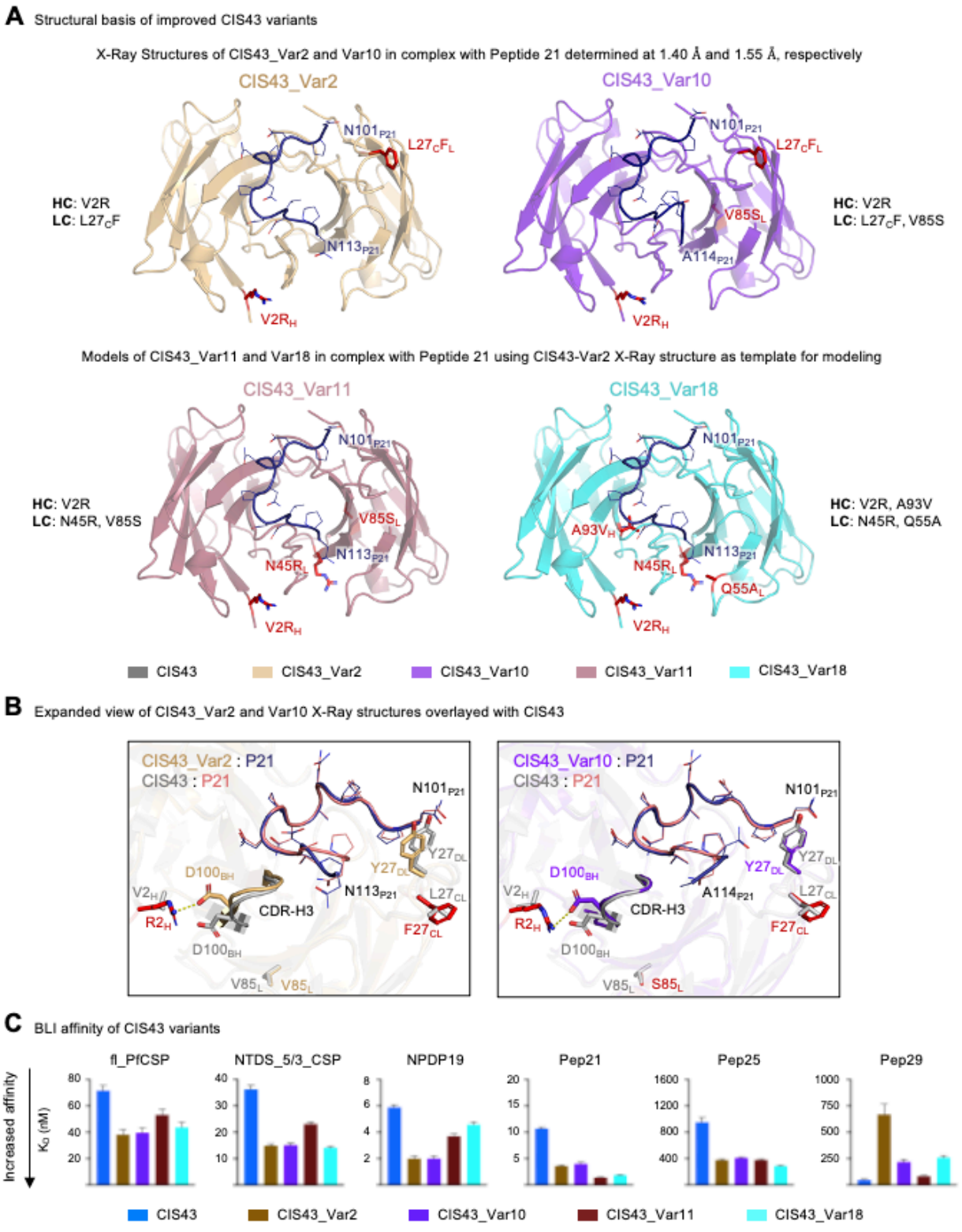


Figure 17. Structural and functional analysis of the mechanisms of improved anti-malaria protection.

Fig. 21. Structural and functional analysis of the mechanisms of improved anti-malaria protection.

(A) X-ray structures of CIS43_Var2 and CIS43_Var10 in complex with the junctional epitope (Peptide 21) determined at 1.33 Å and 1.55 Å resolution. CIS43_Var11 and CIS43_Var18 structures were modeled using CIS43_Var2 as template. The location of the mutations is shown in red. The Pep21 numbering corresponds to PfCSP sequence.

(B) Overlay of CIS43_Var2 with CIS43 showed interactions of Arg2 with $Asp100_B$, resulting in significant movement of CDR-H3 (left); CIS43_Var10 overlayed with CIS43 showed similar interactions between Arg2 and $Asp100_B$, without substantial changes to the CDR-H3 (right).

(C) Biolayer interferometry (BLI) affinity of improved CIS43 variant Fabs was measured against fl_PfCSP, NTDS_5/3_CSP, NPDP19, Pep21, Pep25, and Pep29 antigens. The affinity of multi-mutation variants was improved over CIS43 for all antigens except Pep 29.

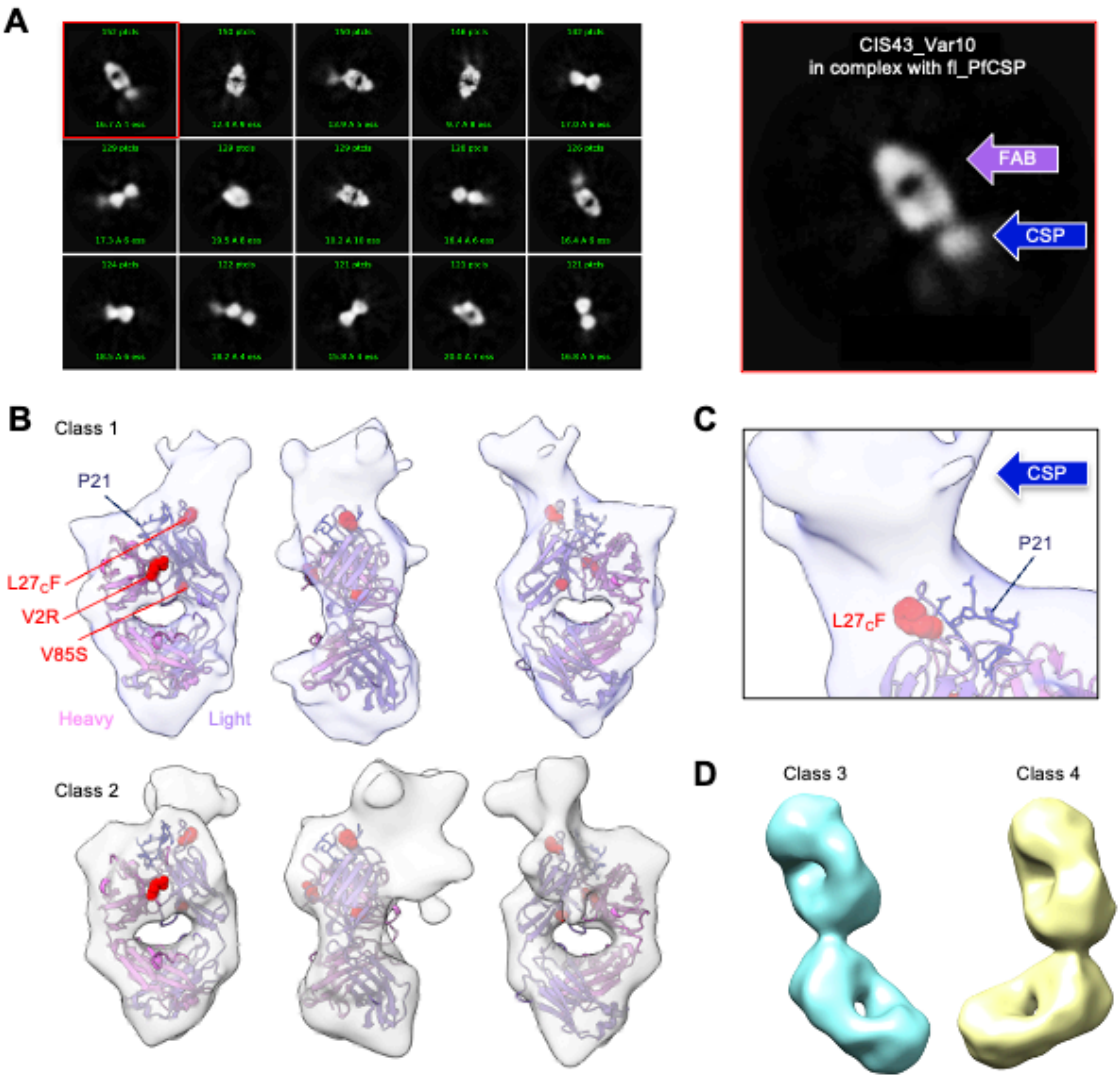


Figure 18. Cryo-EM structures of CIS43_Var10 in complex with CSP.

Fig. 22. Cryo-EM structures of CIS43_Var10 in complex with CSP

(A) 2D class averages are shown for CIS43_Var10 in complex with CSP. The Fab shows additional signal from the CSP protruding from the Fv region. (right) A panel is highlighted showing the two components of the complex.

(B) 3D reconstructions from two classes are shown fit with the crystal structures of CIS43_Var10 bound to peptide 21. The ~6-7 Å resolution maps both show density related to the CSP beyond the bound peptide and encompassing the VL-L27$_C$F mutation on the antibody.

(C) A closer view of the density surrounding the VL-L27$_C$F mutation. The map indicates that regions of CSP beyond peptide 21 extend in the direction of the mutation suggesting additional interactions.

(D) Two classes showed two Fabs linked through binding a single CSP, indicating that two Fabs can bind close to one another in reconstructions of 10-11 Å resolution.

4. Structural mechanism for broad immune recognition of HIV-1 revealed by directed evolution of fusion peptide targeting antibody[3]

4.1 Introduction and Rationale: Antibody Engineering

Antibody Engineering – Evaluate the mutational fitness landscape of protective monoclonal antibody variable regions to screen for mutations enhancing antibody recognition while revealing molecular features correlated to disease protection. Molecular-level understanding of acquired pathogen immunity can be enhanced by studying protective mAbs (*98, 162, 163*). Here we hypothesize that a limited amount of amino acid substitutions can further improve a neutralizing mAbs breadth vs. a diverse target. We test this hypothesis using a protective anti-HIV-1 mAb: VRC34.01 (*209*).

The fusion peptide (FP) of HIV-1 is a promising target for broadly neutralizing antibody vaccines, however FP-diversity among globally circulating HIV-1 strains presents a functional challenge for eliciting pan-neutralizing antibody responses. In this study, we explored how anti-FP antibodies with limited breadth could gain near-pan neutralization by evaluating the mutant landscape of a previously identified FP-neutralizing antibody: VRC34.01. Using site-saturation mutagenesis we generated DNA libraries encoding all possible amino acid substitutions for each residue in the VRC34.01 variable-region template genes and evaluated each mutation's influence on antigen recognition *en masse* using yeast-library-display to screen against HIV-1 trimer probes containing artificially diversified fusion peptide sequences. Successive rounds of combinatorial mutant shuffling and yeast-display screening revealed the most potent variant, VRC34.01_mm28, utilized three mutations to gain 20-fold improved neutralization potency and pan-neutralizing breadth as measured by a diverse 20-virus HIV-1 panel. Structural studies of beneficial mutations revealed optimized antibody paratopes accommodate recognition of the HIV-1 Env backbone

[3] Bailey B. Banach (BBB), *et al*. Structural mechanism for broad immune recognition of HIV-1 revealed by directed evolution of fusion peptide targeting antibody. Designed the experiments: BBB, Tatsiana Bylund, Thuy Duong Nguyen, John R. Mascola, Peter D. Kwong, and Brandon J. DeKosky; Performed the experiments: BBB, Kai Xu, Chen-Hsiang Shen, Baoshan Zhang, Tatsiana Bylund, Thuy Duong Nguyen, Jacy R. Wolfe, Natalie Bui, Gwo-Yu Chuang, Nicole Doria-Rose, Mark Louder, Bob C. Lin, Sijy O'Dell, Reda Rawi, Mallika Sastry, Tracy Liu, Adam Olia, Sergei Pletnev, Arne Schon, Shuishu Wang; Analysed the data: BBB, Ahmed S. Fahad, Bharat Madan, Peter D. Kwong, and Brandon J. DeKosky; Writing: BBB and Brandon J. DeKosky. Reviewing & Editing: all authors. *In preparation for submission.*

and fusion peptide sequences of different lengths. These data define important antibody features for best-in-class anti-FP breadth and potency against circulating HIV-1 strains to accelerate the clinical development of novel FP-targeting vaccines and therapeutics.

Background

A major goal of HIV-1 vaccine development is to elicit broadly neutralizing antibody responses; however several molecular mechanisms constrain immune recognition of the Envelope (Env) trimer and impede *in vivo* development of broadly neutralizing antibodies (bNAbs). Viral mutation rates produce a highly diverse, unstable, fusion-capable, glycan-shielded Env structure capable of evading immune recognition and delaying the onset of bNAb-pre-cursor B-cell activation. Once activated, broad antibody development can be further hindered by natural limitations in the mutations readily sampled during antigen dependent maturation, resulting in sub-optimized immune responses. Prolonged viral exposure has been shown to support increased bNAb development as approximately half of individuals infected with HIV-1 produce bNAbs neutralizing up to ~50% of circulating HIV-1 strains following several years of infection, and highly potent bNAbs have been discovered in cases of chronic HIV-1 infection onset in fetal development (*210, 211*). Preliminary investigations using prefusion-stabilized versions of Env immunogens have elicited neutralizing antibody responses at potencies relevant for protection against natural infection, but these responses were primarily strain- or clade-specific leading to gaps in protection against globally circulating strains. These constraints on natural HIV-1 bNAb development underscore an urgent need to identify critical molecular pathways that can be readily translated into broadly neutralizing antibody vaccine strategies.

Building on insights from well characterized bNAbs, a handful of neutralizing epitopes on the Env trimer have been discovered to induce rare, cross-clade reactive antibodies by targeting immune responses to vulnerable viral machinery. Accumulating evidence for this targeted vaccine approach has revealed unique challenges associated with translating each known vulnerable epitope into effective immunogens. For example, broad and potent recognition of the CD4 binding site requires extensive antibody somatic

hypermutation (*212-214*), whereas the quaternary V1V2 site at the trimer apex requires unusual and rare antibody recombination events (*215-220*). Likewise, neutralization against the glycan-V3 supersite requires complex recognition of a N-linked glycan (*221-224*), and the membrane-proximal external region (MPER) entails close proximity with the viral membrane that may require artificial breaks in immune tolerance for effective antibody elicitation (*225-232*). The recent discovery of bNAb VRC34.01, isolated from a chronically HIV-1-infected donor (*209*), identified the fusion peptide (FP) of the HIV-1 Env gp41 subunit as a specific target for broad immune recognition. A majority of VRC34.01's total interactive surface area (55%) binds to solvent-exposed FP residues 512-519 at the N-terminus of gp41, while 26% of the surface co-recognizes glycan N88 on gp120. With this unique mode of interaction, trimer bound VRC34.01 prevents Env subunits from entering a co-receptor engagement conformation required for viral entry into host cell membranes, leading to 50% neutralization breadth of 208 globally representative HIV-1 isolates (*209*). These biophysical insights defined by VRC34.01 FP-recognition mechanisms sparked new strategies in broad HIV-1 vaccine design; an experimental vaccine strategy "priming" an immune response first using an 8-amino acid FP-based immunogen (FP8-KLH) followed by multiple Env trimer immunogen boosts increased recognition of the vulnerable FP epitope and enhanced vaccine-elicited antibody neutralization breadth (*233*). First generation prime-boost vaccine-elicited FP-mAbs showed modest breadth of up to 31% on a 208 HIV-1 isolate panel (*234*), whereas second generation studies further refined the FP-prime trimer-boost immunization regimen to produce cross-reactive neutralizing antibodies achieving up to 59% breadth on a 208 HIV-1 isolate panel (*235*).

These preliminary studies have demonstrated the potential of FP-specific prime-boost immunization strategies to elicit broad antibody responses on par with natural bNAb development but at a fraction of the timescale (*235*), however, establishing a pan-neutralizing FP-directed epitope vaccine remains an outstanding challenge in the field. Analysis of diverse fusion peptide sequences from 3,942 HIV-1 isolates reveals considerable variation at multiple FP residues (*209*) can enhance neutralization resistance for strains encoding different FP sequences, and prior structural analyses of vaccine-elicited FP-antibodies revealed cross-clade neutralization is hampered by FP conformational diversity (*209, 236, 237*). An enhanced

understanding of antibody development pathways which aggregate FP-recognition for cumulative protection against circulating HIV-1 strains is urgently needed to translate rational FP-vaccine design toward clinical practice.

While natural immune responses are often limited by the sampling of rare mutations and inter-clonal competition dynamics in secondary lymphoid organs, synthetic mutational screening and directed evolution techniques offer *in vitro* alternatives to identify new neutralizing mechanisms that have not been observed to date. Prior investigations by our group and others demonstrated that functional improvements in HIV-1 neutralization can be achieved by precision site-saturation mutagenesis (SSM) and precision yeast display screening of antibody gene variant libraires to identify synergistic combinations of beneficial mutations that increase antibody potency and/or breadth (*171, 175, 238-240*). Based on these promising data and the urgent need to improve the neutralization breadth of anti-FP vaccines and antibodies, we applied a comprehensive *in vitro* antibody gene variant screening approach to reveal additional molecular mechanisms accumulating in broad anti-FP recognition.

In this study, we hypothesized that enhanced antibody recognition of neutralization-resistant FP residues would lead to improvements in protective breadth. We experimentally sampled FP-specific antibody maturation pathways for VRC34.01 using precise directed evolution techniques to determine genetic, structural, and biophysical antibody features associated with potent and broad HIV-1 FP-neutralization. Analysis of mutant antibody libraries screened *en masse* at high-throughput using yeast display defined three critical antibody mutations that enhanced FP recognition dynamics to increase breadth from ~60% (template antibody) to 100% while enhancing potency 20-fold on a globally representative 20-virs panel. Structural analyses revealed key amino acid mutations overcame deficiencies in antibody breadth by adapting paratope conformations accommodating FP sequences of different lengths to enhance neutralization function. These efforts define broad neutralization features for the first pan-neutralizing antibody targeting the HIV-1 fusion peptide epitope, and can be translated into new strategies for FP-focused broadly neutralizing HIV-1 vaccines.

.

4.2 Results

Precise directed evolution of VRC34.01 enhances antibody recognition of diverse fusion peptides presented by HIV-1 SOSIP trimers

We engineered a suite of broadly neutralizing anti-FP HIV-1 antibodies by precisely identifying affinity-enhancing mutations in the VRC34.01 variable region that could lead to improved HIV-1 neutralization potency and breadth (**Fig. 29**). Beginning with the VRC34.01 antibody template variable region genes, site-saturation mutagenesis was used to generate single-mutant DNA libraries comprising all possible 7,328 single amino acid substitutions across the VRC34.01 variable heavy and variable light chains (**Appendix U Figure 33** and **Appendix Y Table 11**) (*81, 196, 209*). Single-mutant DNA libraries were cloned into a yeast-surface-display plasmid vector with a galactose-induced bi-directional promoter, leucine zipper, and FLAG tag to express, assemble, and monitor Fab libraries via fluorescence activated cell sorting (FACS) (**Fig. 29A**) (*20, 78*). We hypothesized that engineering for enhanced recognition of diversified FP residues 512-519 presented in context of the full Env trimer could improve antibody breadth. Accordingly, Fab-expressing yeast libraries were stained with BG505.SOSIP gp160-trimers containing diverse FP sequences: FP- v1 (AVGIGAVF), Thai (AVGIGAMI), and v3 (AIGLGAMF) (**Fig. 29B**). Single mutant libraries of the heavy and light chains (VH-SSM and VL-SSM, respectively) were separately enriched over three rounds of FACS to fractionate libraries into variants with high-, medium-, and low- binding affinity phenotypes (**Fig. 29B**). Sorted libraries were analyzed with next generation sequencing (NGS) to reveal the composition of pre- and post-sort libraries and bioinformatically track enriched mutant sequences (**Fig. 30A** and **Appendix U Figure 33**) (*78*). Several key residue substitutions with high enrichment and selectivity for multiple FP antigens were identified in NGS data and selected for follow-up performance analysis as soluble IgG (**Fig. 30A)**.

Single mutation variants increased antibody neutralization breadth and potency

A total of 47 single mutant variants were selected for expression and characterization as IgG protein (**Appendix Z Table 12**). First, a limited 8-virus neutralization panel was designed as a first-pass screen to identify the most promising single mutation candidates (Supplemental Fig. 2A). The 8-virus panel was selected so that neutralization breadth would correlate with breadth on a larger 208-virus panel by using a 10^7 randomized search for an 8-virus sub-group within the large panel that maximized the correlation between select control anti-FP antibody's neutralization outcomes (**Appendix V Figure 34A** and **Appendix AA Table 13**). Functional evaluation of the 47-single mutant variants on the predictive first-pass 8-virus panel revealed that most mutations had no effect or detrimental effect on neutralization, however, eight of the single-mutant variants showed important potency improvements. One variant (VH_E2P) showed enhanced HIV-1 neutralization potency and breadth (**Appendix V Figure 34B** and **Appendix AB Table 14**). Next, we designed a 20-virus neutralization panel for maximal correlation with FP antibody neutralization breadth on the 208-virus panel, using the same method as in the previous 8-virus panel design, for analysis of the top single- and rational combination-mutant variants (**Fig. 30B, Appendix V Figure 34C** and **Appendices Z, AC, and AD Supplemental Tables 12, 15, and 16**; VRC34.01_combo3 and VRC34.01_combo4 did not express). These neutralization data demonstrated that single mutations to VRC34.01 could increase neutralization breadth from 60% (template mAb) up to 75%, whereas rational combinations of the top single mutations resulted in up to 85% breadth and ~10-fold increase in overall potency for the most improved combination variant, VRC34.01_combo1 (VH_E2K_T59F) (**Fig. 30C**).

VRC34.01_combo1 showed enhanced performance, but with several gaps in neutralization of circulating HIV-1 strains

VRC34.01_combo1 was next evaluated on a 208-virus panel to characterize its neutralization breadth against a globally represented panel of fusion peptide diversity (**Appendix AE Table 17**). VRC34.01-combo1 comprises two heavy chain substitutions identified from yeast display (VH_E2K_T59F), and showed an improvement in breadth from 52% up to 70% on the total 208-virus panel (**Fig. 30D**). However, the enhanced breadth was not evenly distributed amongst HIV-1 variants when

grouped by FP sequence, as shown in figure 2D. We found that the VRC34.01_combo1 neutralized FP_v1 (AVGIGAVF) very effectively, like template VRC34.01 (**Appendix W Figure 35C**) whereas VRC34.01_combo1 showed substantially improved neutralization breadth against variants encoding the FP_Thai (AVGIGAMI), and FP_v3 (AIGLGAMF) sequences screened in FACS analysis (**Figs. 29** and **30**). VRC34.01 also acquired new neutralization gain of function against some FP variants not screened for such as FP_v4 (AVGTIGAMF) and other rare FP-variants (**Fig. 30D**). However, despite starting from the 0% of FP_v4 strains neutralized by VRC34.01, the improved VRC34.01_combo1 neutralized a low fraction of FP_v4 strains (11%), and we hypothesized that a limited ability to recognize the 1 amino acid insertion in FP_v4 (a 9-mer, compared to most other FP sequences being 8-mers) may be limiting the neutralization capacity that could be achieved against strains carrying longer fusion peptide sequences.

Multi-mutation variants selected against diverse FP sequences of different lengths can achieve near pan-neutralization breadth

To address the gaps that we observed in neutralization function against FP_v4 and other diverse HIV-1 variants, we next undertook subsequent rounds of screening against multiple HIV-1 BG505 SOSIP trimers with different FP seqs, specifically including FP_v4 to help address the difficulty of VRC34.01_combo1 had shown with recognizing HIV-1 variants that utilized FP_v4. To generate library diversity, we performed DNA shuffling to combine the most promising single mutations identified from our earlier yeast screens, and we also repeated SSM on the shuffled genes to introduce and screen additional potentially synergistic multi-mutation combinations (**Appendix Z Table 12**). In total, five multi-mutation libraries were generated and screened against four antigen probes: FP- v1 (AVGIGAVF), Thai (AVGIGAMI), v3 (AIGLGAMF), and v4 (AVGTIGAMF) to address gaps in anti-FP neutralization function (**Fig. 31B** and **Appendix Z Table 12**). Multi-mutation libraries were enriched for three rounds of FACS, and potentially beneficial combinations of mutations were bioinformatically identified by NGS of heavy and light chain genes (**Fig. 31C**). 28 rational multi-mutant combinations were inferred from NGS multi-mutation performance data (**Appendix Z Table 12**). Soluble IgG expression and characterization of

these multi-mutant variants compared with unmutated VRC34.01 template on the 20-virus panel revealed that VRC34.01 neutralized up to 75% of variants tested, whereas 10 of the 28 multi-mutant variants showed a neutralization breadth of 90% or more. The most improved multi-mutation variant, VRC34.01_mm28 (VH_E2K_A33P_T59F), neutralized 100% of the 20 viruses in the panel, with a nearly 10-fold enhanced IC50 geomean neutralization potency compared to the template antibody (**Fig. 31D**). As VRC34.01_mm28 was now identified to be the first near pan-neutralizing antibody against the HIV-1 FP, we next sought to characterize its structural interactions and define the mechanisms by which VRC34.01_mm28 achieved its remarkable potency and neutralization breadth.

Enhancing mutations influence both FP recognition interactions and interactions with gp41 to achieve superior FP-based neutralization

Structural analysis of VRC34.01_combo1 and VRC34.01_mm28 revealed three synergistic interactions that enabled enhanced antibody recognition through accommodation of common insertion mutations in the fusion peptide (**Fig. 32**). Super-imposed crystal structures of VRC34.01-combo1 Fab and VRC34.01 Fab in complex with FP shows how the T58F substitution enhances inter-chain hydrophobic interactions to VH_W50 and VL_Y94, while also removing unfavorable interactions between the T58 hydroxyl group and I515 on the Fusion Peptide (**Fig. 32A**). Cryo-EM mapping of the HIV-1 trimer in complex with VRC34.01-combo1 Fab appears similar in structure to the interactions previously observed for the VRC34.01 template antibody in complex with HIV-1 trimer (**Fig. 32B**). The VRC34.01-combo1 mutations introduced to template antibody also support electrostatic interactions between VRC34.01-combo1 and HIV-1 trimer to increase affinity to fusion peptide; the E2K substitution appears to remove the electrostatic conflict with HIV-1 E87, while the T59F substitution on VRC34.01-combo1 alters the position of VH_Y94, moving it close to the carbonyl of FP residue G514 (**Fig. 32C-E**). Comparison of epitope topography between template VRC34.01 and a previously characterized germline intermediate: VRC34 I3, reveal how the A33P substitution widens the hydrophobic binding groove in the antibody epitope to accommodate broad recognition of fusion peptide sequences with insertions, such as FP_v4

(AVGTIGAMF). These antibody-based insights provide a molecular pathway to inform the design of FP-targeted vaccines that can elicit potent neutralizing anti-HIV-1 antibody responses against the breadth of circulating FP variants, and can be further translated into potent, broadly neutralizing anti-FP drug design and development.

4.3 Discussion

We designed a new best-in-class anti-FP neutralizing antibody: VRC34.01_mm28, by applying precision antibody engineering of the FP-targeting antibody VRC34.01. Our directed evolution inspired strategy of site-saturated mutagenesis antibody libraries and combinatorial DNA shuffling, paired with yeast display selection, was able to identify rare combinations of mutations in antibody framework and CDR regions to select for synergistic improvements to anti-HIV-1 antibody performance. The most improved variant, VRC34.01_mm28, achieved 100% neutralization breadth on a 20-isolate panel designed to be predictive of FP-based antibody neutralization of circulating HIV-1 strains. We defined specific antibody:antigen structural interactions for improved FP-based recognition of HIV-1, and these engineered bNAbs demonstrate the potential for truly broad protection against HIV-1 through targeting the FP epitope.

Prior structures of anti-FP-antibodies demonstrated that viral mutations in the fusion peptide and distal interacting sites of the Env trimer can mediate viral escape from natural- and vaccine-elicited antibodies (*241*). In addition to the diversity of FP and viral mutation escape mechanisms, the statistical biases in mutational accumulation during *in vivo* antibody development processes prevent comprehensive sampling of potentially beneficial rare mutations in the antibody mutational-fitness landscape (*26, 78, 209, 211, 212, 214, 219, 224, 233-235, 241-243*). Mutations that require multiple nucleotide substitutions per amino acid codon may not be readily sampled during *in vivo* antibody maturation. In contrast, our *in vitro* antibody engineering efficiently demonstrated that improvements in protective breadth can be achieved with combinations of a very few, but infrequently sampled, sets of mutations to the antibody sequence. Our use of site saturation mutagenesis and combinatorial shuffling with yeast display screening was shown to substantially enhance antibody performance, even for human bNAbs like VRC34.01 elicited from long term, chronic infection and natural exposure to viral diversity.

Structural analysis of the most improved variants enhanced our understanding of the molecular interactions required to achieve broad fusion-peptide recognition. Cryo-EM and crystal structure analyses of VRC34.01_combo1 and VRC34.01_mm28 bound to FP and Env revealed broad antibody recognition correlated with the ability to bind diverse FP variants (*244*). The three mutations comprising the most

improved variant: VH_E2K_A33P_T59F in VRC34.01_mm28, work synergistically for enhanced neutralization, each providing a benefit to a different area of anti-HIV-1 trimer recognition. The VH substitutions E2K and T59F both influence electrostatic interactions to increase favorable antibody interactions with the fusion peptide, whereas A33P opens up the binding groove that interacts with the mid-region of FP, where a large amount of sequence diversity is observed in circulating strains. A quarter of the interactive surface area on template VRC34.01 recognizes glycan N88 on gp120 by a hollow formed by CDRs L1, L2, and H3, whereas the VRC34.01_mm28 paratope relies primarily on the heavy chain for over 75% of HIV-1 interactions (*209*).

Our efforts complement other recent data in the field that indicate the high promise of targeting FP for HIV-1 vaccine designs. The linear fusion peptide target sequence is linked to critical viral entry function, common among Class I viral fusion proteins, and conserved FP features can be recognized by antibodies elicited from both natural infection and vaccination indicating promise for targeted immunogen design (*24, 211, 215, 217, 219, 241, 245*). The N-terminus of the fusion peptide on HIV-1 Env is accessible when exposed in the closed, prefusion state, and several of these hydrophobic residues can enable high-affinity binding interactions with the antibody paratope. FP-directed vaccines have shown increasing promise in multiple animal models (*234, 235*), yet still present several challenges for broad HIV-1 neutralization. Importantly, the data presented here provide the molecular blueprint for very broad and potent anti-FP targeting, and we anticipate these broadly neutralizing antibody-based insights will guide FP-based vaccine design. These findings support immunization strategies using both 8-mer and 9-mer FP variants, to elicit antibodies that can recognize more diverse FPs. Targeted elicitation of a single broad and potent antibody lineage resulting in humoral immunity against both FP variant lengths, similar to the A33P mutation function described here, would be an ideal vaccine strategy. However, the elicitation of separate sets of antibodies targeting 8-mer and 9-mer FPs could be plausible through multiple antibody maturation pathways, as shown by the directed evolution analyses here, and may also be sufficient to achieve high neutralization breadth and potency via vaccination.

In recent decades, an enhanced understanding of effective bNAb development pathways to achieve broad neutralization has fueled rational vaccine design and therapeutic drug development against HIV-1. The recombinantly engineered bNAb VRC34.01_mm28 presented here likely increases HIV-1 neutralization breadth from 50% up to potentially 90% on a 208-isolate panel, which is the first anti-FP mAb reported with a neutralization breadth similar to traditionally broader recognition epitopes like the CD4 binding site (*242, 245-247*), lending further promise to the FP epitope for targeted vaccine design, The template version of VRC34.01 was recently demonstrated to provide passive protection against SHIV exposure when administered as soluble IgG in non-human primates (*248*), and further potency improvements to pan-neutralizing mAbs like VRC34.01_mm28 will support translational efforts developing and administering prophylactic passive vaccines to prevent HIV-1 transmission (*249, 250*). Our results have outlined the molecular features of truly broad HIV-1 neutralization targeting the FP epitope, and we expect that future efforts will build on these data to establish highly potent and near-universally protective FP antibodies and vaccines to suppress the ongoing HIV-1 pandemic.

4.4 Methods

SSM Library Generation and Cloning into Yeast Display

Single-mutation plasmid libraries were constructed using a one-pot site saturation mutagenesis (SSM) protocol (*81*). VRC34.01 heavy chain variable region (VH) and light chain variable region (VL) template gene were mutated separately using mutagenic primers containing degenerate single codons to express all 20 amino acids at each residue of the VH and VL. A process of single-strand nicking, exonuclease digestion, and degenerate oligonucleotide-primed PCR was used to generate each library (VH and VL) containing the comprehensive set of single-residue substitutions across the VRC34.01 variable regions (*196*). The separate (VH-SSM):VL and VH:(VL-SSM) libraries were cloned into a yeast display vector containing a FLAG-marker to quantify Fab surface expression, as well as a leucine zipper and protein disulfide isomerase expression for enhanced expression of diverse antibody libraries (*20*). Plasmid DNA libraries were used to transform yeast populations, and library sizes of at least 3×10^6 were maintained in all cloning steps, as described previously (*78*).

Fluorescence-Activated Cell Sorting (FACS) of Single-Mutation Libraries

Transformed yeast libraries were cultured in SGDCAA medium to induce antibody fragment antigen binding (Fab) surface expression, (20 g/L galactose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4.H_2O$; SGCAA from TEKnova, Hollister, CA) supplemented with 2 g/L dextrose (SGDCAA) for 36 h at 20°C and 225 rpm. Yeast libraries were washed and stained with an anti-FLAG FITC monoclonal to quantify Fab expression (F4049, Clone M2, Sigma-Aldrich, Burlington, MA). Trimer antigen probes were generated by fluorescently labeling biotinylated constructs with an anti-biotin-PE label (12-9895-82, clone BK-1/39, Thermo Scientific, Waltham, MA). Anti-FLAG-FITC-labeled libraries were co-stained with fluorescently conjugated antigen probes to screen single mutation Fab libraries for antigen-recognition. In the first round of FACS, 3e7 yeast cells were stained and sorted using three gates to sort and collect low, medium, and high affinity yeast library groups by the ratio of Fab surface-expression versus Ag binding for the cell population, as described previously (*77, 78*). For all sorting

experiments, control libraries of FITC+ yeast were also collected to track the identity and initial prevalence of library variants with successful Fab expression. Collected cells were grown in low-pH SDCAA (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 10.4 g/L trisodium citrate, and 7.4 g/L citric acid monohydrate, pH 4.5) for 24-48 h in a 30°C incubator shaking at 225 rpm. After sorted yeast collection and culture, each sort was repeated (using either low, medium, or high affinity, to match the initial Round 1 sort) for an additional two rounds.

Multi-Mutation Library Design and Screening

Plasmid libraries were isolated from sorted yeast cells using previously described DNA extraction methods (*196*). VH and VL genes were amplified from library plasmids using Kapa Hifi HotStart ReadyMix (Kapa Biosystems, Roche. Wilmington, MA) (*20*). Five VH:VL paired combinatorial VRC34.01 multi-mutation libraries were designed using template DNA from the enriched single-mutation high-affinity screens. Library 1 ($VH_{Shuffled}$:VL) and Library 2 (VH:$VL_{Shuffled}$) were generated by shuffling enriched single-mutations on VH and VL genes, respectively, using previously defined protocols (*79*). Template DNA was fragmented with DNAseI, followed by homologous reassembly and reamplification of the shuffled genetic material_ENREF_7. Library 3 (pooled-VH:VL) was generated by pooling enriched single-mutation VH and VL gene libraries and combining inter-chain mutations via restriction enzyme digest and sub-cloning into a shared expression vector. Library 4 (shuffled-pooled-VH:VL) was created by performing DNA shuffling on Library 3, and Library 5 (re-SSM-shuffled-pooled-VH:VL) was generated by repeating another round of SSM on Library 4.

Yeast libraries were next generated using multi-mutation plasmids and transformation, as described above for the single mutation library screens, and library sizes exceeding 1×10^7 were maintained in all cloning steps. Expression was induced and yeast were stained as above using an anti-FLAG FITC monoclonal antibody (F4049, Clone M2, Sigma-Aldrich. Burlington, MA) and a fluorescently labeled antigen probe. antigen probes were fluorescently labeled with an anti-biotin-PE label (12-9895-82, clone

BK-1/39, Thermo Scientific. Waltham, MA). For screening, libraries were stained with 0.3 nM FP-v1, 0.7 nM FP-Thai, 0.7 nM FP-v3, and 7 nm FP-v4 individually and enriched consistently against the same antigen across multiple rounds. Multi-mutation libraries were enriched for high-affinity binding as described above for single mutation library screening; 3×10^7 yeast cells were stained and the 0.1% highest affinity variants in the library were collected, as determined by the highest ratio of antigen binding to Fab expression (*77, 78*). Control libraries of FITC+ yeast from the pre-sort libraries were also collected to track the identity and initial prevalence of multi-mutation variants with successful Fab expression. Multi-mutation libraries were similarly enriched for high-affinity antigen binders for an additional three rounds using FACS. Collected cells were cultured in low-pH SDCAA (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 10.4 g/L trisodium citrate, and 7.4 g/L citric acid monohydrate, pH 4.5) for 24-48 h at 30 °C and 225 rpm.

NGS and Bioinformatic Analysis of Sorted Antibody Libraries

VH and VL genes were extracted and amplified from sorted yeast libraries and submitted for Next Generation Sequencing (NGS) on an 2x300bp Illumina MiSeq platform. First, plasmids were isolated from cultured yeast cells as described previously (*196*). VH and VL genes were amplified from library plasmids using Kapa Hifi HotStart ReadyMix (Kapa Biosystems, Roche. Wilmington, MA) as described previously (*20, 78, 197*). Amplified libraries were prepared for NGS with an addition round of PCR to incorporate barcodes and adapters for Illumina sequencing.

Raw Illumina fastq sequence reads were processed as described previously (*78, 197, 198*). NGS reads were quality-filtered for a score of 30 over 90% of the raw reads using Fastxtoolkit (v0.0.14 http://hannonlab.cshl.edu/fastx_toolkit/). Filtered reads were then processed using IgBlast software to reference the IMGT database and determine complete variable region gene alignments (*148, 199*). Once variable region alignments were obtained, mutant VRC34.01 sequences were aligned to the template VRC34.01 antibody sequence using Usearch software (*251*). Mutations were defined by determining the percent identity match to the template CIS43 gene and denoting the substituted amino acid residue according to the template CIS43 gene sequence, with Base 1 corresponding to the start of the variable

framework region. Kabat numbering identifiers are also listed for key mutations and referred to in structural data (**Appendix Z Table 12**). The number of reads for unique sequences were enumerated in each library, and used to determine the prevalence, or frequency, of each variant in the sorted libraries:

$$Prevalence_{\,Variant\ X\ in\ Library\ Y} = \frac{Number\ of\ Reads\ of\ Sequence\ X\ in\ Library\ Y}{Total\ number\ of\ Reads\ in\ Library\ Y}$$

We defined a variant's enrichment ratio (ER) as the change in sequence prevalence from the initial Fab expressing (VL-FITC+) sorted control library to the experimentally screened library:

$$ER_{\,Sequence\ X\ in\ Library\ Y} = \frac{Prevalence\ of\ Sequence\ X\ in\ Library\ Y}{Prevalence\ of\ Sequence\ X\ in\ Fab\ Expressing\ Control\ Library}$$

Each single mutation variant was binned into high, medium, and low affinity screened groups by comparing prevalence and enrichment ratio values across different screening conditions, as we reported previously (*77, 78*). Multi-mutation variants were analyzed by their ER in high-affinity sorted library screens, using the multi-mutation VL-FITC+ prevalence in the denominator.

Antibody Expression

Antibody variable heavy chain and light chain sequences were codon optimized, synthesized and cloned into a VRC8400 (CMV/R expression vector)-based IgG1 vector as previously described (*200*). The variants were expressed by transient transfection in Expi293 cells (ThermoFisher Scientific, Waltham, MA) using Turbo293 transfection reagent (SPEED BioSystems, Gaithersburg, MD) according to the manufacturer's recommendation. 50 μg plasmid encoding heavy-chain and 50 μg plasmid encoding light-chain variant genes were mixed with the transfection reagents, added to 100 ml of cells at 2.5×10^6/ml, and incubated in a shaker incubator at 120 rpm, 37°C, 9% CO2. At 5 days post-transfection, cell culture supernatant was harvested and purified with a Protein A (GE Healthcare, Chicago, IL) column. The antibody was eluted using IgG Elution Buffer (ThermoFisher Scientific, Waltham, MA) and were brought to neutral pH with 1 M Tris-HCl, pH 8.0. Eluted antibodies were dialyzed against PBS overnight and were confirmed by SDS-PAGE before use.

Virus Neutralization Assays

8- and 20-virus panels

Monoclonal antibodies were assessed one on one against multiple viral strains using entry neutralization assays as previously described neutralization (*78, 234*). Five-folding serial dilutions of antibodies starting at 500 μg/mL were mixed in 50 μL volumes with stocks of viruses carrying fluorescent luciferase reporter genes. Mixtures were incubated at 37°C for 1 h, followed by the addition of 20 μL TZM-bI cells ($0.5x10^6$ cells/mL) and incubation overnight at 37°C. After 24hrs (day 2) an additional 130 μL of complete Dulbecco's modified Eagle medium was added to the neutralization test reactions and incubated at 37°C overnight. On day 3 cells were lysed and assessed for luciferase activity indicative of viral infection by measuring in relative light units. The concentration of antibody required to inhibit 50 and 80% of virus infection as determined by comparing relative light units between samples to negative non-neutralized controls were determined using a hill-slope regression analysis as described (*78, 234*).

208-virus panels

To model monoclonal antibody function against globally circulating fusion peptide diversity an automated large-batch neutralization panel of 208 HIV-1 Env-psequdotyped viruses was performed using 384-well microneutralization assays as described previously (*252*).

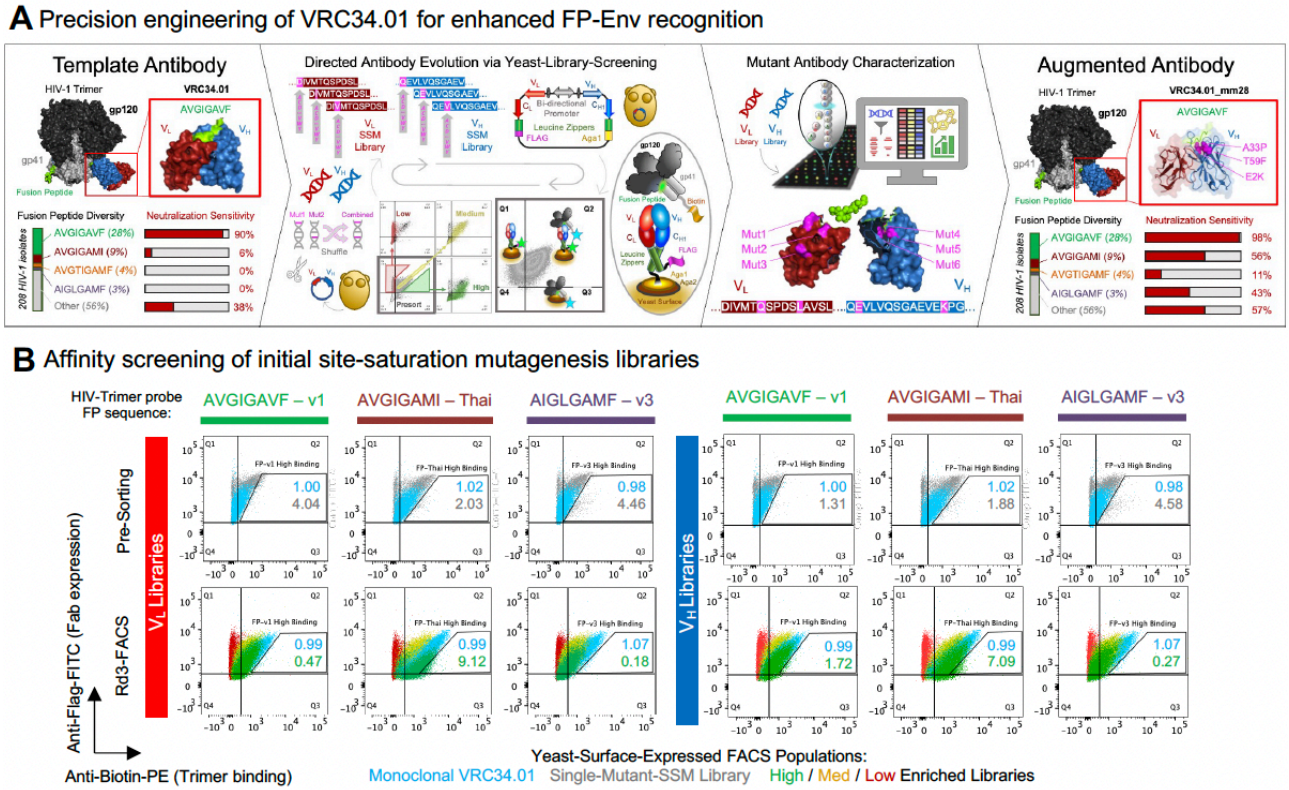


Figure 19. Precision library generation and yeast display screening were applied to the template antibody VRC34.01 to enhance recognition of HIV-1 Envelope (Env)-displayed Fusion Peptide (FP).

(A) Illustrative workflow of precision engineering for antibody performance. VRC34.01 antibody function was engineered for enhanced recognition of diverse HIV-1 FP sequences using precise directed evolution techniques. Template antibody genes were synthetically evolved for directed FP recognition via successive rounds of site-saturated mutagenesis (SSM) and DNA shuffling. Genetic libraries containing mutant antibody sequences were functionally assayed using yeast display and FACS technologies to fractionate mutant variants by antigen-binding affinity phenotypes. Next generation sequencing (NGS) was used to evaluate sorted rounds and construct bioinformatically-inferred molecular maps predicting trimer binding function for individually enriched mutant-antibody sequences. Biophysical characterization of precision engineered single- and multi-

mutation antibody variants revealed important anti-FP antibody sequence-structure-function relationships defining potent gain-of-function mechanisms enabling near-pan HIV-1 neutralization.

(B) Yeast libraries expressing surface-bound antibody fragments (Fab) visualized as fluorescent activated cell sorting (FACS) populations stained with fluorescence markers to measure Fab-surface expression (Y-axis) versus Fab-antigen binding (X-axis). Data shown represents single-mutant amino-acid substitution libraries generated via SSM across VRC34.01 variable light (VL) and variable heavy (VH) genes bound to diverse fusion peptide sequences displayed on HIV-1 trimer probes (Top, Pre-Sorting graphs, gray population represent antigen-stained single-mutant libraries with yeast-displayed antigen-stained monoclonal VRC34.01 shown superimposed in blue). Single-mutant libraries were fractionated and thrice enriched into high- (green), medium- (yellow), and low-(red) affinity performance bins using FACS (Bottom, Rd3-FACS) resulting in sorted libraries with phenotypic differences in trimer binding.

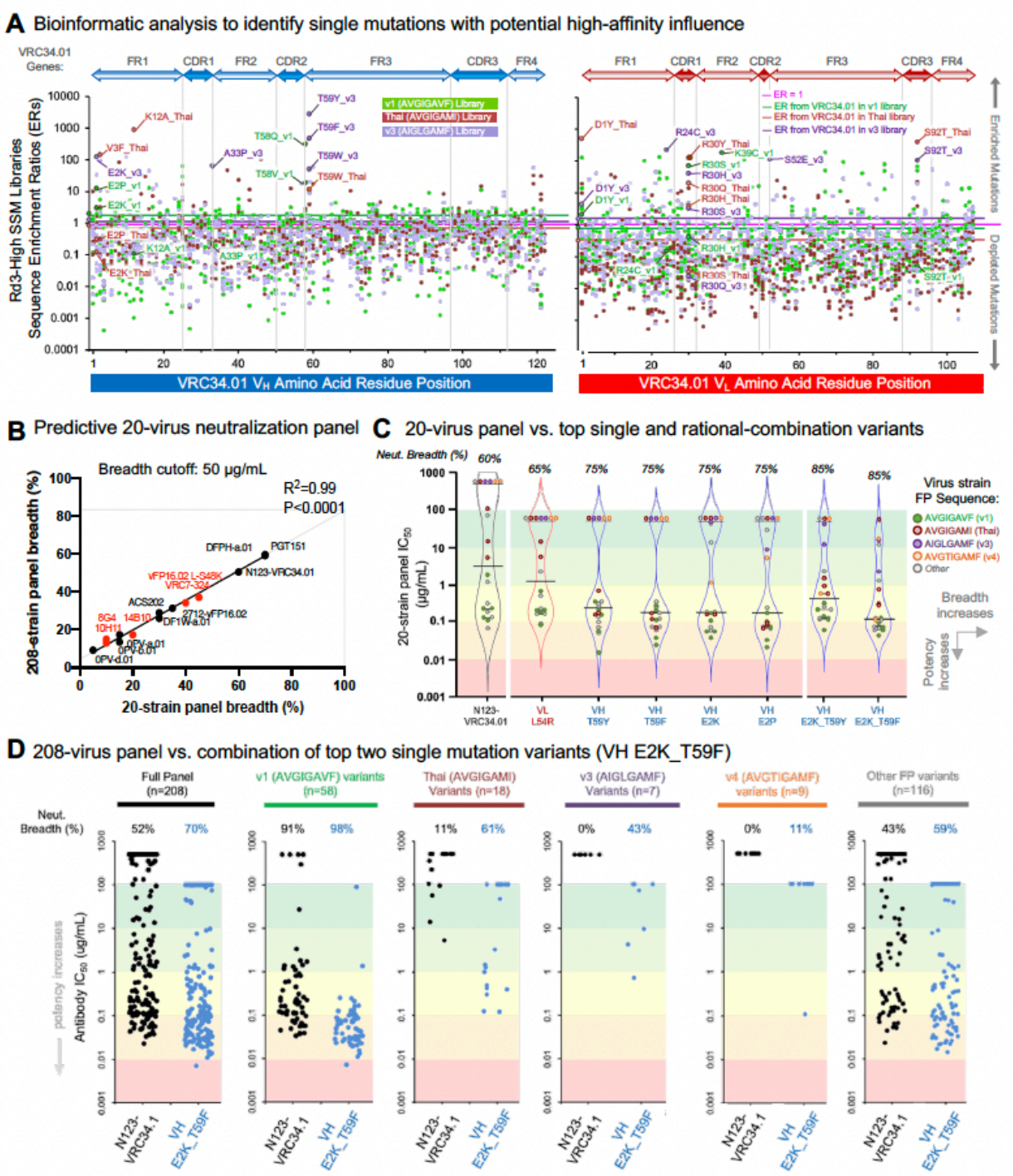


Figure 20. Bioinformatic mining of single-mutation NGS data from SSM-library screens revealed multiple mutations that provided enhanced HIV-1 neutralization potency and breadth.

Figure 30. Bioinformatic mining of single-mutation NGS data from SSM-library screens revealed multiple mutations that provided enhanced HIV-1 neutralization potency and breadth.

(A) Enrichment ratios (ER) are plotted for single mutant antibody sequences derived from round three high-affinity sorted libraries against the mutant residue location in the antibody variable region. NGS data analysis of single mutant library screens highlight multiple amino acid substitutions across VRC34.01 antibody variable regions predicted to enhanced diverse fusion peptide recognition.

(B) A predictive virus panel was designed as a pre-screening tool to select potentially broad candidates prior to 208-virus panel neutralization analysis. Template VRC34.01 recognized 60% of strains in the 20-virus panel compared to PGT151, the most broad previously reported antibody that interacts with FP, which recognized 70% of the 20 virus panel.

(C) Four single mutations (VH_T59Y, VH_T59F, VH_E2K, VH_E2P) discovered by SSM screening and expressed as soluble IgG increased neutralization to 75% of the 20-virus panel and rational combinations of top single-mutant variants further improved neutralization breadth, with a maximum 85% breadth achieved for VH_E2K_T59F (also referred to as VRC34.01-combo1). Antibodies were considered neutralizing if their IC50 potency was less than 50 μg/mL.

(D) 208-virus panel data revealed enhanced breadth for the most-improved VH_E2K_T59F antibody variant. VH_E2K_T59F showed strong neutralization of FP_v1 strains, moderate neutralization of FP_Thai strains, and poor neutralization of FP_v3 and FP_v4 strains, despite improvements against the template VRC34.01 antibody against all HIV-1 FP subclasses. These data suggested that gaps in recognition remained against HIV-1 strains that encoded diverse FP sequences, especially FP_v4 (AVGTIGAMF).

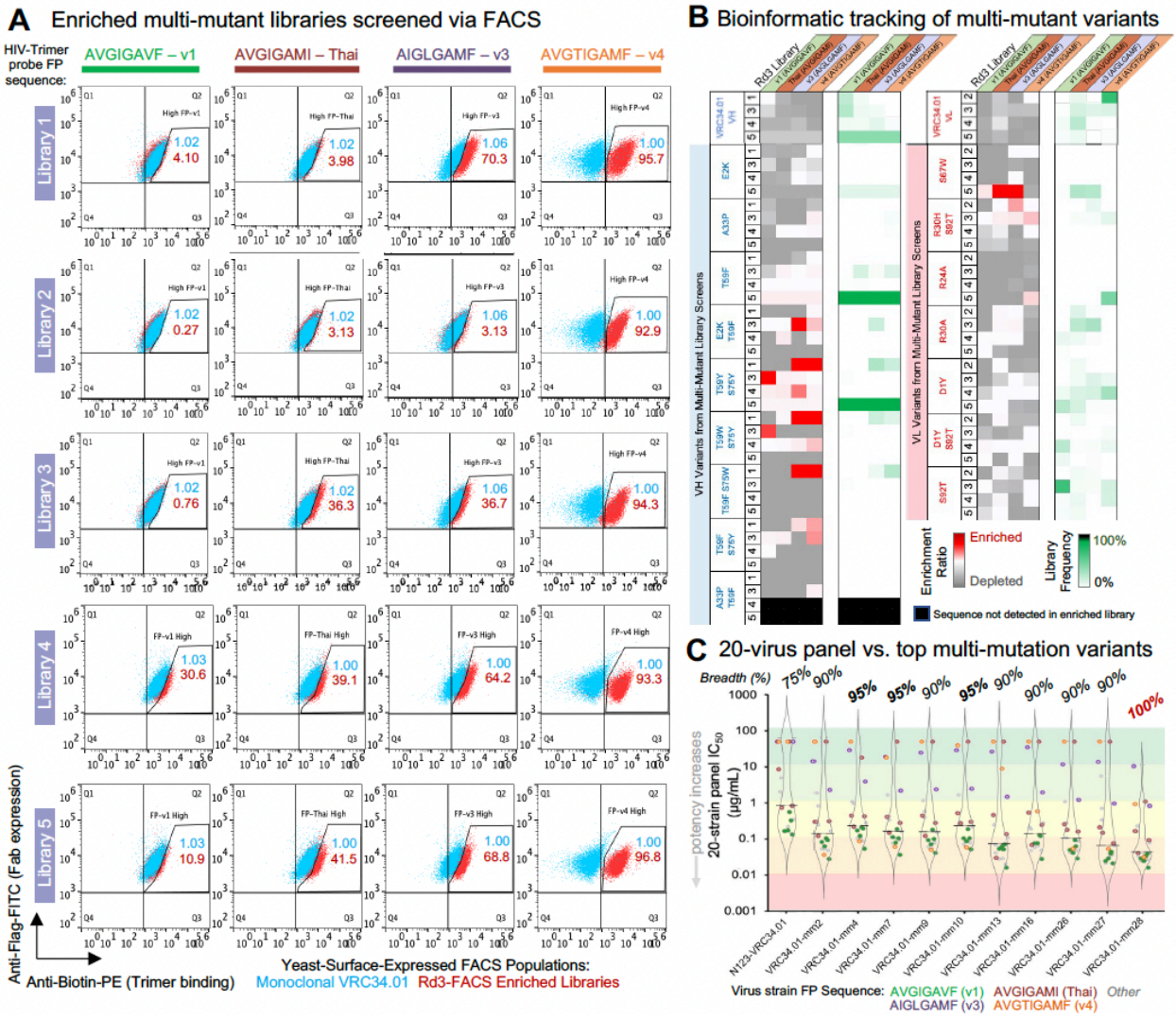


Figure 21. Targeted multi-mutation screening enhanced HIV-1 neutralization of FP8_v3 and FP8_v4 strains, enabling best-in-class anti-FP neutralization breadth and potency.

(A) Multi-mutation yeast libraries expressing surface-bound antibody fragments (Fab) visualized as fluorescent activated cell sorting (FACS) populations (red) compared to surface-expressed VRC34.01 template Fab. Multi-mutant library generation and FACS screening provided visible affinity improvements against BG505 SOSIP encoding FP_Thai, FP_v3, and FP_v4 relative to template mAb VRC34.01, potentially addressing the gaps observed in HIV_1 neutralization of VH_E2K_T59F (Fig. 2D).

(B) Heat maps of enriched and prevalent mutant sequences from screening multi-mutation library 1, 2, 3, 4, and 5 against trimer probes displaying four different FP variants. Bioinformatic data mining of multi-mutation yeast display library screens revealed several mutations enriched against diverse fusion peptide sequences.

(C) When expressed solubly as IgG, several mutational variants achieved 95% neutralization in the 20-virus panel, while one variant (VRC34.01_mm28, VH_E2K_A33P_T59F) with three heavy chain mutations from the template antibody achieved 100% neutralization of HIV-1 strains in the panel. Antibodies were considered neutralizing if their IC50 potency was less than 50 μg/mL.

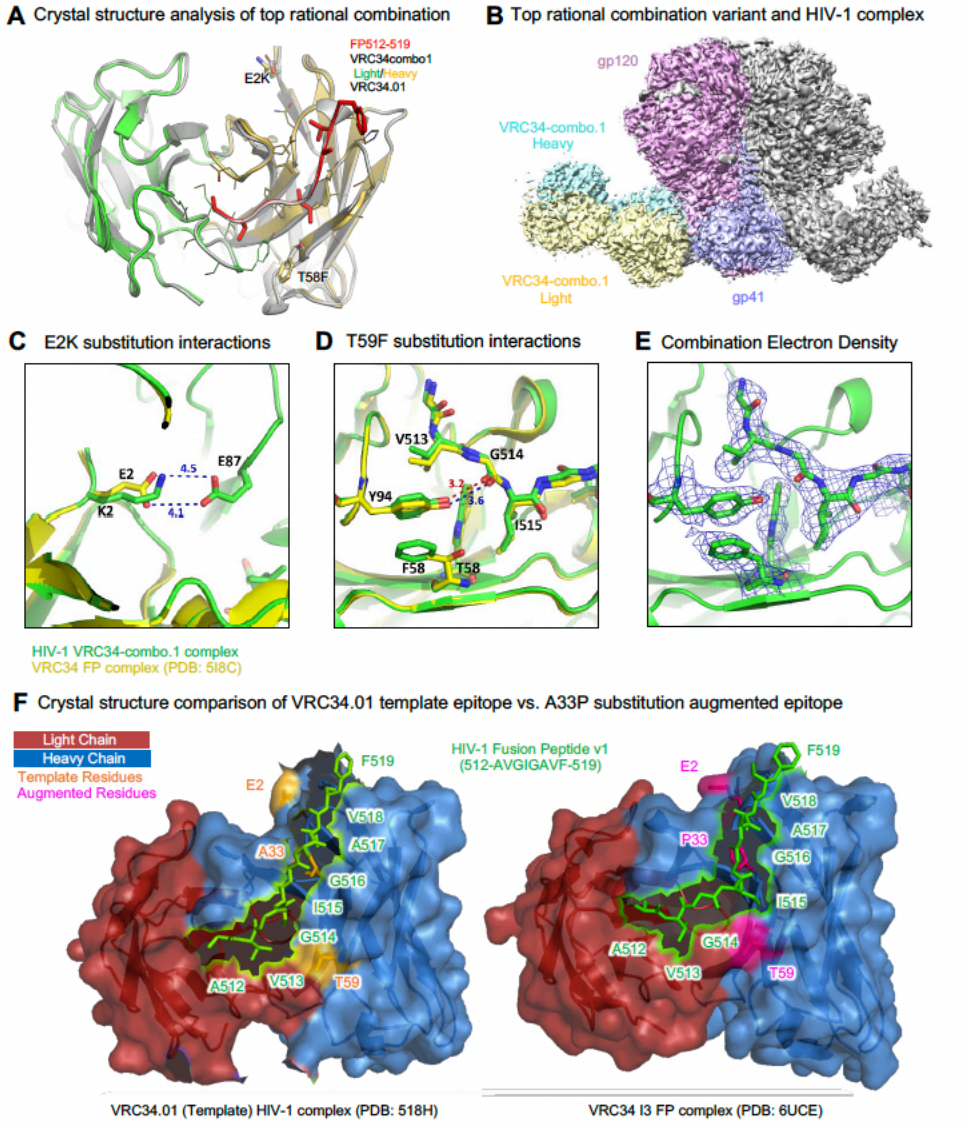


Figure 22. Structural analysis of Combo-1 and mm28 revealed three synergistic interactions that enabled enhanced anti-FP breadth and potency, including accommodation of the amino acid insertion in FP_v4.

(A) Super-imposed crystal structures of VRC34.01-combo1 Fab (1.2 Å) and VRC34.01 Fab. The T58F substitution enhances the hydrophobic interaction to W50heavy and Y94light, and removes an unfavorable interaction between T58 hydroxyl group and I515 on the gp41 FP. Cryo-EM map of HIV-1 trimer in complex with VRC34-combo.1 Fab.

(B) The E2K substitution on VRC34-combo1 removes electrostatic conflict with HIV-1 E87 residue, thereby increasing the affinity VRC34-combo.1 to HIV-1 trimer.

(C) The T58F substitution on VRC34-combo1-FP alters the position of Y94, moving it closer to the carbonyl of G514 increasing the affinity of VRC34-combo.1 to FP.

(D) An example of electron density for the cryo-EM structure of HIV-1 VRC34-combo.1 contoured at 8.0 sigma.

(E) Crystal structure comparison of template VRC34.01 vs. VRC34.01_mm28 shows how the substitution A33P widens the fusion peptide binding groove to enable recognition of diverse fusion peptide sequences that contain insertions, such as FP_v4 (AVGTIGAMF).Template antibody VRC34.01 is modeled in complex with HIV-1 trimer (PDB: 518H), and VRC34.01_mm28 is modeled using VRC34 I3 in complex with fusion peptide v1 (AVGIGAVF) (PDB: 6UCE).

5. Dissertation Conclusions and Future Perspectives

Antibody discovery, optimization, and application: translational protein engineering for precision medicines. Translational "omics-based" immune engineering leverages features of genetically determined pathophysiology to design precise solutions against pathogens of clinical significance. Collaborations between scientists, researchers, clinical doctors, and public health officials have led to the invention of renewable, high-throughput antibody engineering methods for detailing humoral immunity at an extraordinary rate. This in-depth biomolecular characterization of monoclonal antibody proteins can be used to generate translational scientific knowledge addressing critical gaps in our understanding of immunity that can be used in turn to generate new clinical medicines. Application of this knowledge can be achieved by defining the relationship between antibody-mediated infection and disease outcomes.

Here we provide data on antibody gene-specific functions related to the targeted elimination of clinically significant pathogens to inform the design of advanced precision medicines. Three separate scientific studies mapping the capacity of genetic antibody immunity features to impact infection outcomes are presented. These research studies provide detailed functional knowledge of protective antibody features to help inform new clinical disease intervention strategies related to SARS-CoV-2, malaria, and HIV-1, respectively. In the first study, a novel SARS-CoV-2 neutralizing monoclonal antibody: 910-30, was discovered from a patient infected during an acute COVID-19 global pandemic outbreak. Through the generation and high-throughput screening of natively paired antibody libraries derived from this patient's antigen-experienced B cells, monoclonal antibody 910-30 and other SARS-CoV-2 specific antibodies were selected and characterized to describe shared community features of humoral immunity against COVID-19 infection in the larger population.

Also described are two translational antibody engineering methods used to define new, functionally optimized, genetic sequences from previously identified monoclonal antibodies: CIS43 and VRC34.01, against the clinically relevant pathogens: parasitic *plasmodium falciparum* and viral HIV-1, respectively. The molecular insights gained from these antibody engineering studies were used to produce functionally

improved antibody variants that generate insights into forward pathways for highly potent protective humoral immunity against *p. falciparum* parasite infection and defined structural mechanisms for broad humoral immune recognition of highly diverse HIV-1 viral variants. Altogether, the scientific outcomes from this research have immediate clinical impact for the development of therapeutic, prophylactic, and diagnostic and research reagents for COVID-19, HIV/AIDS, and malaria, as well as broader scientific significance for the development of precision medicines against various pathogens of clinical relevance.

5.1 Limitations

The limitations of this research are defined by the capacity of our screening system to recapitulate natural *in vivo* antibody function. As the goal for this research is to produce biological materials with high clinical translatability, a key feature of these methods is the rational experimental design based on fundamental knowledge of biological interactions. One important feature to note about this antibody screening platform is that phenotypic selection occurs via affinity-based binding, which is not always a direct correlation to a given biological functional outcome of interest. Antibody sequences which bind well to antigen but do not provide functional improvement limit the true positive hit-rate of our screening approach and may reduce the number or quality of antibody candidates identified for clinical translation. In addition, the screening platform described herein focused on antibody fragment antigen binding (Fab) formats which do not include an Fc region, limiting evaluation of screened hits' effector functions to downstream soluble characterization via low-throughput formats, further impeding the rate of clinical translation. Modified screening platforms which could capture a range of desired drug-ability features prior to phenotype selection would be able to enhance the rate of translational antibody hits compared to the limited Fab-affinity based screening methods applied here.

For many of the library generation and screening experiments, obstacles surrounding antigen probe fabrication and conjugation influenced the way libraries could be functionally evaluated. For example, at the onset of the SARS-CoV-2 outbreak (Chapter 2), the first generation of antigen probes used for patient library screening and anti-SARS-CoV-2 antibody discovery contained a high number of biotinylated sites,

generating large amount of noise and off-target binding in subsequent screening experiments. Additionally, for the directed antibody evolution projects (Chapters 3 and 4) screening was limited to the handful of available antigens in panel screening formats, which requires time-consuming multiplexed wet-lab methods and does not fully reflect the large biodiversity among global clinical populations. In addition, yeast library transformation methods constrain the transformed library size which can be evaluated by yeast display to around 1 million cells (10^6 sequences /yeast transformation). In addition, throughput limitations from single-cell approaches dictate the wet-lab time required to screen a given library size. As we are limited in screening by the quality of antigen and size of library which can be evaluated in a reasonable time scale, careful consideration should then be given for library design and antigen selection pre-screening based on native antibody development insights to increase screening hits.

In addition to these limitations, two of the projects described in the body of work here intended to evaluate paired combinations of multi-mutation variants (Chapters 2 and 3), however, the MiSeq sequencing platform read length (500 bp) limits the size of amplicon sequencing to either variable region, and does not reveal how antibody mutations were paired naturally during library screening (paired sequences from yeast display libraries are up to ~1200 bp in length, >500bp MiSeq read capacity). Thus, the risk of incorrectly mismatching functionally optimized paired VH:VL sequences from separate VH and VL datasets derived from MiSeq analysis VH:VL display libraries can further delay the rate of improved phenotype discovery and characterization.

5.2 Considerations for Future Research

Future research into translational antibody engineering should serve to address the current limitations in experimental methods discussed while expanding on these methods to probe new features of the human immune system. To make significant conclusions about the native immune response to a given pathogen or selective pressure, it is critical that researchers have access to study the functional immune repertoire form clinical phenotypes of interest. Due to the dynamic fundamental nature of disease progression and adaptive immunity it is often necessary to collect samples of clinical interest across

multiple time points to gain an important temporal layer of depth for repertoire development analysis. Access to these kinds of precious clinical samples is not readily available to many researchers. Thus, studies aiming to elucidate sophisticated immune mechanisms hinge upon collaborative efforts between clinicians, specialists, or independent parties that have built the necessary patient-sample banks and accurately logged clinical phenotype data across time scales long enough to capture humoral immune evolution. These collaborations are largely built on pre-existing networks or relationships among researchers; therefore, inter-disciplinary communication and community are key in advancing the field.

The research presented herein focused exclusively on antibody discovery and characterization against infectious disease, however, these translational antibody engineering methods could provide broad scientific benefit to many other immune-modulated clinical dilemmas related to autoimmunity, cancer, allergies, transplantation rejection, etc. as well as a variety of applications beyond the clinic. Future considerations for translational antibody engineering in precision medicine design may include replacing antigen-screening panels with antigen libraries for comprehensive library versus library immune profiling capabilities. Considerations for experimental antigen probe design and selection should reflect Ag features correlated to desired functional outcomes, or the screening system itself should be modified further to select for more comprehensive drug features beyond antigen-recognition improvement alone. Sequencing capacity should be expanded to accommodate the length of paired antibody VH:VL sequences from library samples using platforms such as PacBio to obtain high throughput paired genetic information. Finally, immediate continuation of the intellectual contributions of this work should serve to further evaluate and advance the discovery and engineering of clinically relevant antibody molecules related to SARS-CoV-2, malaria, and HIV-1. This can be achieved by delivering new antibody-based drugs to the body as therapeutics through injection, infusion, or delivery of exogenous mRNA encoding antibody genes of interest; in addition, antibody lineages with similar mechanistic features could be elicited through antibody-informed immunogens designed to generate antibody-based immune protection.

6. PUBLICATION LIST

6.1 Peer-reviewed articles related to dissertation

Banach, B. B., Cerutti, G., Fahad, A. S., Shen, C. H., Oliveira De Souza, M., Katsamba, P. S., Tsybovsky, Y., Wang, P., Nair, M. S., Huang, Y., Francino-Urdániz, I. M., Steiner, P. J., Gutiérrez-González, M., Liu, L., López Acevedo, S. N., Nazzari, A. F., Wolfe, J. R., Luo, Y., Olia, A. S., Teng, I. T., … DeKosky, B. J. (2021). Paired heavy- and light-chain signatures contribute to potent SARS-CoV-2 neutralization in public antibody responses. *Cell reports*, *37*(1), 109771. https://doi.org/10.1016/j.celrep.2021.109771

Madan, B., Reddem, E. R., Wang, P., Casner, R. G., Nair, M. S., Huang, Y., Fahad, A. S., de Souza, M. O., **Banach, B. B.**, López Acevedo, S. N., Pan, X., Nimrania, R., Teng, I. T., Bahna, F., Zhou, T., Zhang, B., Yin, M. T., Ho, D. D., Kwong, P. D., Shapiro, L., … DeKosky, B. J. (2021). Antibody screening at reduced pH enables preferential selection of potently neutralizing antibodies targeting SARS-CoV-2. *AIChE journal. American Institute of Chemical Engineers*, *67*(12), e17440. https://doi.org/10.1002/aic.17440

Zhou, T., Teng, I. T., Olia, A. S., Cerutti, G., Gorman, J., Nazzari, A., Shi, W., Tsybovsky, Y., Wang, L., Wang, S., Zhang, B., Zhang, Y., Katsamba, P. S., Petrova, Y., **Banach, B. B.,** Fahad, A. S., Liu, L., Lopez Acevedo, S. N., Madan, B., Oliveira de Souza, M., … Kwong, P. D. (2020). Structure-Based Design with Tag-Based Purification and In-Process Biotinylation Enable Streamlined Development of SARS-CoV-2 Spike Molecular Probes. *Cell reports*, *33*(4), 108322. https://doi.org/10.1016/j.celrep.2020.108322

6.2 Peer-reviewed articles prior to University of Kansas

Wiley, L. A., Burnight, E. R., Drack, A. V., **Banach, B. B**., Ochoa, D., Cranston, C. M., Madumba, R. A., East, J. S., Mullins, R. F., Stone, E. M., & Tucker, B. A. (2016). Using Patient-Specific Induced Pluripotent Stem Cells and Wild-Type Mice to Develop a Gene Augmentation-Based Strategy to Treat CLN3-Associated Retinal Degeneration. *Human gene therapy*, *27*(10), 835–846. https://doi.org/10.1089/hum.2016.049

6.3 Patents related to dissertation

DeKosky BJ, **Banach BB**, de Souza, MO, Wolfe J, Madan B, Fahad A, Acevedo SL, Pan X, Ho D, Wang P, Liu L, Yin M, "Anti-SARS-CoV-2 antibodies and uses thereof," U.S. Provisional Patent Appl. 63/050226, July 10, 2020.

Kwong PD, DeKosky BJ, …, **Banach BB**, et al, "Neutralizing antibodies to plasmodium falciparum circumsporozoite protein and their use," U.S. Provisional Patent Appl.63/299,730, January 14, 2022

7. References

1. H. Matthews, J. Hanison, N. Nirmalan, "Omics"-Informed Drug and Biomarker Discovery: Opportunities, Challenges and Future Perspectives. *Proteomes* **4**, (2016).
2. V. M. Chernov, O. A. Chernova, A. A. Mouzykantov, L. L. Lopukhov, R. I. Aminov, Omics of antimicrobials and antimicrobial resistance. *Expert Opin Drug Discov* **14**, 455-468 (2019).
3. P. D. Kwong *et al.*, Antibodyomics: bioinformatics technologies for understanding B-cell immunity to HIV-1. *Immunol Rev* **275**, 108-128 (2017).
4. D. Ribatti, From the discovery of monoclonal antibodies to their therapeutic application: an historical reappraisal. *Immunol Lett* **161**, 96-99 (2014).
5. K. Strebhardt, A. Ullrich, Paul Ehrlich's magic bullet concept: 100 years of progress. *Nat Rev Cancer* **8**, 473-480 (2008).
6. S. F. de StGroth, D. Scheidegger, Production of monoclonal antibodies: strategy and tactics. *J Immunol Methods* **35**, 1-21 (1980).
7. R. Kunert, D. Reinhart, Advances in recombinant antibody manufacturing. *Appl Microbiol Biotechnol* **100**, 3451-3461 (2016).
8. M. R. Dyson, Fundamentals of Expression in Mammalian Cells. *Adv Exp Med Biol* **896**, 217-224 (2016).
9. A. Six *et al.*, The past, present, and future of immune repertoire biology - the rise of next-generation repertoire analysis. *Front Immunol* **4**, 413 (2013).
10. G. Georgiou *et al.*, The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* **32**, 158-168 (2014).
11. W. H. Robinson, Sequencing the functional antibody repertoire--diagnostic and therapeutic discovery. *Nat Rev Rheumatol* **11**, 171-182 (2015).
12. J. Glanville *et al.*, Deep sequencing in library selection projects: what insight does it bring? *Curr Opin Struct Biol* **33**, 146-160 (2015).
13. J. A. Weinstein, X. Zeng, Y. H. Chien, S. R. Quake, Correlation of gene expression and genome mutation in single B-cells. *PLoS One* **8**, e67624 (2013).
14. J. L. Xu, M. M. Davis, Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* **13**, 37-45 (2000).
15. S. Friedensohn, T. A. Khan, S. T. Reddy, Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends Biotechnol* **35**, 203-214 (2017).
16. M. A. Turchaninova *et al.*, Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol* **43**, 2507-2515 (2013).
17. M. McHeyzer-Williams, S. Okitsu, N. Wang, L. McHeyzer-Williams, Molecular programming of B cell memory. *Nat Rev Immunol* **12**, 24-34 (2011).
18. B. J. DeKosky *et al.*, High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* **31**, 166-169 (2013).
19. J. R. McDaniel, B. J. DeKosky, H. Tanno, A. D. Ellington, G. Georgiou, Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. *Nat Protoc* **11**, 429-442 (2016).
20. B. Wang *et al.*, Functional interrogation and mining of natively paired human VH:VL antibody repertoires. *Nat Biotechnol* **36**, 152-155 (2018).
21. Z. Sheng *et al.*, Gene-Specific Substitution Profiles Describe the Types and Frequencies of Amino Acid Changes during Antibody Somatic Hypermutation. *Front Immunol* **8**, 537 (2017).

22. K. Imkeller *et al.*, Antihomotypic affinity maturation improves human B cell responses against a repetitive epitope. *Science* **360**, 1358-1362 (2018).
23. J. Zhu *et al.*, Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. *Proc Natl Acad Sci U S A* **110**, 6470-6475 (2013).
24. J. Zhu *et al.*, De novo identification of VRC01 class HIV-1-neutralizing antibodies by next-generation sequencing of B-cell transcripts. *Proc Natl Acad Sci U S A* **110**, E4088-4097 (2013).
25. J. S. Klein, P. J. Bjorkman, Few and far between: how HIV may be evading antibody avidity. *PLoS Pathog* **6**, e1000908 (2010).
26. K. O. Saunders *et al.*, Targeted selection of HIV-specific antibody mutations by engineering B cell maturation. *Science* **366**, (2019).
27. C. A. Schramm, D. C. Douek, Beyond Hot Spots: Biases in Antibody Somatic Hypermutation and Implications for Vaccine Design. *Front Immunol* **9**, 1876 (2018).
28. R. Murugan *et al.*, Evolution of protective human antibodies against Plasmodium falciparum circumsporozoite protein repeat motifs. *Nat Med* **26**, 1135-1145 (2020).
29. V. Jawa *et al.*, T-Cell Dependent Immunogenicity of Protein Therapeutics Pre-clinical Assessment and Mitigation-Updated Consensus and Review 2020. *Front Immunol* **11**, 1301 (2020).
30. A. Vaisman-Mentesh, M. Gutierrez-Gonzalez, B. J. DeKosky, Y. Wine, The Molecular Mechanisms That Underlie the Immune Biology of Anti-drug Antibody Formation Following Treatment With Monoclonal Antibodies. *Front Immunol* **11**, 1951 (2020).
31. C.-H. Shen *et al.*, VRC34-Antibody Lineage Development Reveals How a Required Rare Mutation Shapes the Maturation of a Broad HIV-Neutralizing Lineage. *Cell Host & Microbe* **27**, 531-543.e536 (2020).
32. M. Keller *et al.*, Ancient Yersinia pestis genomes from across Western Europe reveal early diversification during the First Pandemic (541-750). *Proc Natl Acad Sci U S A* **116**, 12363-12372 (2019).
33. P. V'Kovski, A. Kratzel, S. Steiner, H. Stalder, V. Thiel, Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol* **19**, 155-170 (2021).
34. C. Theves, E. Crubezy, P. Biagini, History of Smallpox and Its Spread in Human Populations. *Microbiol Spectr* **4**, (2016).
35. T. H. M. L. o. T. C. o. P. o. Philadelphia, in *Lecture Memoranda*. (https://www.historyofvaccines.org/content/representation-smallpox, 1913), vol. 2022, pp. A.M.A. Meeting, Minneapolis. Burroughs Wellcome and Co. London.
36. S. M. Helfert, Historical aspects of immunization and vaccine safety communication. *Curr Drug Saf* **10**, 5-8 (2015).
37. C. C. Chen *et al.*, From germ theory to germ therapy. *Kaohsiung J Med Sci* **35**, 73-82 (2019).
38. S. H. Kaufmann, Remembering Emil von Behring: from Tetanus Treatment to Antibody Cooperation with Phagocytes. *mBio* **8**, (2017).
39. N. C. Sharma *et al.*, Diphtheria. *Nat Rev Dis Primers* **5**, 81 (2019).
40. T. Hifumi *et al.*, Clinical Serum Therapy: Benefits, Cautions, and Potential Applications. *Keio J Med* **66**, 57-64 (2017).
41. S. H. Institute. (2017), vol. 2021.
42. I. M. Nobel Lectures. (Elsevier Publishing Company, Amsterdam, 1967).

43. S. L. Nutt, P. D. Hodgkin, D. M. Tarlinton, L. M. Corcoran, The generation of antibody-secreting plasma cells. *Nat Rev Immunol* **15**, 160-171 (2015).
44. D. Ribatti, Edelman's view on the discovery of antibodies. *Immunol Lett* **164**, 72-75 (2015).
45. P. Portin, The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA. *J Genet* **93**, 293-302 (2014).
46. G. Goldstein, Overview of the development of Orthoclone OKT3: monoclonal antibody for therapeutic use in transplantation. *Transplant Proc* **19**, 1-6 (1987).
47. F. Matsuda *et al.*, The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J Exp Med* **188**, 2151-2162 (1998).
48. F. Carrasco-Ramiro, R. Peiro-Pastor, B. Aguado, Human genomics projects and precision medicine. *Gene Ther* **24**, 551-561 (2017).
49. J. S. Lanchbury, The Human Genome Project. *Br J Rheumatol* **37**, 119-121 (1998).
50. J. A. Schloss, R. A. Gibbs, V. B. Makhijani, A. Marziali, Cultivating DNA Sequencing Technology After the Human Genome Project. *Annu Rev Genomics Hum Genet* **21**, 117-138 (2020).
51. A. S. De Groot *et al.*, Better Epitope Discovery, Precision Immune Engineering, and Accelerated Vaccine Design Using Immunoinformatics Tools. *Front Immunol* **11**, 442 (2020).
52. B. Pulendran, M. M. Davis, The science and medicine of human immunology. *Science* **369**, (2020).
53. S. Sell, How vaccines work: immune effector mechanisms and designer vaccines. *Expert Rev Vaccines* **18**, 993-1015 (2019).
54. P. Gogesch, S. Dudek, G. van Zandbergen, Z. Waibler, M. Anzaghe, The Role of Fc Receptors on the Effectiveness of Therapeutic Monoclonal Antibodies. *Int J Mol Sci* **22**, (2021).
55. C. Vettermann, K. Herrmann, H. M. Jack, Powered by pairing: the surrogate light chain amplifies immunoglobulin heavy chain signaling and pre-selects the antibody repertoire. *Semin Immunol* **18**, 44-55 (2006).
56. D. S. Chahar, S. Ravindran, S. S. Pisal, Monoclonal antibody purification and its progression to commercial scale. *Biologicals* **63**, 1-13 (2020).
57. A. Hayhurst, G. Georgiou, High-throughput antibody isolation. *Curr Opin Chem Biol* **5**, 683-689 (2001).
58. P. Zuo, Capturing the Magic Bullet: Pharmacokinetic Principles and Modeling of Antibody-Drug Conjugates. *AAPS J* **22**, 105 (2020).
59. P. J. Hudson, Recombinant antibody fragments. *Curr Opin Biotechnol* **9**, 395-402 (1998).
60. Z. A. Ahmad *et al.*, scFv antibody: principles and clinical application. *Clin Dev Immunol* **2012**, 980250 (2012).
61. S. Steeland, R. E. Vandenbroucke, C. Libert, Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discov Today* **21**, 1076-1113 (2016).
62. P. Bemani, M. Mohammadi, A. Hakakian, ScFv Improvement Approaches. *Protein Pept Lett* **25**, 222-229 (2018).
63. X. Wang, M. Mathieu, R. J. Brezski, IgG Fc engineering to modulate antibody effector functions. *Protein Cell* **9**, 63-73 (2018).
64. T. T. Wang, IgG Fc Glycosylation in Human Immunity. *Curr Top Microbiol Immunol* **423**, 63-75 (2019).

65. M. J. Coloma, S. L. Morrison, Design and production of novel tetravalent bispecific antibodies. *Nat Biotechnol* **15**, 159-163 (1997).
66. A. M. Cuesta, N. Sainz-Pastor, J. Bonet, B. Oliva, L. Alvarez-Vallina, Multivalent antibodies: when design surpasses evolution. *Trends Biotechnol* **28**, 355-362 (2010).
67. J. Golay *et al.*, Design and Validation of a Novel Generic Platform for the Production of Tetravalent IgG1-like Bispecific Antibodies. *J Immunol* **196**, 3199-3211 (2016).
68. M. Takahashi, S. A. Fuller, S. Winston, Design and production of bispecific monoclonal antibodies by hybrid hybridomas for use in immunoassay. *Methods Enzymol* **203**, 312-327 (1991).
69. A. Thakur, M. Huang, L. G. Lum, Bispecific antibody based therapeutics: Strengths and challenges. *Blood Rev* **32**, 339-347 (2018).
70. Q. Wang *et al.*, Design and Production of Bispecific Antibodies. *Antibodies (Basel)* **8**, (2019).
71. K. L. Wu *et al.*, Precision Modification of Native Antibodies. *Bioconjug Chem* **32**, 1947-1959 (2021).
72. A. Beck, L. Goetsch, C. Dumontet, N. Corvaia, Strategies and challenges for the next generation of antibody-drug conjugates. *Nat Rev Drug Discov* **16**, 315-337 (2017).
73. M. J. Birrer, K. N. Moore, I. Betella, R. C. Bates, Antibody-Drug Conjugate-Based Therapeutics: State of the Science. *J Natl Cancer Inst* **111**, 538-549 (2019).
74. D. R. Goulet, W. M. Atkins, Considerations for the Design of Antibody-Based Therapeutics. *J Pharm Sci* **109**, 74-103 (2020).
75. W. R. Strohl, Current progress in innovative engineered antibodies. *Protein Cell* **9**, 86-120 (2018).
76. L. Benatuil, J. M. Perez, J. Belk, C. M. Hsieh, An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng Des Sel* **23**, 155-159 (2010).
77. A. S. Fahad *et al.*, Functional Profiling of Antibody Immune Repertoires in Convalescent Zika Virus Disease Patients. *Front Immunol* **12**, 615102 (2021).
78. B. Madan *et al.*, Mutational fitness landscapes reveal genetic and structural improvement pathways for a vaccine-elicited HIV-1 broadly neutralizing antibody. *Proc Natl Acad Sci U S A* **118**, (2021).
79. A. J. Meyer, J. W. Ellefson, A. D. Ellington, Library generation by gene shuffling. *Curr Protoc Mol Biol* **105**, Unit 15 12 (2014).
80. R. Rouet, K. J. L. Jackson, D. B. Langley, D. Christ, Next-Generation Sequencing of Antibody Display Repertoires. *Front Immunol* **9**, 118 (2018).
81. E. E. Wrenbeck *et al.*, Plasmid-based one-pot saturation mutagenesis. *Nat Methods* **13**, 928-930 (2016).
82. Y. Cao *et al.*, Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73-84 e16 (2020).
83. B. Ju *et al.*, Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119 (2020).
84. L. Liu *et al.*, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450-456 (2020).

85. E. Traggiai *et al.*, An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nat Med* **10**, 871-875 (2004).
86. D. Cucinotta, M. Vanelli, WHO Declares COVID-19 a Pandemic. *Acta Biomed* **91**, 157-160 (2020).
87. Y. Liu *et al.*, Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature* **582**, 557-560 (2020).
88. V. Coronaviridae Study Group of the International Committee on Taxonomy of, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).
89. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181-192 (2019).
90. E. de Wit, N. van Doremalen, D. Falzarano, V. J. Munster, SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol* **14**, 523-534 (2016).
91. R. L. Graham, R. S. Baric, Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol* **84**, 3134-3146 (2010).
92. T. G. Ksiazek *et al.*, A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953-1966 (2003).
93. A. M. Zaki, S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus, R. A. Fouchier, Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814-1820 (2012).
94. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).
95. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
96. X. Ou *et al.*, Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* **11**, 1620 (2020).
97. T. Tang, M. Bidon, J. A. Jaimes, G. R. Whittaker, S. Daniel, Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Res* **178**, 104792 (2020).
98. D. J. Benton *et al.*, Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **588**, 327-330 (2020).
99. Y. Cai *et al.*, Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586-1592 (2020).
100. C. O. Barnes *et al.*, Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-842 e816 (2020).
101. P. J. M. Brouwer *et al.*, Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643-650 (2020).
102. X. Chi *et al.*, A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650-655 (2020).
103. S. Du *et al.*, Structurally Resolved SARS-CoV-2 Antibody Shows High Efficacy in Severely Infected Hamsters and Provides a Potent Cocktail Pairing Strategy. *Cell* **183**, 1013-1023 e1013 (2020).
104. J. Hansen *et al.*, Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010-1014 (2020).

105. N. K. Hurlburt *et al.*, Structural basis for potent neutralization of SARS-CoV-2 and role of antibody affinity maturation. *Nat Commun* **11**, 5413 (2020).
106. T. F. Rogers *et al.*, Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956-963 (2020).
107. E. Seydoux *et al.*, Analysis of a SARS-CoV-2-Infected Individual Reveals Development of Potent Neutralizing Antibodies with Limited Somatic Mutation. *Immunity* **53**, 98-105 e105 (2020).
108. R. Shi *et al.*, A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120-124 (2020).
109. Y. Wu *et al.*, A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **368**, 1274-1278 (2020).
110. M. Yuan *et al.*, Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119-1123 (2020).
111. N. C. Wu *et al.*, An Alternative Binding Mode of IGHV3-53 Antibodies to the SARS-CoV-2 Receptor Binding Domain. *Cell Rep* **33**, 108274 (2020).
112. J. Catalan-Dibene, Human antibodies can neutralize SARS-CoV-2. *Nat Rev Immunol* **20**, 350 (2020).
113. C. Wang *et al.*, A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat Commun* **11**, 2251 (2020).
114. D. Wrapp *et al.*, Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **181**, 1436-1441 (2020).
115. D. F. Robbiani *et al.*, Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442 (2020).
116. S. J. Zost *et al.*, Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nat Med* **26**, 1422-1427 (2020).
117. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **183**, 1735 (2020).
118. T. Zhou *et al.*, Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. *Cell Host Microbe* **28**, 867-879 e865 (2020).
119. B. Korber *et al.*, Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812-827 e819 (2020).
120. S. W. Long *et al.*, Molecular Architecture of Early Dissemination and Massive Second Wave of the SARS-CoV-2 Virus in a Major Metropolitan Area. *mBio* **11**, (2020).
121. E. Volz *et al.*, Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell* **184**, 64-75 e11 (2021).
122. L. Yurkovetskiy *et al.*, Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **183**, 739-751 e738 (2020).
123. L. Zhang *et al.*, The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv*, (2020).
124. N. D. Grubaugh, W. P. Hanage, A. L. Rasmussen, Making Sense of Mutation: What D614G Means for the COVID-19 Pandemic Remains Unclear. *Cell* **182**, 794-795 (2020).
125. Y. J. Hou *et al.*, SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **370**, 1464-1468 (2020).
126. R. A. Mansbach *et al.*, The SARS-CoV-2 Spike variant D614G favors an open conformational state. *Sci Adv* **7**, (2021).

127. K. K. To *et al.*, Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* **20**, 565-574 (2020).
128. P. Wang *et al.*, SARS-CoV-2 neutralizing antibody responses are more robust in patients with severe disease. *Emerg Microbes Infect* **9**, 2091-2093 (2020).
129. J. Huo *et al.*, Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe* **28**, 497 (2020).
130. J. ter Meulen *et al.*, Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med* **3**, e237 (2006).
131. T. N. Starr *et al.*, Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310 e1220 (2020).
132. M. P. Lefranc *et al.*, IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* **27**, 55-77 (2003).
133. A. Brauninger, T. Goossens, K. Rajewsky, R. Kuppers, Regulation of immunoglobulin light chain gene rearrangements during early B cell development in the human. *Eur J Immunol* **31**, 3631-3637 (2001).
134. Z. Sethna, Y. Elhanati, C. G. Callan, A. M. Walczak, T. Mora, OLGA: fast computation of generation probabilities of B- and T-cell receptor amino acid sequences and motifs. *Bioinformatics* **35**, 2974-2981 (2019).
135. C. Soto *et al.*, High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* **566**, 398-402 (2019).
136. T. Zhou *et al.*, Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245-258 (2013).
137. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
138. P. Wang *et al.*, Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130-135 (2021).
139. Y. Weisblum *et al.*, Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* **9**, (2020).
140. D. Weissman *et al.*, D614G Spike Mutation Increases SARS CoV-2 Susceptibility to Neutralization. *Cell Host Microbe* **29**, 23-31 e24 (2021).
141. M. A. Tortorici *et al.*, Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **370**, 950-957 (2020).
142. Y. Han *et al.*, Identification of SARS-CoV-2 inhibitors using lung and colonic organoids. *Nature* **589**, 270-275 (2021).
143. J. Nie *et al.*, Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. *Emerg Microbes Infect* **9**, 680-686 (2020).
144. M. A. Whitt, Generation of VSV pseudotypes using recombinant DeltaG-VSV for studies on virus entry, identification of entry inhibitors, and immune responses to vaccines. *J Virol Methods* **169**, 365-374 (2010).
145. B. J. DeKosky *et al.*, In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med* **21**, 86-91 (2015).
146. C. E. Lagerman *et al.*, Ultrasonically-guided flow focusing generates precise emulsion droplets for high-throughput single cell analyses. *J Biosci Bioeng* **128**, 226-233 (2019).

147. T. A. Whitehead *et al.*, Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat Biotechnol* **30**, 543-548 (2012).
148. J. Ye, N. Ma, T. L. Madden, J. M. Ostell, IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* **41**, W34-40 (2013).
149. E. T. Boder, K. D. Wittrup, Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* **15**, 553-557 (1997).
150. G. Chao *et al.*, Isolating and engineering human antibodies using yeast surface display. *Nat Protoc* **1**, 755-768 (2006).
151. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012).
152. C. Suloway *et al.*, Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005).
153. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
154. E. A. Sickmier *et al.*, The Panitumumab EGFR Complex Reveals a Binding Mechanism That Overcomes Cetuximab Induced Resistance. *PLoS One* **11**, e0163366 (2016).
155. P. D. Adams *et al.*, Recent developments in the PHENIX software for automated crystallographic structure determination. *J Synchrotron Radiat* **11**, 53-55 (2004).
156. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
157. B. A. Barad *et al.*, EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* **12**, 943-946 (2015).
158. I. W. Davis, L. W. Murray, J. S. Richardson, D. C. Richardson, MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32**, W615-619 (2004).
159. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
160. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021).
161. D. R. Burton, What Are the Most Powerful Immunogen Design Vaccine Strategies? Reverse Vaccinology 2.0 Shows Great Promise. *Cold Spring Harb Perspect Biol* **9**, (2017).
162. S. A. Plotkin, Correlates of protection induced by vaccination. *Clin Vaccine Immunol* **17**, 1055-1065 (2010).
163. R. Rappuoli, M. J. Bottomley, U. D'Oro, O. Finco, E. De Gregorio, Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *J Exp Med* **213**, 469-481 (2016).
164. N. K. Kisalu *et al.*, A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite. *Nat Med* **24**, 408-416 (2018).
165. "World Malaria Report 2018," (World Health Organization, 2018).
166. J.-P. Julien, H. Wardemann, Antibodies against Plasmodium falciparum malaria at the molecular level. *Nature Reviews Immunology* **19**, 761-775 (2019).
167. N. K. Kisalu *et al.*, A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite. *Nature Medicine*, (2018).
168. J. Tan *et al.*, A public antibody lineage that potently inhibits malaria infection through dual binding to the circumsporozoite protein. *Nature Medicine* **24**, 401-407 (2018).

169. L. T. Wang *et al.*, A Potent Anti-Malarial Human Monoclonal Antibody Targets Circumsporozoite Protein Minor Repeats and Neutralizes Sporozoites in the Liver. *Immunity* **53**, 733-744.e738 (2020).
170. M. R. Gaudinski *et al.*, A Monoclonal Antibody for Malaria Prevention. *N Engl J Med* **385**, 803-814 (2021).
171. B. Madan *et al.*, Mutational fitness landscapes reveal genetic and structural improvement pathways for a vaccine-elicited HIV-1 broadly neutralizing antibody. *Proceedings of the National Academy of Sciences* **118**, e2011653118 (2021).
172. A. V. Medina-Cucurella, Y. Zhu, S. J. Bowen, L. M. Bergeron, T. A. Whitehead, Pro region engineering of nerve growth factor by deep mutational scanning enables a yeast platform for conformational epitope mapping of anti-NGF monoclonal antibodies. *Biotechnology and Bioengineering* **115**, 1925-1937 (2018).
173. J. A. Van Deventer, K. D. Wittrup, in *Monoclonal Antibodies,* V. Ossipow, N. Fischer, Eds. (Humana Press, Totowa, NJ, 2014), vol. 1131, pp. 151-181.
174. R. M. Adams, T. Mora, A. M. Walczak, J. B. Kinney, Measuring the sequence-affinity landscape of antibodies with massively parallel titration curves. *eLife* **5**, e23156 (2016).
175. K. K. Chan *et al.*, Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **369**, 1261-1265 (2020).
176. A. Cassotta *et al.*, A single T cell epitope drives the neutralizing anti-drug antibody response to natalizumab in multiple sclerosis patients. *Nature Medicine* **25**, 1402-1407 (2019).
177. V. Jawa *et al.*, T-Cell Dependent Immunogenicity of Protein Therapeutics Pre-clinical Assessment and Mitigation–Updated Consensus and Review 2020. *Frontiers in Immunology* **11**, (2020).
178. A. Vaisman-Mentesh, M. Gutierrez-Gonzalez, B. J. DeKosky, Y. Wine, The Molecular Mechanisms That Underlie the Immune Biology of Anti-drug Antibody Formation Following Treatment With Monoclonal Antibodies. *Frontiers in Immunology* **11**, (2020).
179. E. E. Wrenbeck *et al.*, Plasmid-based one-pot saturation mutagenesis. *Nature Methods* **13**, 928-930 (2016).
180. B. Wang *et al.*, Functional interrogation and mining of natively paired human V_H:V_L antibody repertoires. *Nature Biotechnology* **36**, 152-155 (2018).
181. A. J. Meyer, J. W. Ellefson, A. D. Ellington, Library Generation by Gene Shuffling. *Current Protocols in Molecular Biology* **105**, 15.12.11-15.12.17 (2014).
182. R. Murugan *et al.*, Evolution of protective human antibodies against Plasmodium falciparum circumsporozoite protein repeat motifs. *bioRxiv*, 798769 (2019).
183. K. Imkeller *et al.*, Antihomotypic affinity maturation improves human B cell responses against a repetitive epitope. *Science* **360**, 1358-1362 (2018).
184. D. Oyen *et al.*, Cryo-EM structure of P. falciparum circumsporozoite protein with a vaccine-elicited antibody is stabilized by somatically mutated inter-Fab contacts. *Science Advances* **4**, eaau8529 (2018).
185. D. Oyen *et al.*, Structure and mechanism of monoclonal antibody binding to the junctional epitope of Plasmodium falciparum circumsporozoite protein. *PLOS Pathogens* **16**, e1008373 (2020).
186. T. Pholcharee *et al.*, Diverse Antibody Responses to Conserved Structural Motifs in Plasmodium falciparum Circumsporozoite Protein. *J Mol Biol* **432**, 1048-1063 (2020).

187. C. A. Schramm, D. C. Douek, Beyond Hot Spots: Biases in Antibody Somatic Hypermutation and Implications for Vaccine Design. *Frontiers in Immunology* **9**, (2018).
188. Z. Sheng *et al.*, Gene-Specific Substitution Profiles Describe the Types and Frequencies of Amino Acid Changes during Antibody Somatic Hypermutation. *Frontiers in Immunology* **8**, (2017).
189. P. Dosenovic *et al.*, Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. *Cell* **161**, 1505-1515 (2015).
190. M. Tian *et al.*, Induction of HIV Neutralizing Antibody Lineages in Mice with Diverse Precursor Repertoires. *Cell* **166**, 1471-1484.e1418 (2016).
191. X. Wang *et al.*, Multiplexed CRISPR/CAS9-mediated engineering of pre-clinical mouse models bearing native human B cell receptors. *Embo j* **40**, e105926 (2021).
192. N. Kose *et al.*, A lipid-encapsulated mRNA encoding a potently neutralizing human monoclonal antibody protects against chikungunya infection. *Science Immunology* **4**, (2019).
193. F. P. Polack *et al.*, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **0**, null (2020).
194. A. B. Balazs *et al.*, Vectored immunoprophylaxis protects humanized mice from mucosal HIV transmission. *Nature Medicine* **20**, 296-300 (2014).
195. W. Zhan, M. Muhuri, P. W. L. Tai, G. Gao, Vectored Immunotherapeutics for Infectious Diseases: Can rAAVs Be The Game Changers for Fighting Transmissible Pathogens? *Frontiers in Immunology* **12**, (2021).
196. A. V. Medina-Cucurella *et al.*, User-defined single pot mutagenesis using unamplified oligo pools. *Protein Eng Des Sel* **32**, 41-45 (2019).
197. B. J. DeKosky *et al.*, Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proc Natl Acad Sci U S A* **113**, E2636-2645 (2016).
198. J. Lee *et al.*, Molecular-level analysis of the serum antibody repertoire in young adults before and after seasonal influenza vaccination. *Nat Med* **22**, 1456-1464 (2016).
199. X. Brochet, M. P. Lefranc, V. Giudicelli, IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* **36**, W503-508 (2008).
200. R. Kong *et al.*, Antibody Lineages with Vaccine-Induced Antigen-Binding Hotspots Develop Broad HIV Neutralization. *Cell* **178**, 567-584.e519 (2019).
201. Y. Flores-Garcia *et al.*, Optimization of an in vivo model to study immunity to Plasmodium falciparum pre-erythrocytic stages. *Malar J* **18**, 426 (2019).
202. E. Freire, A. Schon, A. Velazquez-Campoy, Isothermal titration calorimetry: general formalism using binding polynomials. *Methods Enzymol* **455**, 127-155 (2009).
203. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
204. N. R. Voss, C. K. Yoshioka, M. Radermacher, C. S. Potter, B. Carragher, DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J Struct Biol* **166**, 205-213 (2009).
205. G. C. Lander *et al.*, Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J Struct Biol* **166**, 95-102 (2009).

206. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
207. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015).
208. K. Zhang, Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12 (2016).
209. R. Kong *et al.*, Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* **352**, 828-833 (2016).
210. E. S. Gray *et al.*, 4E10-resistant variants in a human immunodeficiency virus type 1 subtype C-infected individual with an anti-membrane-proximal external region-neutralizing antibody response. *J Virol* **82**, 2367-2375 (2008).
211. S. I. Richardson, P. L. Moore, The antibody response in HIV-1-infected donors. *Curr Opin HIV AIDS* **14**, 233-239 (2019).
212. X. Wu *et al.*, Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856-861 (2010).
213. J. F. Scheid *et al.*, Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633-1637 (2011).
214. X. Wu *et al.*, Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* **333**, 1593-1602 (2011).
215. L. M. Walker *et al.*, Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285-289 (2009).
216. J. S. McLellan *et al.*, Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336-343 (2011).
217. B. S. Briney, J. R. Willis, J. E. Crowe, Jr., Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. *PLoS One* **7**, e36750 (2012).
218. N. A. Doria-Rose *et al.*, Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**, 55-62 (2014).
219. R. Andrabi *et al.*, Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. *Immunity* **43**, 959-973 (2015).
220. J. Gorman *et al.*, Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. *Nat Struct Mol Biol* **23**, 81-90 (2016).
221. R. Pejchal *et al.*, A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science* **334**, 1097-1103 (2011).
222. L. M. Walker *et al.*, Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466-470 (2011).
223. L. Kong *et al.*, Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat Struct Mol Biol* **20**, 796-803 (2013).
224. F. Garces *et al.*, Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* **43**, 1053-1063 (2015).
225. T. Muster *et al.*, A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* **67**, 6642-6647 (1993).
226. T. Muster *et al.*, Cross-neutralizing activity against divergent human immunodeficiency virus type 1 isolates induced by the gp41 sequence ELDKWAS. *J Virol* **68**, 4031-4034 (1994).

227. G. Stiegler *et al.*, A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res Hum Retroviruses* **17**, 1757-1765 (2001).
228. G. Ofek *et al.*, Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol* **78**, 10724-10737 (2004).
229. B. F. Haynes, M. A. Moody, L. Verkoczy, G. Kelsoe, S. M. Alam, Antibody polyspecificity and neutralization of HIV-1: a hypothesis. *Hum Antibodies* **14**, 59-67 (2005).
230. G. Ofek *et al.*, Relationship between antibody 2F5 neutralization of HIV-1 and hydrophobicity of its heavy chain third complementarity-determining region. *J Virol* **84**, 2955-2962 (2010).
231. J. Huang *et al.*, Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491**, 406-412 (2012).
232. A. Irimia, A. Sarkar, R. L. Stanfield, I. A. Wilson, Crystallographic Identification of Lipid as an Integral Component of the Epitope of HIV Broadly Neutralizing Antibody 4E10. *Immunity* **44**, 21-31 (2016).
233. A. R. Corrigan *et al.*, Fusion peptide priming reduces immune responses to HIV-1 envelope trimer base. *Cell Rep* **35**, 108937 (2021).
234. K. Xu *et al.*, Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1. *Nat Med* **24**, 857-867 (2018).
235. R. Kong *et al.*, Antibody Lineages with Vaccine-Induced Antigen-Binding Hotspots Develop Broad HIV Neutralization. *Cell* **178**, 567-584 e519 (2019).
236. S. Kumar *et al.*, Capturing the inherent structural dynamics of the HIV-1 envelope glycoprotein fusion peptide. *Nat Commun* **10**, 763 (2019).
237. M. Yuan *et al.*, Conformational Plasticity in the HIV-1 Fusion Peptide Facilitates Recognition by Broadly Neutralizing Antibodies. *Cell Host Microbe* **25**, 873-883 e875 (2019).
238. R. M. Adams, T. Mora, A. M. Walczak, J. B. Kinney, Measuring the sequence-affinity landscape of antibodies with massively parallel titration curves. *Elife* **5**, (2016).
239. A. V. Medina-Cucurella, Y. Zhu, S. J. Bowen, L. M. Bergeron, T. A. Whitehead, Pro region engineering of nerve growth factor by deep mutational scanning enables a yeast platform for conformational epitope mapping of anti-NGF monoclonal antibodies. *Biotechnol Bioeng* **115**, 1925-1937 (2018).
240. N. Mehta *et al.*, An engineered antibody binds a distinct epitope and is a potent inhibitor of murine and human VISTA. *Sci Rep* **10**, 15171 (2020).
241. A. S. Dingens *et al.*, Complete functional mapping of infection- and vaccine-elicited antibodies against the fusion peptide of HIV. *PLoS Pathog* **14**, e1007159 (2018).
242. S. Conti *et al.*, Design of immunogens to elicit broadly neutralizing antibodies against HIV targeting the CD4 binding site. *Proc Natl Acad Sci U S A* **118**, (2021).
243. C. H. Shen *et al.*, VRC34-Antibody Lineage Development Reveals How a Required Rare Mutation Shapes the Maturation of a Broad HIV-Neutralizing Lineage. *Cell Host Microbe* **27**, 531-543 e536 (2020).
244. J. R. Mascola, B. F. Haynes, HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunol Rev* **254**, 225-244 (2013).

245. M. Caskey, F. Klein, M. C. Nussenzweig, Broadly neutralizing anti-HIV-1 monoclonal antibodies in the clinic. *Nat Med* **25**, 547-553 (2019).
246. F. Benjelloun, P. Lawrence, B. Verrier, C. Genin, S. Paul, Role of human immunodeficiency virus type 1 envelope structure in the induction of broadly neutralizing antibodies. *J Virol* **86**, 13152-13163 (2012).
247. Y. Liu, W. Cao, M. Sun, T. Li, Broadly neutralizing antibodies for HIV-1: efficacies, challenges and opportunities. *Emerg Microbes Infect* **9**, 194-206 (2020).
248. A. Pegu *et al.*, Potent anti-viral activity of a trispecific HIV neutralizing antibody in SHIV-infected monkeys. *Cell Rep* **38**, 110199 (2022).
249. A. B. Balazs *et al.*, Vectored immunoprophylaxis protects humanized mice from mucosal HIV transmission. *Nat Med* **20**, 296-300 (2014).
250. T. Schlake *et al.*, mRNA: A Novel Avenue to Antibody Therapy? *Mol Ther* **27**, 773-784 (2019).
251. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).
252. M. Sarzotti-Kelsoe *et al.*, Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J Immunol Methods* **409**, 131-146 (2014).

8. Appendices

Appendix A: Figure 14. Supplemental Chapter 2.

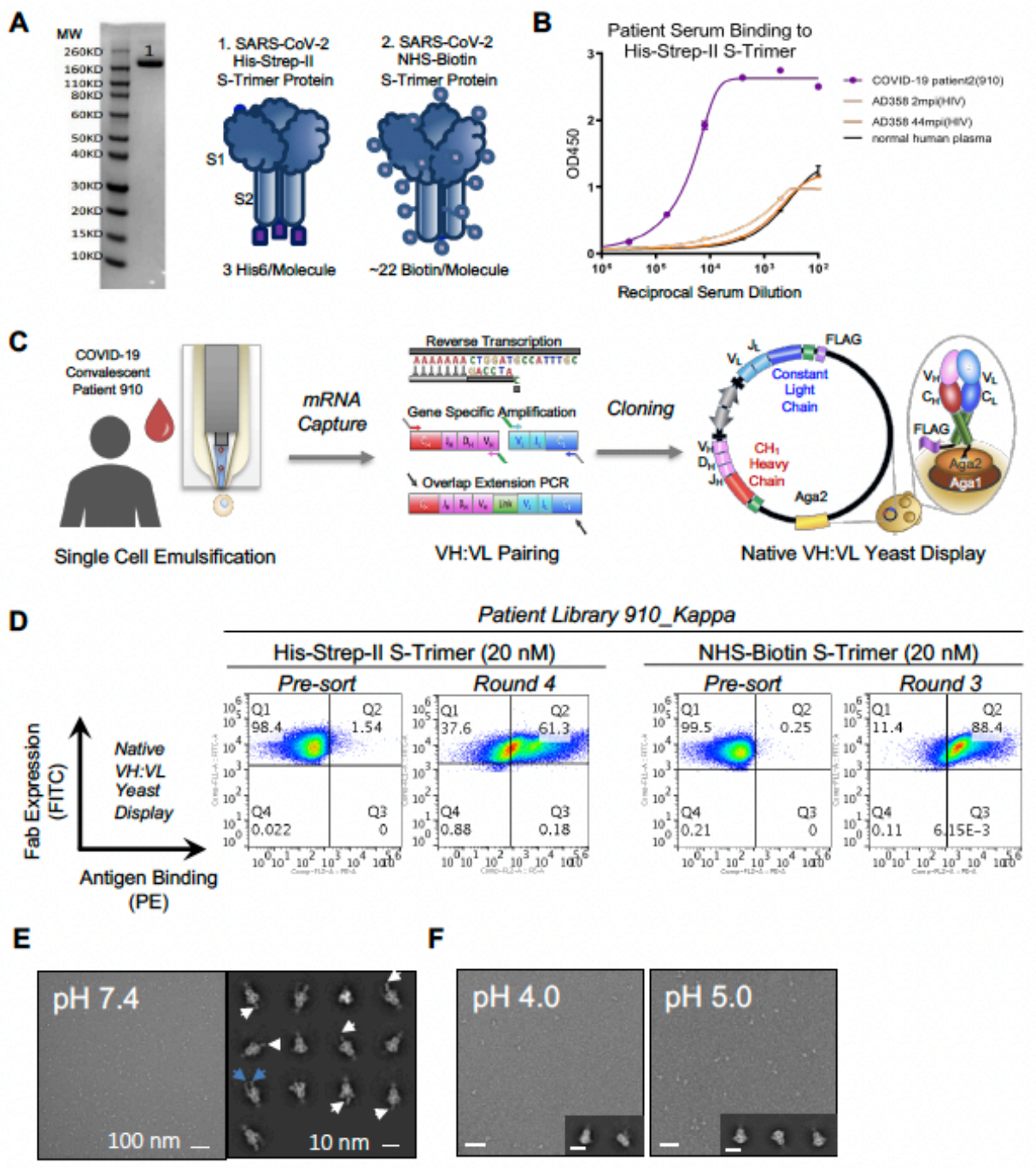


Figure 23. Supplemental Chapter 2. Overview of 910-30 discovery from a convalescent COVID-19 patient utilizing natively paired antibody fragment yeast display, FACS bio-panning, and soluble characterization related to Figure 10 and Chapter 2 STAR Methods.

Fig. 14. Supplemental Chapter 2. Overview of 910-30 discovery from a convalescent COVID-19 patient utilizing natively paired antibody fragment yeast display, FACS bio-panning, and soluble characterization related to Figure 10 and Chapter 2 STAR Methods. (A) Reduced SDS-PAGE gel shows SARS-CoV-2 His-Strep-II S-Trimer monomer protein at approximately 142 kDa. Schematics highlight unique features of SARS-CoV-2 Spike S2P antigen probes used for FACS bio panning of antibody yeast display libraries from a COVID-19 convalescent donor. **(B)** Hong Kong University convalescent Donor 910 serum showed strong binding to SARS-CoV-2 His-Strep-II S-Trimer protein compared to controls. Data are represented as mean ± SEM. **(C)** Workflow overview used to generate native VH:VL libraries from the COVID-19 convalescent donor HKU910 for functional antibody screening using yeast display. **(D)** Donor-derived antibody library bio-panning via FACS shows significant library enrichment after multiple rounds of sorting. *Left* Donor 910 pre-sort yeast library vs. sorted yeast library for His-S-Trimer antigen. *Right* Donor 910 pre-sort yeast library vs. sorted yeast library for Biotin-S-Trimer antigen. **(E)** Negative-staining electron microscopy at pH 7.4 resolved complexes between SARS-CoV-2 S2P and 910-30 Fab. Left: representative micrograph; right: representative 2D class averages. White arrows point to Fab fragments in complexes formed between one Fab and one spike trimer; blue arrows point to Fab fragments in complexes formed between two Fab fragments and one spike trimer. **(F)** Negative-staining electron microscopy at pH 4.0, and 5.0 reveal no 910-30 Fab bound to SARS-CoV-2 S2P protein at given pH values. Representative micrographs are shown. Insets show representative 2D class averages. Scale bars: 50 nm (micrographs), 20 nm (2D class averages).

Appendix B: Figure 15. Supplemental Chapter 2.

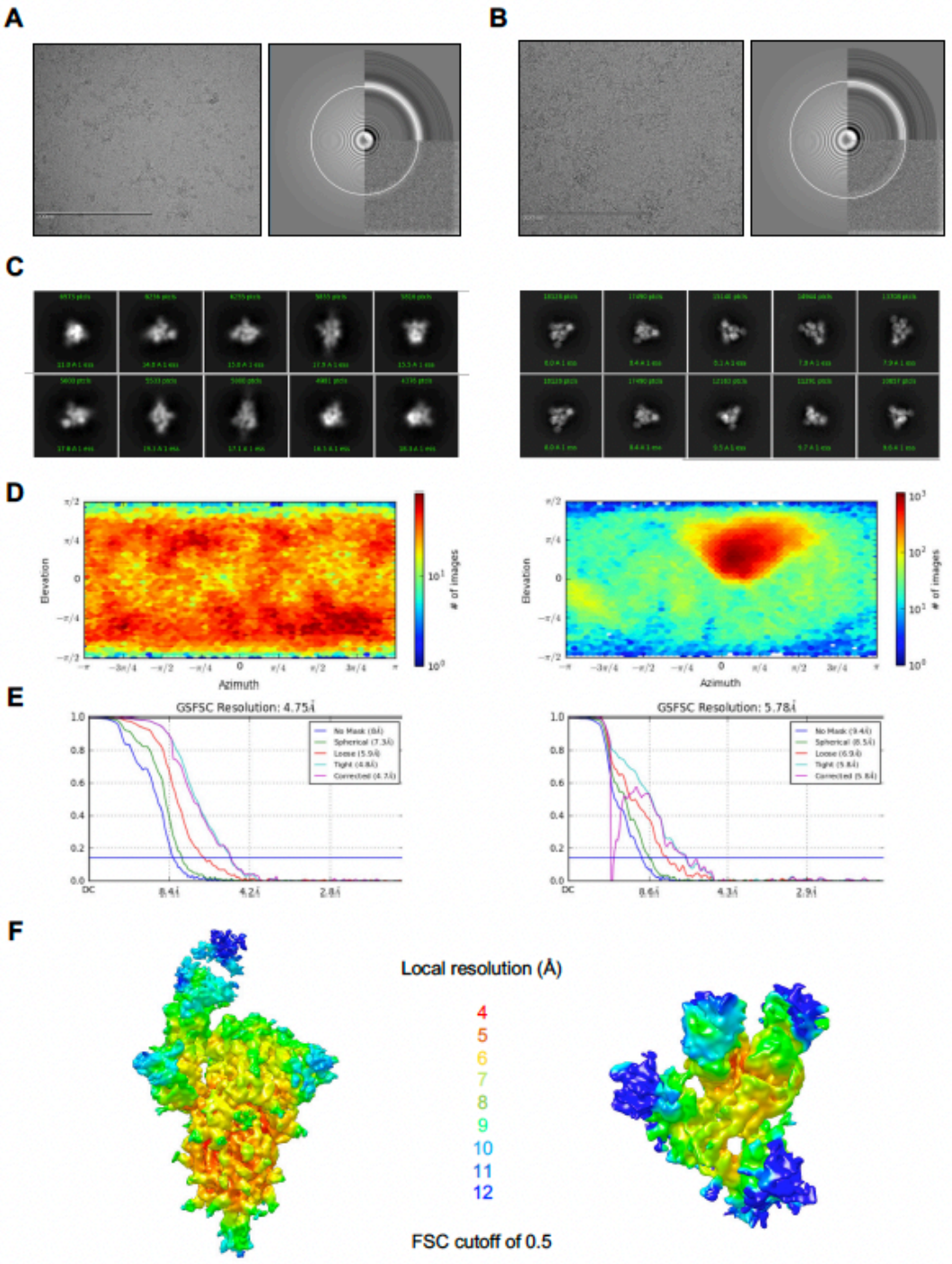


Figure 24. Supplemental Chapter 2. Cryo-EM analysis of 910-30 Fab in complex with SARS-CoV-2 spike at pH 5.5. Sample 1 obtained mixing 910-30 Fab and spike in a 1:1 molar ratio, sample 2 obtained mixing 910-30 Fab and spike in a 9:1 molar ratio. Related to Fig10.

Fig. 15. Supplemental Chapter 2. Cryo-EM analysis of 910-30 Fab in complex with SARS-CoV-2 spike at pH 5.5. Sample 1 obtained mixing 910-30 Fab and spike in a 1:1 molar ratio, sample 2 obtained mixing 910-30 Fab and spike in a 9:1 molar ratio. Related to Figures 10C and 10D. (A) Representative micrograph and CTF of the micrograph for sample 1. **(B)** Representative micrograph and CTF of the micrograph for sample 2. **(C)** Representative 2D classes for sample 1 (left) and sample 2 (right). **(D)** The orientations of all particles used in the final refinement are shown as a heatmap for sample 1 (left) and sample 2 (right). **(E)** The gold-standard Fourier shell correlation resulted in a resolution of 4.75 Å for sample 1 (left) and 5.78 Å for sample 2 (right). **(F)** The local resolution of the two final maps are shown generated through cryoSPARC using an FSC cutoff of 0.5; left: sample 1, right: sample 2.

Appendix C: Figure 16. Supplemental Chapter 2.

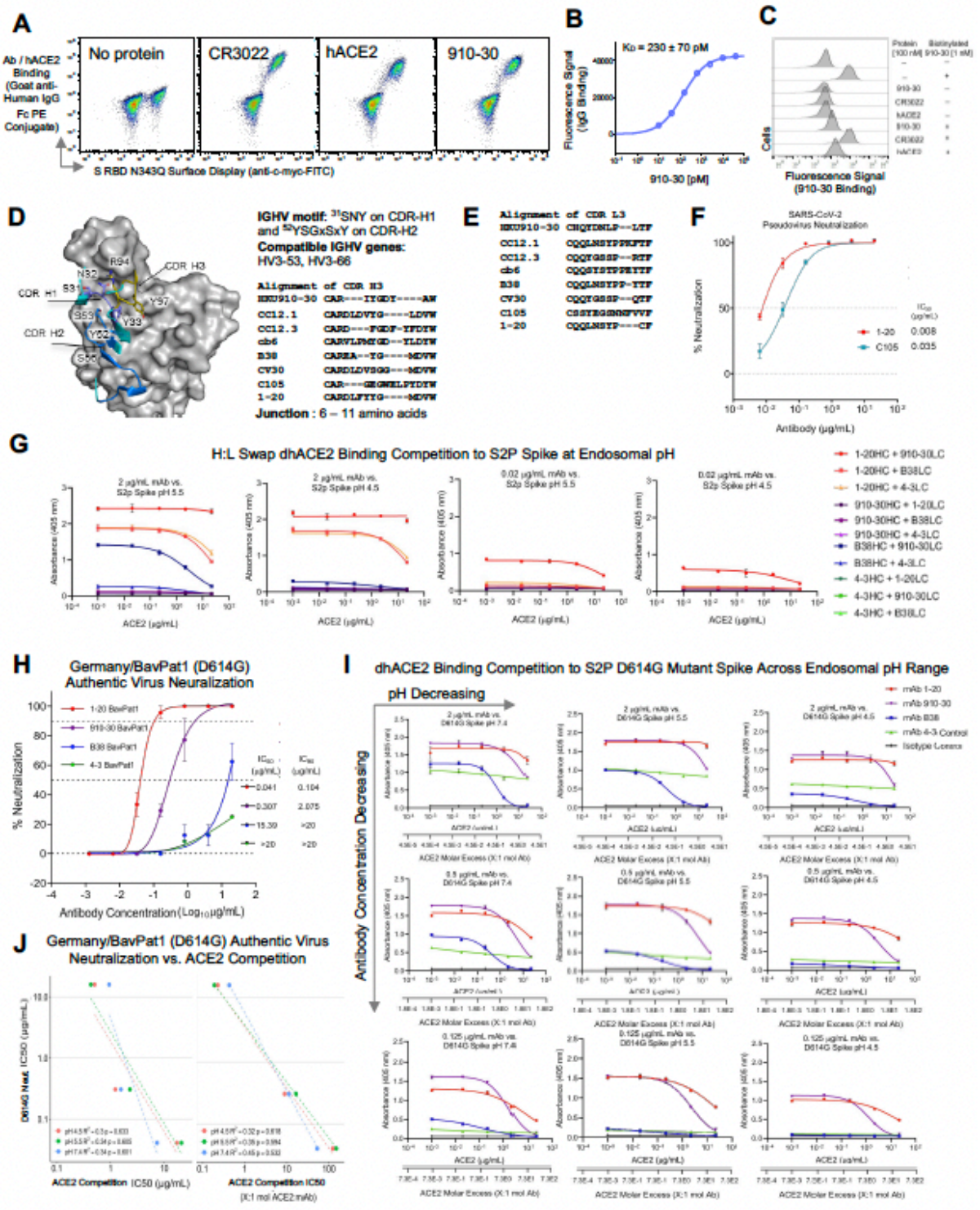


Figure 25. Supplemental Chapter 2. IGHV3-53/3-66 class member extended characterization and biophysical analysis related to Figures 11, 12, and 13.

Fig. 16. Supplemental Chapter 2. IGHV3-53/3-66 class member extended characterization and biophysical analysis related to Figures 11, 12, and 13. **(A)** Yeast-displayed aglycosylated RBD demonstrates 910-30 recognition is glycan-independent. S RBD N343Q with a C-terminal myc epitope tag was displayed on the surface of yeast and labeled with no protein or 1 nM of CR3022, human ACE2-Fc (hACE2), or 910-30. Cells were washed, secondarily labeled with anti-c-myc-FITC and Goat anti-Human IgG Fc PE conjugate, and read on a Sony SH800 cell sorter. Biological replicates were performed on two different days. **(B)** Yeast cell surface titrations of 910-30 IgG against aglycosylated S RBD yield a K_D of 230 ± 38 pM. Technical triplicates were performed for two biological replicates (n = 6), and error reported is 2 s.e.m. **(C)** Yeast-displayed RBD competition binding experiments of free 910-30, hACE2 and CR3022 vs. biotinylated or unbiotinylated 910-30. Technical triplicates were performed for two biological replicates (n = 6). **(D)** Heavy chain genetic elements associated with the IGHV3-53/3-66 antibody class. **(E)** Light chain CDR3 alignment of IGHV3-53/3-66 antibody class. **(F)** Lambda chain IGHV3-53/3-66 class member C105 shows moderate neutralizing capacity compared to potent kappa chain IGHV3-53/3-66 class member 1-20. Data are represented as mean ± SEM. **(G)** pH mediated dhACE2 competition measured by ELISA showing constant concentrations of heavy-light-swapped IgG binding to SARS-CoV-2 S2P protein versus increasing dhACE2 (ACE2) concentrations. Potently neutralizing heavy-light swap variants (Fig. 12A) show higher affinity binding to S2P spike and stronger ACE2 competition relative to less potent Abs. Data are represented as mean ± SEM. **(H)** 1-20, 910-30, and B38 show equivalent neutralization in a D614G authentic virus assay as for D614 authentic virus (Fig. 11D), with 4-3 included as a gene-matched control. Data are represented as mean ± SEM. **(I)** SARS-CoV-2 D614G S2P protein mutant variant pH mediated dhACE2 (ACE2) competition measured by ELISA showing constant concentrations of heavy-light-swapped IgG versus increasing dhACE2 concentrations. Potently neutralizing heavy-light swap variants show higher affinity binding to D614G S2P mutant spike and stronger ACE2 competition relative to less potent Abs. The concentration of dhACE2 required to outcompete antibody binding to spike is given as both μg/mL and as ACE2 molar excess units. Data are represented as mean ± SEM. **(J)** D614G authentic

virus neutralization potency and dhACE2 competition IC_{50} show a correlation between potent neutralization and stronger ACE2 competition.

Appendix D: Figure 17. Supplemental Chapter 2.

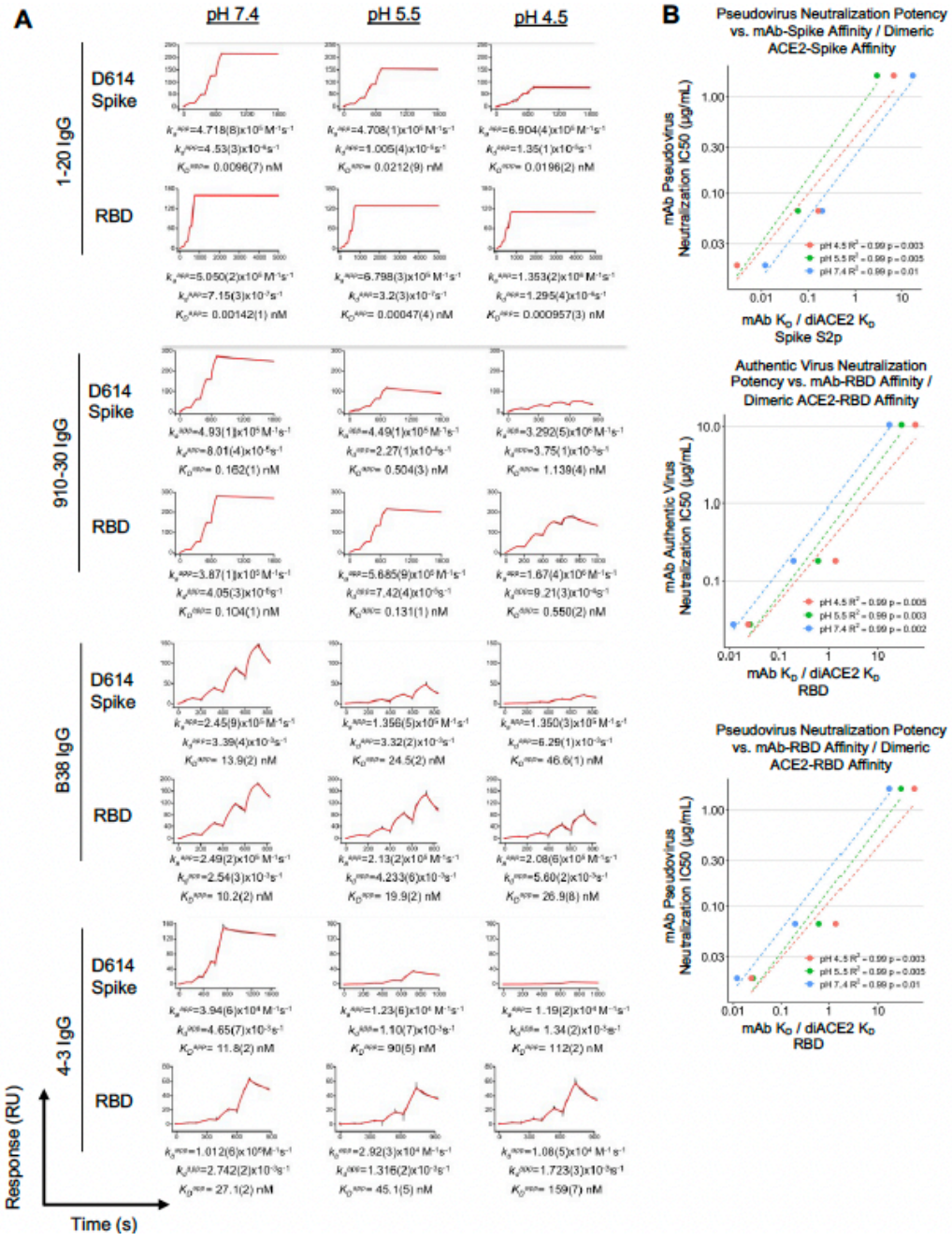


Figure 26. Supplemental Chapter 2. Extended binding and neutralization analysis across multiple pH values related to Figures 13C and 13D.

Fig. 17. Supplemental Chapter 2. Extended binding and neutralization analysis across multiple pH values related to Figures 13C and 13D. (A) pH mediated SPR single-cycle kinetic experiments for 910-30, B38, 4-3, and 1-20, for IgG binding to biotinylated spike (top row) and to biotinylated-RBD (bottom row) in each of the four panels. Black traces represent the experimental data and red traces represent the fit to a 1:1 interaction model. The number in brackets represents the error of the fit in the last significant digit. **(B)** Correlations between both authentic and pseudovirus neutralization vs. the ratio of antibody IgG affinity to RBD or Spike divided by dimeric ACE2 affinity to RBD or Spike.

Appendix E: Table 1. Supplemental Chapter 2.

Table 1. Supplemental Chapter 2. Cryo-Em collection and refinement statistics for 910-30 Fab in complex with SARS-CoV-2 spike at pH 5.5 related to Figure 10D

	SARS-CoV-2 spike in complex with 910-30 Fab at pH 5.5 (folded spike)	SARS-CoV-2 spike in complex with 910-30 Fab at pH 5.5 (disrupted spike)
EMDB ID	EMD-23016	EMD-23039
PDB ID	7KS9	
Data Collection		
Microscope	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300
Electron dose ($e^-/Å^2$)	41.92	41.92
Detector	Gatan K3 BioQuantum	Gatan K3 BioQuantum
Pixel Size (Å)	1.07	1.07
Defocus Range (μm)	-0.8/-2.5	-0.8/-2.5
Magnification	81000	81000
Reconstruction		
Software	cryoSPARC v2.15	cryoSPARC v2.15
Particles	88,315	188,269
Symmetry	C1	C1
Box size (pix)	400	400
Resolution (Å) ($FSC_{0.143}$)	4.75	5.78
Refinement		
Software	Phenix 1.18	
Protein residues	3189	
Chimera CC	0.88	
EMRinger Score	0.55	
R.m.s. deviations		
Bond lengths (Å)	0.003	
Bond angles (°)	0.711	
Validation		
Molprobity score	1.15	
Clash score	3.59	
Favored rotamers (%)	100	
Ramachandran		
Favored regions (%)	98.0	
Allowed regions (%)	2.0	
Disallowed regions (%)	0	

Appendix F: Table 2. Supplemental Chapter 2.

Table 2. Supplemental Chapter 1. List of IGHV3-53 / IGHV3-66 anti-SARS-CoV-2 antibodies in previously published articles related to Figure 11D. *Supplemental Table provided separately as an Excel file.* Hyperlink to automatic download: mmc2.xlsx

Appendix G: Table 3. Supplemental Chapter 2.

Table 3. Supplemental Chapter 2. Features of the IGHV3-53/3-66 antibodies investigated in this study related to Figures 11, 12, and 13.

		Antibody			
Feature	Units	1-20	910-30	B38	4-3
Relevant Spike S2P Affinity Values (pH 7.4)	*ka* (M^-1 s^-1)	4.718×10^5	4.39×10^5	2.45×10^5	3.94×10^4
	kd (s^-1)	4.53×10^{-6}	8.01×10^{-6}	3.39×10^{-3}	4.65×10^{-3}
	KD (nM)	0.0096	0.162	13.9	11.8
Relevant RBD Affinity Values (pH 7.4)	*ka* (M^-1 s^-1)	5.050×10^5	3.87×10^5	2.49×10^5	1.012×10^5
	kd (s^-1)	7.15×10^{-7}	4.05×10^{-6}	2.54×10^{-3}	2.742×10^{-3}
	KD (nM)	0.00142	0.104	10.2	27.1
WT (D614) Pseudo Virus Neutralization potency (µg/mL)	IC50	0.018	0.066	1.668	5.582
WT (D614) Authentic Virus Neutralization potency (µg/mL)	IC50	0.027	0.183	10.571	18.54
	IC90	0.219	1.436	>20	>20
Mutant D614G Authentic Virus Neutralization potency (µg/mL)	IC50	0.041	0.307	15.39	>20
	IC90	0.104	2.075	>20	>20
Heavy chain	IGHV Gene	IGHV3-53*01	IGHV3-53*04	IGHV3-53*04	IGHV3-66*01
	IGHJ Gene	IGHJ6*02	IGHJ5*02	IGHJ6*02	IGHJ6*01 F
	IGHD Gene	IGHD2-2*02	IGHD4-17*01	N/A	IGHD1-26*01
	IGHV identity aa	97.9%	99.0%	99.0%	99.0%
	CDR-H3 seq (aa)	CARDLFYYGMDVW	CARIYGDYAW	CAREAYGMDVW	CARDSSEGPGYYGMDVW
	CDR-H3 len (aa)	13	10	11	17
Light chain	IGKV Gene	IGKV1-9*01	IGKV1-33*01	IGKV1-9*01	IGKV1-9*01
	IGKJ Gene	IGKJ3*01	IGKJ4*01	IGKJ2*01	IGKJ4*01
	IGKV identity aa	100.0%	97.9%	97.9%	100.0%
	CDR-L3 seq (aa)	CQQLNSYPCF	CHQYDNLPLTF	CQQLNSYPPYTF	CQQLNSYLPLTF
	CDR-L3 len (aa)	10	11	12	12

Appendix H: Table 4. Supplemental Chapter 2.

Table 4. Supplemental Chapter 2. Heavy chain and light chain CDR1 and CDR 2 sequence alignment for recognition signature related to Figures 12C, Appendix C Figure 16D-E. *Supplemental Table provided separately as an Excel file.* Hyperlink to automatic download: mmc3.xlsx

Appendix I: Figure 23. Supplemental Chapter 3.

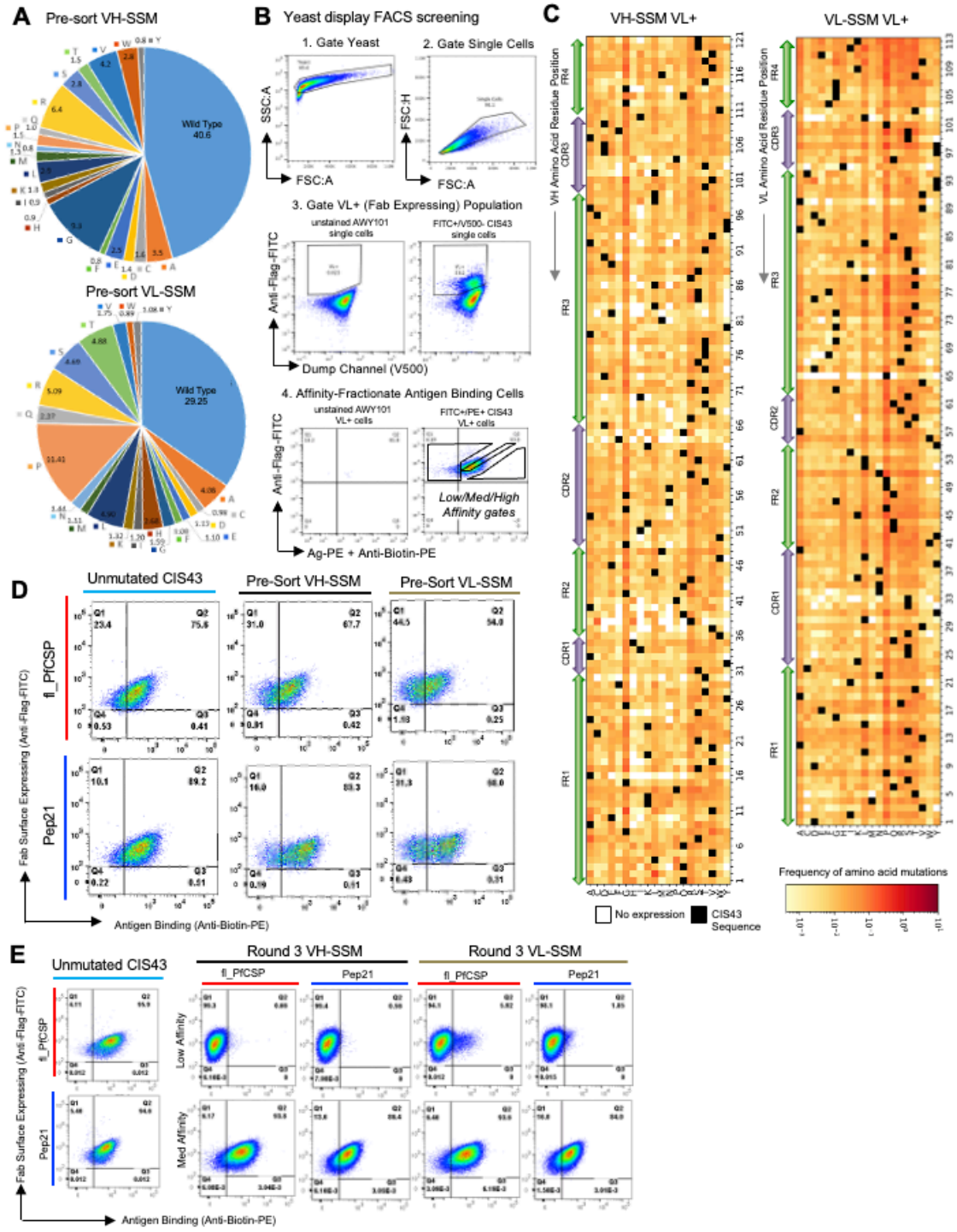


Figure 27. Supplemental Chapter 3. Single mutant library synthesis and screening revealed the sequence-function relationship between CIS43 antibody mutants and anti-malarial antigen recognition.

Fig. 23. Supplemental Chapter 3. Single mutant library synthesis and screening revealed the sequence-function relationship between CIS43 antibody mutants and anti-malarial antigen recognition.

(A) Distribution of single-mutation VH_NNK and VL_MNN DNA codons in pre-sort SSM libraries, as quantified by NGS.

(B) FACS gating strategy for affinity-based sorting of SSM libraries.

(C) CIS43 single-mutation heat maps showing library frequencies for mutants expressed in the VL+ Fab-expressing VH-SSM and VL-SSM libraries, prior to sorting for anti-malarial antigen recognition.

(D) Flow cytometric analysis of CIS43 wild type surface-displayed Fab and Pre-sort single mutation Fab libraries, stained with Pep21 and fl_PfCSP.

(E) Flow cytometric analysis of CIS43 wild type surface-displayed Fab and Round 3 single mutation Fab libraries sorted against Pep21 and fl_PfCSP for low affinity (*upper*) and medium affinity (*lower*) population phenotypes.

Appendix J: Figure 24. Supplemental Chapter 3.

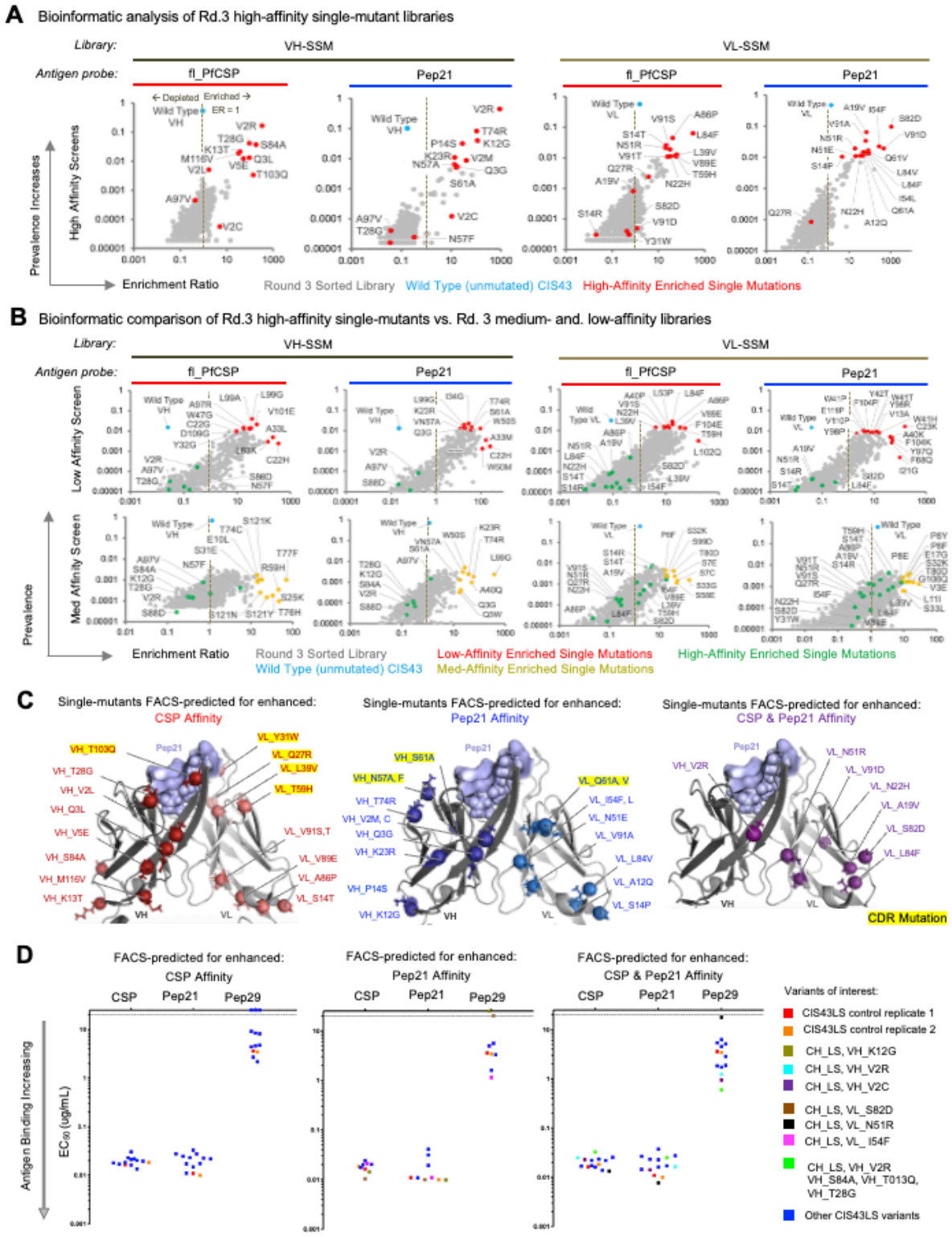


Figure 28. Supplemental Chapter 3. Single-mutation library screening overview.

Figure 24. Supplemental Chapter 3. Single-mutation library screening overview.

(A) Bioinformatic analyses of single mutation sequence enrichment ratios (ER) vs. prevalence after three rounds of FACS revealed the relationship between dominant single mutation sequences (red) versus unmutated CIS43 sequences (cyan) after high-affinity enrichment against fl_PfCSP and Pep21 antigens.

(B) Bioinformatic analyses of single mutation sequence enrichment ratios (ER) vs. prevalence after three rounds of FACS screening revealed the relationship between dominant single mutation sequences (red) versus unmutated CIS43 sequences (cyan) after medium- and low-affinity enrichment against fl_PfCSP and Pep21 antigens. Data are shown for both the low affinity (*upper*) and medium affinity (*lower*) population phenotypes.

(C) Structural mapping of predicted affinity-enhancing mutants based on single mutation NGS data.

(D) ELISA evaluation of expressed and purified single mutation IgGs.

Appendix K: Figure 25. Supplemental Chapter 3.

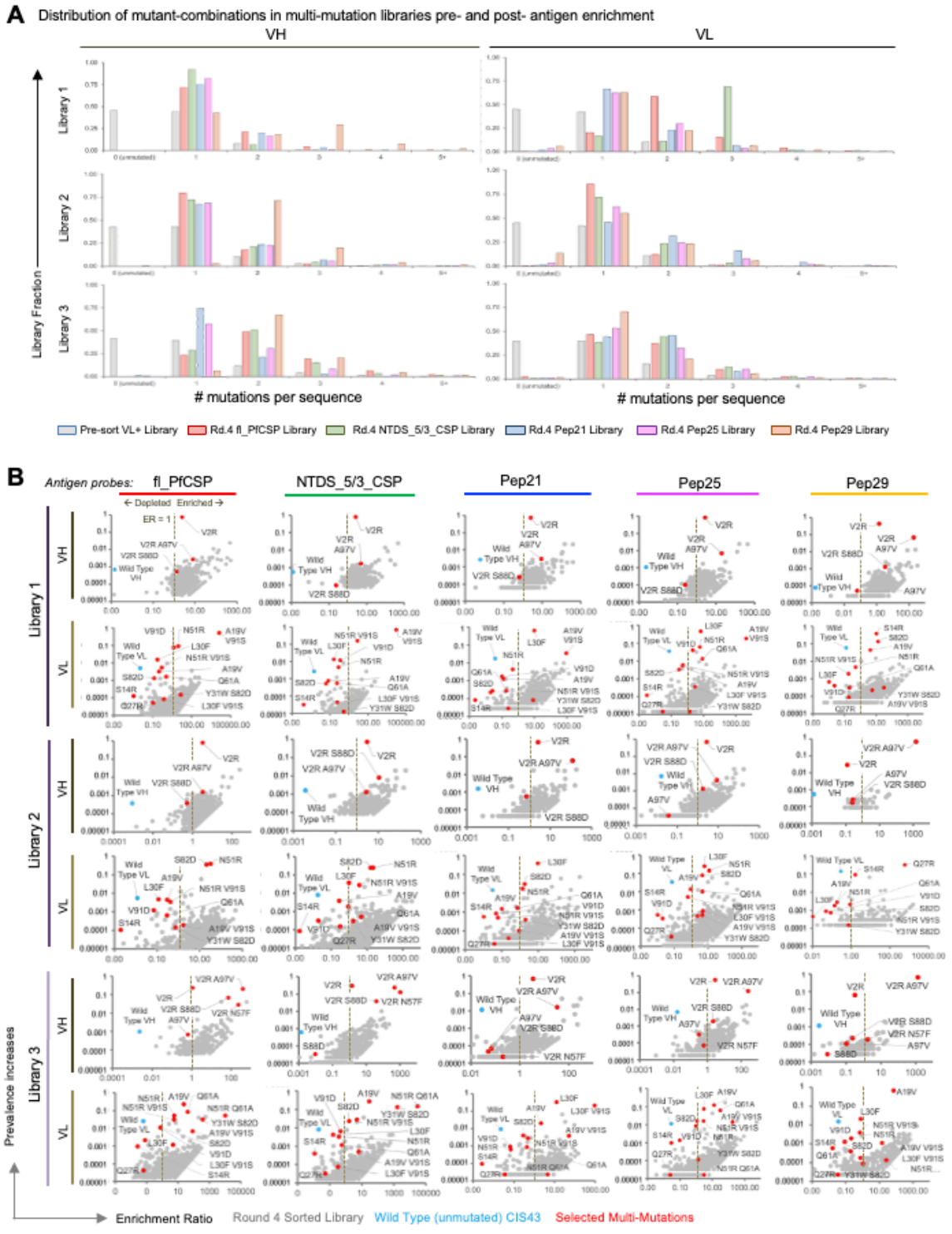


Figure 29. Supplemental Chapter 3. Multi-mutation screening enhances understanding of CIS43 sequence-function relationships.

Fig. 25. Supplemental Chapter 3. Multi-mutation screening enhances understanding of CIS43 sequence-function relationships.

(A) Multi-mutation sequence distribution in pre-sort multi-mutation libraries, and after four rounds of screening for high-affinity variants.

(B) Bioinformatic analyses of multi-mutation sequence enrichment ratios (ER) vs. prevalence after four rounds of FACS revealed the relationship between dominant mutant sequences (red) versus unmutated CIS43 sequences (cyan) after high-affinity enrichment against fl_PfCSP, NTDS_5/3_CSP, Pep21, Pep25, and Pep29 antigens.

Appendix L: Figure 26. Supplemental Chapter 3.

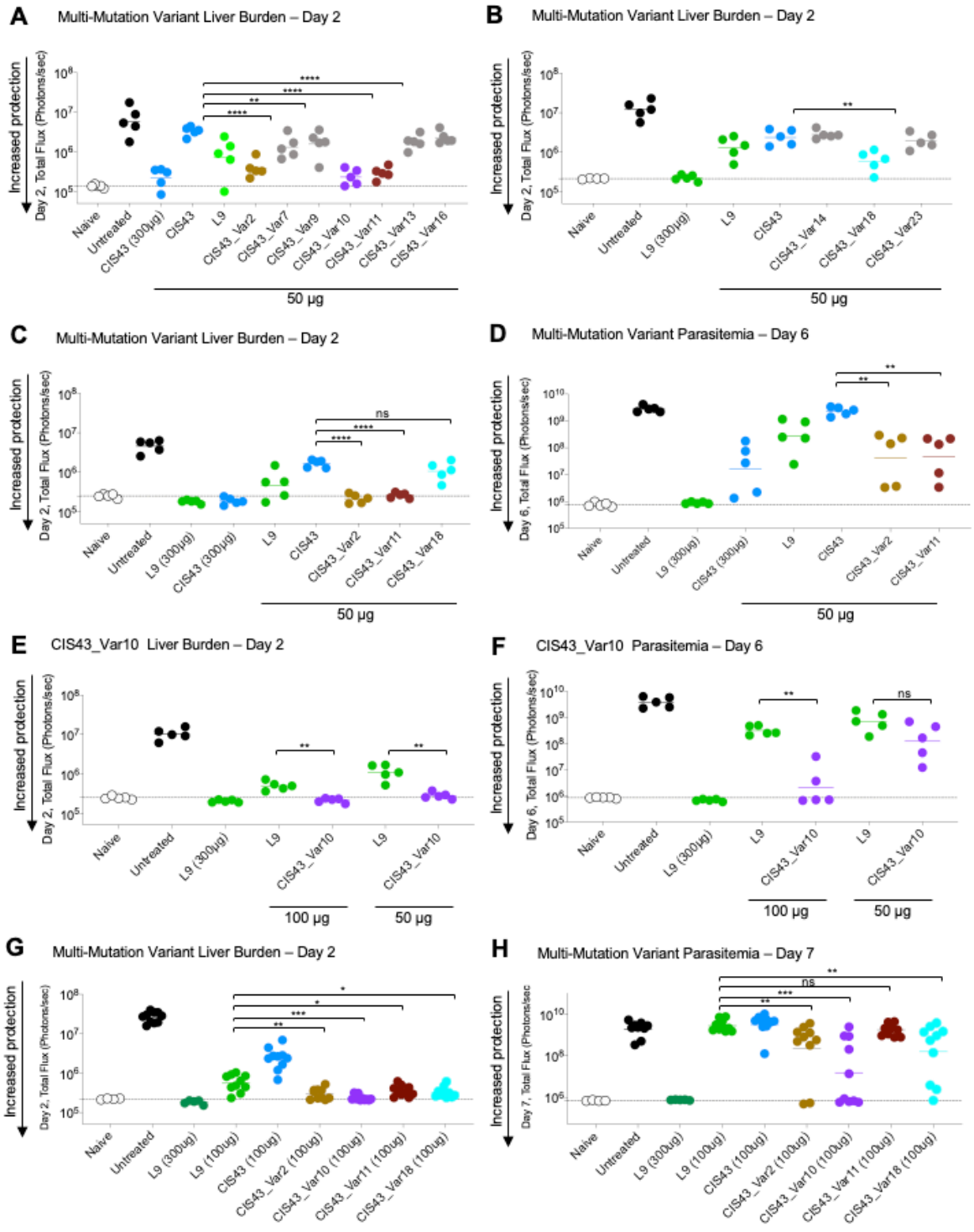


Figure 30. Supplemental Chapter 3. Experimental groups for analysis of multi-mutation variant protection against malaria challenge.

Fig. 26. Supplemental Chapter 3. Experimental groups for analysis of multi-mutation variant protection against malaria challenge.

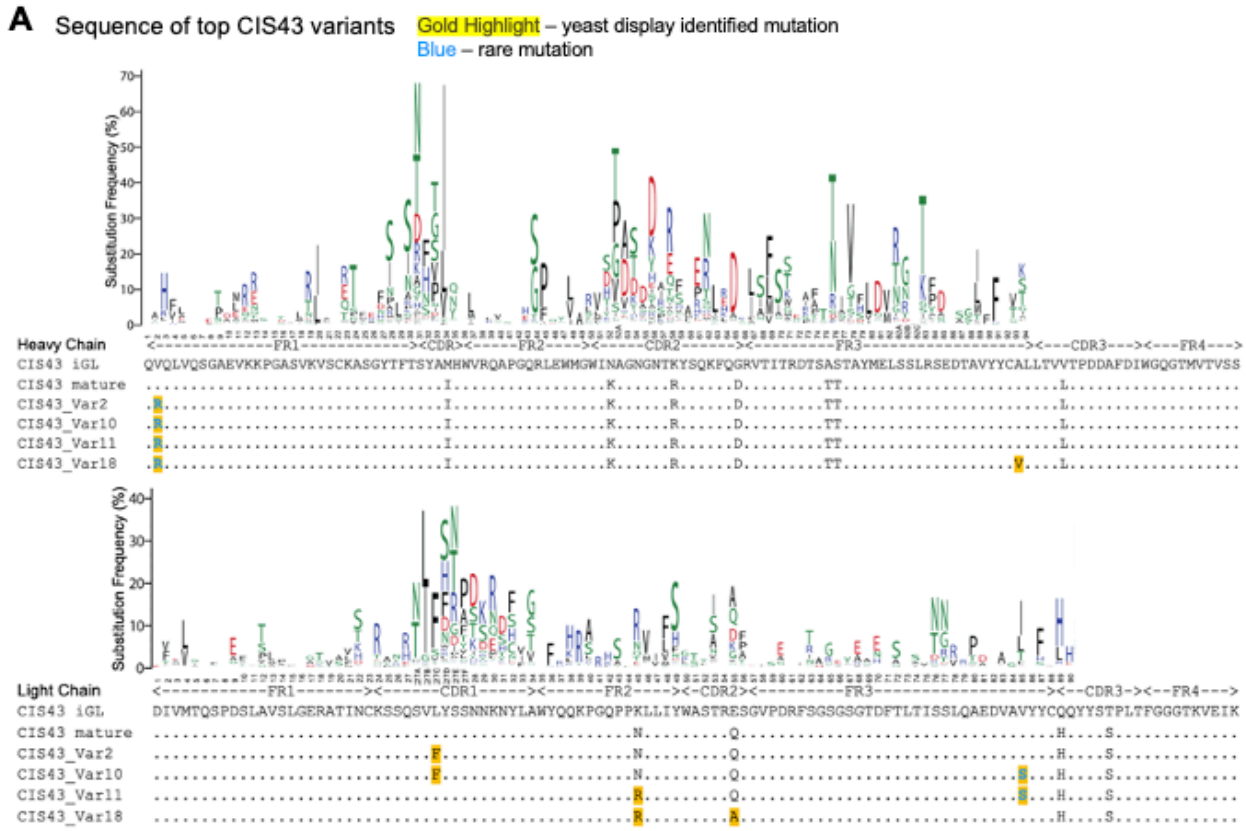
(A) and (B) Following passive transfer of CIS43, L9 and the indicated multi-mutation variants at the specified doses, Albino-B6 mice were intravenously challenged with 2,000 Pb-PfCSP sporozoites, and the liver burden was assessed at Day 2 after challenge. Total flux or bioluminescence expressed as photons/sec is a measure of the liver parasite burden and is shown for naïve (background, white) and untreated (maximum burden, black) groups. Horizontal lines indicate the geometric mean. Statistical differences were assessed using a two-tailed Mann-Whitney test. P-values are indicated by stars, with P-value < 0.05 (*), P-value < 0.01 (**), P-value < 0.001 (***), and P-value < 0.0001 (****).

(C) and **(D)** Select multi-mutation variant antibodies were re-screened at the indicated doses and total flux was measured at Day 2 for liver burden. Infected mice were euthanized, but protected mice were followed out and re-imaged again at Day 6; total flux at this timepoint reflects parasitemia or blood stage infection.

(E) CIS43_Var10 mediated protection as assessed at Day 2, liver burden, and followed to **(F)** Day 6 for Parasitemia assessment as compared to L9 at 50 and 100ug.

(G) Liver Burden, Day 2 assessment and followed to **(H)** Day 7 for Parasitemia assessment for CIS43_Var2, Var10, Var11 and Var18 at 100ug compared to CIS43 and L9.

Appendix M: Figure 27. Supplemental Chapter 3.



B Rarity of CIS43 mutations and yeast-identified mutations in human V-gene segments

VH Mutation* (region)	Frequency (%)	VL Mutation* (region)	Frequency(%)
V2R (FR1)	< 0.0001	L27cF (CDR1)	13.1
M34I (CDR1)	46.1	K45N (FR2)	2.17
N52K (CDR2)	0.94	K45R (FR2)	7.75
K58R (CDR2)	12.7	E55Q (CDR2)	4.82
G65D (CDR2)	16.7	E55A (CDR2)	5.26
A75T (FR3)	3.36	V85S (FR3)	0.13
S76T (FR3)	18.5		

* Kabat antibody numbering used for the position of mutation

HC: V2R
LC: $L27_cF$

HC: V2R
LC: $L27_cF$, V85S

HC: V2R
LC: N45R, V85S

HC: V2R, A93V
LC: N45R, Q55A

Figure 31. Supplemental Chapter 3. Location and natural frequency of mutations improving CIS43.

Fig. 27. Supplemental Chapter 3. Location and natural frequency of mutations improving CIS43.

(A) Sequences of improved CIS43 variant antibodies are shown along with the heavy (top) and light (bottom) antibody V-gene mutational profiles. We define a rare mutation as a frequency < 0.5% in human immune repertoires.

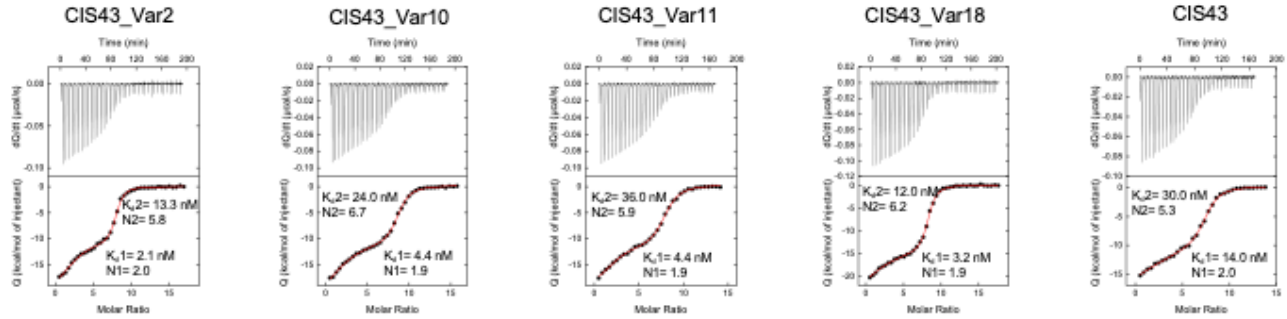
(B) Frequencies of amino acid mutations in the context of their respective germline V-gene, reported from cAb-Rep server.

Appendix N: Figure 28. Supplemental Chapter 3.

A BLI affinity of CIS43 variants

Antibody	fl_PfCSP K_{D1} (nM)	fl_PfCSP K_{D2} (μM)	NTDS_5/3_CSP K_D (nM)	NPDP19 K_D (nM)	Pep21 K_D (nM)	Pep25 K_D (nM)	Pep29 K_D (nM)
CIS43-Var2	38.2 ± 3.8	1.9 ± 0.4	14.9 ± 0.7	2.0 ± 0.2	3.6 ± 0.2	379 ± 16	670. ± 98
CIS43-Var10	39.7 ± 3.8	2.4 ± 0.7	15.1 ± 0.7	2.0 ± 0.2	4.0 ± 0.3	411 ± 14	218 ± 25
CIS43-Var11	53.3 ± 4.3	1.7 ± 0.3	23.1 ± 0.9	3.7 ± 0.2	1.4 ± 0.1	381 ± 14	83.5 ± 9.5
CIS43-Var18	43.9 ± 3.7	1.9 ± 0.4	14.1 ± 0.5	4.6 ± 0.2	1.9 ± 0.1	290. ± 9	261 ± 21
CIS43	71.7 ± 3.9	2.3 ± 0.5	36.3 ± 1.5	5.9 ± 0.2	10.7 ± 0.3	953 ± 72	48.8 ± 5.7

B ITC affinity of CIS43 variants against SA_mut_CSP



mAb	K_d1 (nM)	ΔG1 (kcal/mol)	ΔH1 (kcal/mol)	-TΔS1 (kcal/mol)	N1	Kd_2 (nM)	ΔG2 (kcal/mol)	ΔH2 (kcal/mol)	-TΔS2 (kcal/mol)	N2
Var 2	2.1 ± 0.6	-11.8 ± 0.2	-21.5 ± 2.0	+9.7 ± 1.1	2.0 ± 0.2	13 ± 2	-10.7 ± 0.1	-10.6 ± 0.5	-0.10 ± 0.01	5.8 ± 0.2
Var10	4.4 ± 1.1	-11.4 ± 0.1	-23.8 ± 3.2	+12.4 ± 1.8	1.9 ± 0.2	24 ± 2	-10.4 ± 0.1	-10.9 ± 0.6	+0.50 ± 0.03	6.7 ± 0.3
Var11	4.4 ± 1.4	-11.4 ± 0.2	-20.5 ± 1.4	+9.1 ± 0.8	1.9 ± 0.2	36 ± 3	-10.1 ± 0.1	-11.9 ± 0.5	+1.8 ± 0.1	5.9 ± 0.2
Var 18	3.2 ± 0.6	-11.6 ± 0.1	-26.7 ± 1.9	+15.1 ± 1.2	1.9 ± 0.2	12 ± 1	-10.8 ± 0.1	-13.9 ± 0.6	+3.1 ± 0.2	6.2 ± 0.2
CIS43	14 ± 3	-10.7 ± 0.1	-25.9 ± 3.2	+15.2 ± 2.0	2.0 ± 0.2	30 ± 3	-10.3 ± 0.1	-7.7 ± 0.4	-2.6 ± 0.2	5.3 ± 0.2

Figure 32. Supplemental Chapter 3, Screening of antibody multi-mutation variants by AlphaLISA, BLI, etc.

(A) Bio-layer interferometry (BLI) affinity of improved CIS43 variants; Var2, Var10, Var11, and Var18 was determined against fl_PfCSP, NTDS_5/3_CSP, NPDP19, Pep21, Pep25, and Pep29.

(B) Isothermal titration calorimetry of CSP_SA_mut with various improved CIS43 variants. The affinities and stoichiometry are shown for both K_D1 and K_D2.

Appendix O: Table 5. Supplemental Chapter 3.

Table 5. Supplemental Chapter 3. Cloning & transformation results and calculation of theoretical library coverage for CIS43 gene variant libraries.

Single-Mutant Libraries E. coli Transformations

Library Name	Theoretical Size*	# Transformants	Fold Coverage
CIS43-VH-SSM	3,872	4.5 E+05	116
CIS43-VL-SSM	3,616	4.0 E+05	111

Single-Mutant Libraries Yeast Transformations

Library Name	Theoretical Size*	# Transformants	Fold Coverage
CIS43-VH-SSM	3,872	3.0 E+07	7,748
CIS43-VL-SSM	3,616	8.0 E+06	2,212

*Theoretical library sizes calculated as: (Length of VH or VL gene) × (No. of codons).

CIS43-VH-SSM: 121 residues * 32 codons = 3,872 single mutation variants

CIS43-VL-SSM: 113 residues * 32 codons = 3,616 single mutation variants

Multi-Mutant Libraries E. coli Transformations

Library Name	Transformation	# Transformants
CIS43-VH:VL-Library1	1	1.6 E+08
CIS43-VH:VL-Library1	2	6.8 E+08
CIS43-VH:VL-Library2	1	2.4 E+08
CIS43-VH:VL-Library2	2	1.2 E+09
CIS43-VH:VL-Library3	1	1.1 E+08
CIS43-VH:VL-Library3	2	9.2 E+07

Multi-Mutant Libraries Yeast Transformations

Library Name	Transformation	# Transformants
CIS43-VH:VL-Library1	1	1.2 E+07
CIS43-VH:VL-Library1	2	1.0 E+07
CIS43-VH:VL-Library2	1	3.0 E+07
CIS43-VH:VL-Library2	2	2.9 E+07
CIS43-VH:VL-Library3	1	2.0 E+07
CIS43-VH:VL-Library3	2	1.8 E+06
CIS43-VH:VL-Library3	3	1.0 E+07

Appendix P: Table 6. Supplemental Chapter 3.

Table 6. Supplemental Chapter 3. Sequences of PfCSP-dervied antigen probes used to engineer and assess enhanced antibody protection.

Antigen Probe	AA Sequence
fl_PfCSP	QEYQCYGSSSNTRVLNELNYDNAGTNLYNELEMNYYGKQENWYSLKKNSRSLGENDDGNNEDNEKLRKPK HKKLKQPADGNPDPNANPNVDPNANPNVDPNANPNVDPNANPNANPNANPNANPNANPNANPNANPNANP NANPNANPNANPNANPNANPNANPNANPNANPNANPNVDPNANPNANPNANPNANPNANPNANPNANPNA NPNANPNANPNANPNANPNANPNANPNANPNANPNANPNANPNKNNQGNGQGHNMPNDPNRNVDENANA NSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEWSPCSVTCGNGIQVRIKPGSANKPKDELDYANDIEKKICKME KCS
NTDS_fl_CSP	QEYQSYGSSSNTRVLNELNYDNAGTNLYNELEMNYYGKQENWYSLSSNSASLGENDDGNNEDNEKLRKPK HKKLKQPADGNPDPNANPNVDPNANPNVDPNANPNVDPNANPNANPNANPNANPNANPNANPNANPNANP NANPNANPNANPNANPNANPNANPNANPNANPNANPNVDPNANPNANPNANPNANPNANPNANPNANPNA NPNANPNANPNANPNANPNANPNANPNANPNANPNANPNANPNKNNQGNGQGHNMPNDPNRNVDENANA NSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEWSPCSVTCGNGIQVRIKPGSANKPKDELDYANDIEKKICKME KCS
NTDS_5/3_CSP	QEYQSYGSSSNTRVLNELNYDNAGTNLYNELEMNYYGKQENWYSLSSNSASLGENDDGNNEDNEKLRKPK HKKLKQPADGNPDPNANPNVDPNANPNVDPNANPNVDPNANPNANPNKNNQGNGQGHNMPNDPNRNVDE NANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEWSPCSVTCGNGIQVRIKPGSANKPKDELDYANDIEKKIC KMEKCS
NPDP19	KQPADGNPDPNANPNVDPNAN
Peptide 21	NPDPNANPNVDPNAN
Peptide 25	NVDPNANPNVDPNAN
Peptide 29	NANPNANPNANPNAN

All circumsporozoite protein (CSP) sequences based on the *Plasmodium falciparum (Pf)* 3D7 strain

fl: 'full length'
NTDS: N-terminal domain stabilizing mutations
5/3: 5 major NANP repeats, 3 minor NVDP repeats

PEXEL II NTD recognition sequence: RSLGE

Conserved region 1 (R1) sequence: KLKQP
Junctional epitope: NPDP
Major repeat: NANP
Minor repeat: NVDP

Appendix Q: Table 7. Supplemental Chapter 3.

Table 7. CIS43 multi-mutation variants mined from the data in Figure 19 and expressed as soluble human IgG1 antibodies in HEK293 cells for detailed functional evaluation.

	Variants expressed as IgG	VH substitution(s)		VL substitution(s)	
		Residue number	*Kabat numbering*	Residue number	*Kabat numbering*
Single Mutants	CIS43_VH_V2C	V2C	*V2C*	WT	*WT*
	CIS43_VH_V2M	V2M	*V2M*	WT	*WT*
	CIS43_VH_V2R	V2R	*V2R*	WT	*WT*
	CIS43_VH_V2W	V2W	*V2W*	WT	*WT*
	CIS43_VH_Q3L	Q3L	*Q3L*	WT	*WT*
	CIS43_VH_V5E	V5E	*V5E*	WT	*WT*
	CIS43_VH_K12G	K12G	*K12G*	WT	*WT*
	CIS43_VH_K13T	K13T	*K13T*	WT	*WT*
	CIS43_VH_T28M	T28M	*T28M*	WT	*WT*
	CIS43_VH_T28G	T28G	*T28G*	WT	*WT*
	CIS43_VH_T74R	T74R	*T73R*	WT	*WT*
	CIS43_VH_S84A	S84A	*S82aA*	WT	*WT*
	CIS43_VH_T103Q	T103Q	*T99Q*	WT	*WT*
	CIS43_VH_M116V	M116V	*M112V*	WT	*WT*
	CIS43_VL_S10C	WT	*WT*	S10C	*S10C*
	CIS43_VL_N22H	WT	*WT*	N22H	*N22H*
	CIS43_VL_L39V	WT	*WT*	L39V	*L33V*
	CIS43_VL_N51E	WT	*WT*	N51E	*N45E*
	CIS43_VL_N51R	WT	*WT*	N51R	*N45R*
	CIS43_VL_I54F	WT	*WT*	I54F	*I48F*
	CIS43_VL_T59H	WT	*WT*	T59H	*T53H*
	CIS43_VL_S82D	WT	*WT*	S82D	*S76D*
	CIS43_VL_L84F	WT	*WT*	L84F	*L78F*
	CIS43_VL_A86P	WT	*WT*	A86P	*A80P*
	CIS43_VL_V89E	WT	*WT*	V89E	*V83E*
	CIS43_VL_V91D	WT	*WT*	V91D	*V85D*
Multi-Mutants	CIS43_VH_V2R, T28G, S84A	V2R, T28G, S84A	*V2R, T28G, S82aA*	WT	*WT*
	CIS43_VH_V2R, T28G, S84A, T103Q	V2R, T28G, S84A, T103Q	*V2R, T28G, S82aA, T99Q*	WT	*WT*
	CIS43_VL_N22H, L84F	WT	*WT*	N22H, L84F	*N22H, L78F*
	CIS43_VL_N22H, N51R, L84F	WT	*WT*	N22H, N51R, L84F	*N22H, N45R, L78F*
	CIS43_Var1	V2R	*V2R*	A19V	*A19V*
	CIS43_Var2	V2R	*V2R*	L30F	*L27cF*
	CIS43_Var3	V2R	*V2R*	V91D	*V85D*
	CIS43_Var4	V2R	*V2R*	S14R	*S14R*
	CIS43_Var5	V2R	*V2R*	S82D	*S76D*
	CIS43_Var6	V2R	*V2R*	N51R	*N45R*
	CIS43_Var7	V2R	*V2R*	Q61A	*Q55A*
	CIS43_Var8	V2R	*V2R*	N51R, Q61A	*N45R, Q55A*
	CIS43_Var9	V2R	*V2R*	Y31W, S82D	*Y27dW, S76D*
	CIS43_Var10	V2R	*V2R*	L30F, V91S	*L27cF, V85S*
	CIS43_Var11	V2R	*V2R*	N51R, V91S	*N45R, V85S*
	CIS43_Var12	V2R	*V2R*	A19V, V91S	*A19V*
	CIS43_Var13	V2R, A97V	*V2R, A93V*	WT	*WT*
	CIS43_Var14	V2R, A97V	*V2R, A93V*	A19V	*A19V*
	CIS43_Var15	V2R, A97V	*V2R, A93V*	N51R	*N45R*
	CIS43_Var16	V2R, A97V	*V2R, A93V*	Q61A	*Q55A*
	CIS43_Var17	V2R, A97V	*V2R, A93V*	Q27R	*Q27R*
	CIS43_Var18	V2R, A97V	*V2R, A93V*	N51R, Q61A	*N45R, Q55A*
	CIS43_Var19	V2R, A97V	*V2R, A93V*	Y31W, S82D	*Y27dW, S76D*
	CIS43_Var20	V2R, N57F	*V2R, N56F*	A19V	*A19V*
	CIS43_Var21	V2R, N57F	*V2R, N56F*	Y31W, S82D	*Y27dW, S76D*
	CIS43_Var22	V2R, N57F	*V2R, N56F*	N51R, Q61A	*N45R, Q55A*
	CIS43_Var23	V2R, S88D, A97V	*V2R, S84D, A93V*	S14R	*S14R*
	CIS43_Var24	V2R, S88D, A97V	*V2R, S84D, A93V*	S82D	*S76D*

WT = Wild Type CIS43 template VH or VL gene

Appendix R: Table 8. Supplemental Chapter 3.

Table 8. Supplemental Chapter 3. Multi-mutation library mutant distribution data corresponding to Figure 25A.

			# of mutations to VH Gene						# of mutations to VL Gene					
			0 (unmutated)	1	2	3	4	5+	0 (unmutated)	1	2	3	4	5+
Library 1	VL+	# of unique sequences in library w/ # of mutations	1	802	4465	1810	379	160	1	852	14053	5782	908	158
		# of reads in library w/ # of mutations	97742	94827	17268	2705	666	230	231917	217110	54835	8298	1174	249
		Total % of library reads w/ # of mutations	0.46	0.44	0.08	0.01	0.00	0.00	0.45	0.42	0.11	0.02	0.00	0.00
	4x fl_PfCSP	unique sequences	1	116	1186	12381	3574	2421	1	521	5405	13120	20254	6281
		reads	304	321474	95897	20677	5239	3173	5159	211152	611144	158986	40950	10797
		% of library	0.00	0.72	0.21	0.05	0.01	0.01	0.00	0.20	0.59	0.15	0.04	0.01
	4x NTDS_5/3_CSP	unique sequences	1	36	806	2794	466	127	1	79	715	1432	1292	157
		reads	73	101238	7358	673	31	12	424	137517	92493	566998	14626	9081
		% of library	0.00	0.93	0.07	0.01	0.00	0.00	0.00	0.17	0.11	0.69	0.02	0.01
	4x Pep21	unique sequences	1	23	599	452	109	115	1	156	907	1598	468	140
		reads	43	12132	3157	569	127	105	644	24191	8391	2394	569	165
		% of library	0.00	0.75	0.20	0.04	0.01	0.01	0.02	0.67	0.23	0.07	0.02	0.00
	4x Pep25	unique sequences	1	25	713	1120	130	42	1	222	1751	1294	250	41
		reads	86	67202	13368	1293	139	37	2442	44262	21113	2554	271	75
		% of library	0.00	0.82	0.16	0.02	0.00	0.00	0.03	0.63	0.30	0.04	0.00	0.00
	4x Pep29	unique sequences	1	13	658	1123	891	836	1	314	1758	1081	412	99
		reads	3	16901	7200	11539	3061	826	1307	14138	5045	1444	447	104
		% of library	0.00	0.43	0.18	0.29	0.08	0.02	0.06	0.63	0.22	0.06	0.02	0.00
Library 2	VL+	unique sequences	1	727	2588	1056	374	190	1	903	18568	8960	1322	214
		reads	29049	29124	7084	2054	601	188	373627	348104	90965	13316	1712	270
		% of library	0.43	0.43	0.10	0.03	0.01	0.00	0.45	0.42	0.11	0.02	0.00	0.00
	4x fl_PfCSP	unique sequences	1	76	1152	10006	2414	647	1	351	4941	8281	1192	215
		reads	263	563496	124030	16928	2577	581	4466	703928	102425	10487	1272	230
		% of library	0.00	0.80	0.18	0.02	0.00	0.00	0.01	0.86	0.12	0.01	0.00	0.00
	4x NTDS_5/3_CSP	unique sequences	1	55	802	1965	468	405	1	351	4941	8281	1192	215
		reads	92	41631	12187	2626	687	419	1855	149286	49385	7220	61	3
		% of library	0.00	0.72	0.21	0.05	0.01	0.01	0.01	0.72	0.24	0.03	0.00	0.00
	4x Pep21	unique sequences	1	33	727	884	298	211	1	171	1505	2559	1351	429
		reads	36	15930	5572	1590	390	192	703	22425	15438	7940	1974	482
		% of library	0.00	0.67	0.24	0.07	0.02	0.01	0.01	0.46	0.32	0.16	0.04	0.01
	4x Pep25	unique sequences	1	74	743	1122	306	219	1	395	3858	3861	1494	444
		reads	192	18256	5978	1487	387	203	2963	54728	21555	7103	1782	469
		% of library	0.01	0.69	0.23	0.06	0.01	0.01	0.03	0.62	0.24	0.08	0.02	0.01
	4x Pep29	unique sequences	1	7	101	574	383	186	1	245	891	366	106	39
		reads	6	342	7998	2246	449	168	954	3874	1648	401	108	41
		% of library	0.00	0.03	0.71	0.20	0.04	0.01	0.14	0.55	0.23	0.06	0.02	0.01
Library 3	VL+	unique sequences	1	1230	4286	1861	670	408	1	1438	16433	8333	1732	364
		reads	54022	51750	15581	5892	2137	684	154703	154786	62112	13991	2719	715
		% of library	0.42	0.40	0.12	0.05	0.02	0.01	0.40	0.40	0.16	0.04	0.01	0.00
	4x fl_PfCSP	unique sequences	1	105	887	4295	2274	997	1	712	9877	26412	16513	4909
		reads	179	43959	92051	36394	11595	3928	22723	431500	344026	94080	23924	6397
		% of library	0.00	0.23	0.49	0.19	0.06	0.02	0.02	0.47	0.37	0.10	0.03	0.01
	4x NTDS_5/3_CSP	unique sequences	1	34	795	3724	2521	1284	1	277	3048	4828	2796	729
		reads	56	26903	47304	14445	3240	1248	966	46396	53331	15090	3662	859
		% of library	0.00	0.29	0.51	0.15	0.03	0.01	0.01	0.39	0.44	0.13	0.03	0.01
	4x Pep21	unique sequences	1	81	884	780	96	33	1	77	1017	1479	356	47
		reads	495	31927	9049	1110	150	47	396	18922	19546	3360	369	48
		% of library	0.01	0.75	0.21	0.03	0.00	0.00	0.01	0.44	0.46	0.08	0.01	0.00
	4x Pep25	unique sequences	1	126	856	2521	818	572	1	233	3762	4315	1405	381
		reads	388	32509	17352	4714	1026	539	702	33157	20025	6433	1537	410
		% of library	0.01	0.58	0.31	0.08	0.02	0.01	0.01	0.53	0.32	0.10	0.02	0.01
	4x Pep29	unique sequences	1	18	328	832	1326	504	1	168	998	1794	458	174
		reads	42	2337	24455	7392	1631	513	680	31779	9407	2477	590	212
		% of library	0.00	0.06	0.67	0.20	0.04	0.01	0.02	0.70	0.21	0.05	0.01	0.00

Appendix S: Table 9. Supplemental Chapter 3.

Table 9. Supplemental Chapter 3. X-ray crystallography data collection and refinement statistics associated with Figure 21.

	CIS43_Var2 : Pep21	**CIS43_Var10 : Pep21**
PDB accession code	*7SG5*	*7SG6*
Data collection		
Wavelength	1	1
Resolution range (Å)	50.72 - 1.40 (1.45 - 1.40)	44.96 - 1.55 (1.60 - 1.55)
Space group	C 2	C 2
Cell dimensions *a, b, c* (Å) a, b, g (°)	94.8 61.1 75.3 90.0, 106.5, 90.0	93.8 61.6 75.0 90.0, 105.9, 90.0
Unique reflections	61516 (4587)	58909 (5896)
Multiplicity	2.1 (1.6)	3.5 (3.5)
Completeness (%)	75.5 (56.5)	98.6 (98.9)
I / σI	14.7 (2.2)	5.6 (1.4)
Wilson *B*-factor	15.55	13.62
R_{merge}	0.044 (0.349)	0.132 (0.846)
$CC_{1/2}$	0.996 (0.82)	0.985 (0.56)
Refinement		
Reflections used in refinement	61324 (4583)	58823 (5896)
Reflections used for R-free	3047 (232)	2929 (249)
R_{work}	0.18 (0.28)	0.18 (0.24)
R_{free}	0.22 (0.32)	0.21 (0.27)
Number of non-hydrogen atoms	3854	3765
macromolecules	3451	3456
ligands	31	6
solvent	372	303
Protein residues	451	452
RMS(bonds) (Å)	0.007	0.008
RMS(angles) (°)	1.05	1.12
Ramachandran favored (%)	97.98	97.98
Ramachandran outliers (%)	0.00	0.22
Average *B*-factor (Å^2)	19.38	16.39
macromolecules	18.69	15.88
ligands	28.36	17.50
solvent	25.08	22.16

*Values in parentheses are for highest-resolution shell.

Appendix T: Table 10. Supplemental Chapter 3.

Table 10. Supplemental Chapter 3. Cryo-EM data collection and reconstruction statistics associated with Figure 22.

	CIS43 Variant 10 bound to pfCSP Class 1	**CIS43 Variant 10 bound to pfCSP Class 2**	**CIS43 Variant 10 bound to pfCSP Class 3**	**CIS43 Variant 10 bound to pfCSP Class 4**
EMDB ID	EMD-xxxx	EMD-xxxxx	EMD-xxxx	EMD-xxxxx
Data Collection				
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300	300	300
Electron dose ($e^-/Å^2$)	64.12	64.12	64.12	64.12
Detector	Gatan K3	Gatan K3	Gatan K3	Gatan K3
Pixel Size (Å)	1.083	1.083	1.083	1.083
Defocus Range (μm)	-1 to -2.5	-1 to -2.5	-1 to -2.5	-1 to -2.5
Magnification	81000	81000	81000	81000
Reconstruction				
Software	cryoSPARC	cryoSPARC	cryoSPARC	cryoSPARC
particles	67,433	34,637	30,191	28,900
Symmetry	C1	C1	C1	C1
Box size (pix)	320	320	320	320
Resolution (Å) ($FSC_{0.143}$)	6.1	7.1	10.9	9.8

Appendix U: Figure 33. Supplemental Chapter 4.

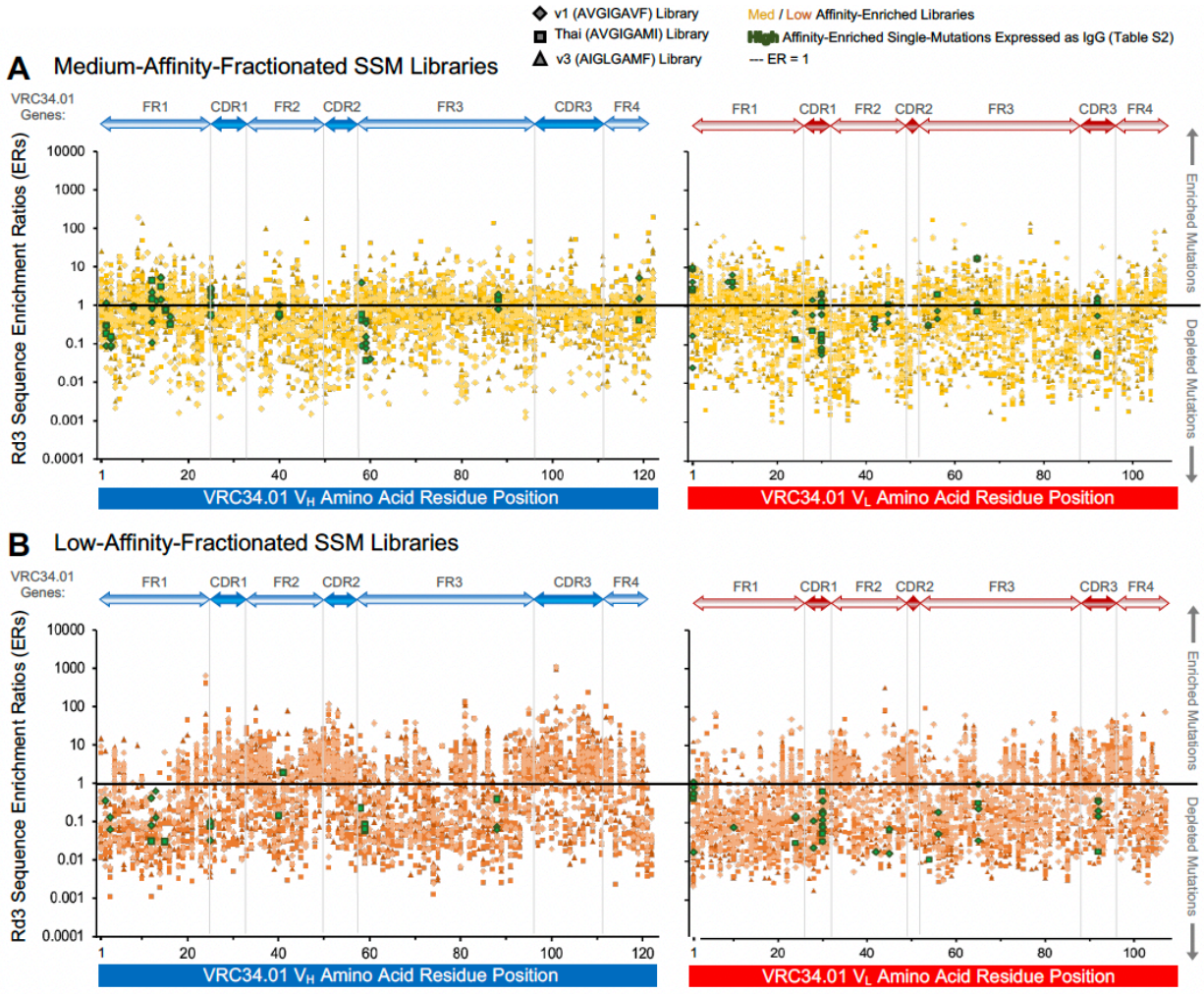


Figure 33. Supplemental Chapter 4. Bioinformatic analysis of single mutant library screens. Enrichment ratios (ER) are plotted for single mutant antibody sequences derived from round three medium affinity. (A) and low-affinity (B) sorted libraries against the mutant residue location in the antibody variable region. Analysis of NGS data reveal the functional impact of single mutations against separate diverse fusion peptide sequences and help determine high-affinity candidates for biophysical evaluation.

Appendix V: Figure 34. Supplemental Chapter 4.

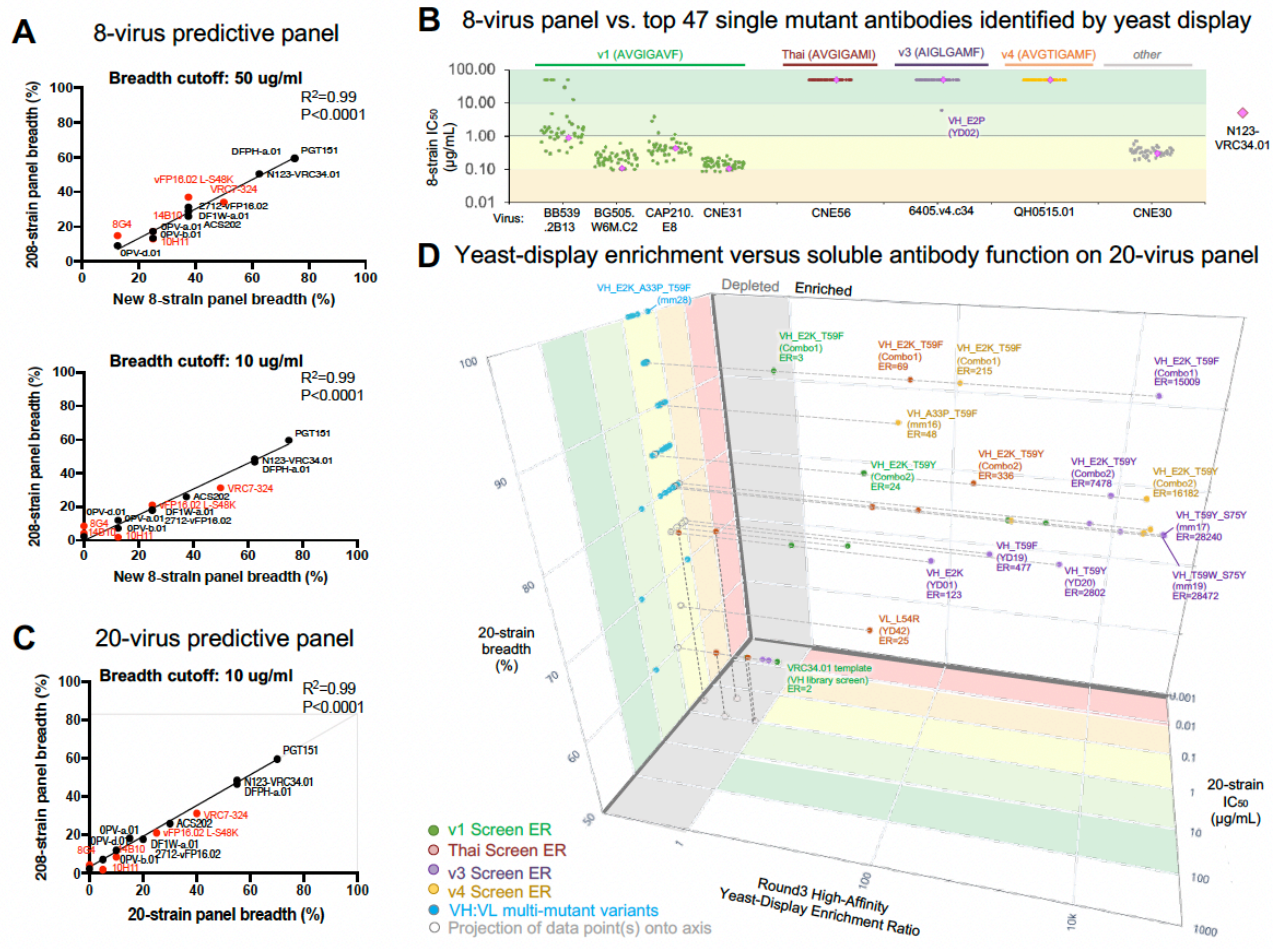


Figure 34. Supplemental Chapter 4. Directed antibody evolution informed by predictive anti-FP neutralization panels.

(A) 8-virus predictive panel design
(B) Top enriched single mutant neutralization on 8 virus panel
(C) 20-virus predictive panel design
(D) Yeast-display enrichment versus soluble antibody function on 20-virus panel

Appendix W: Figure 35. Supplemental Chapter 4.

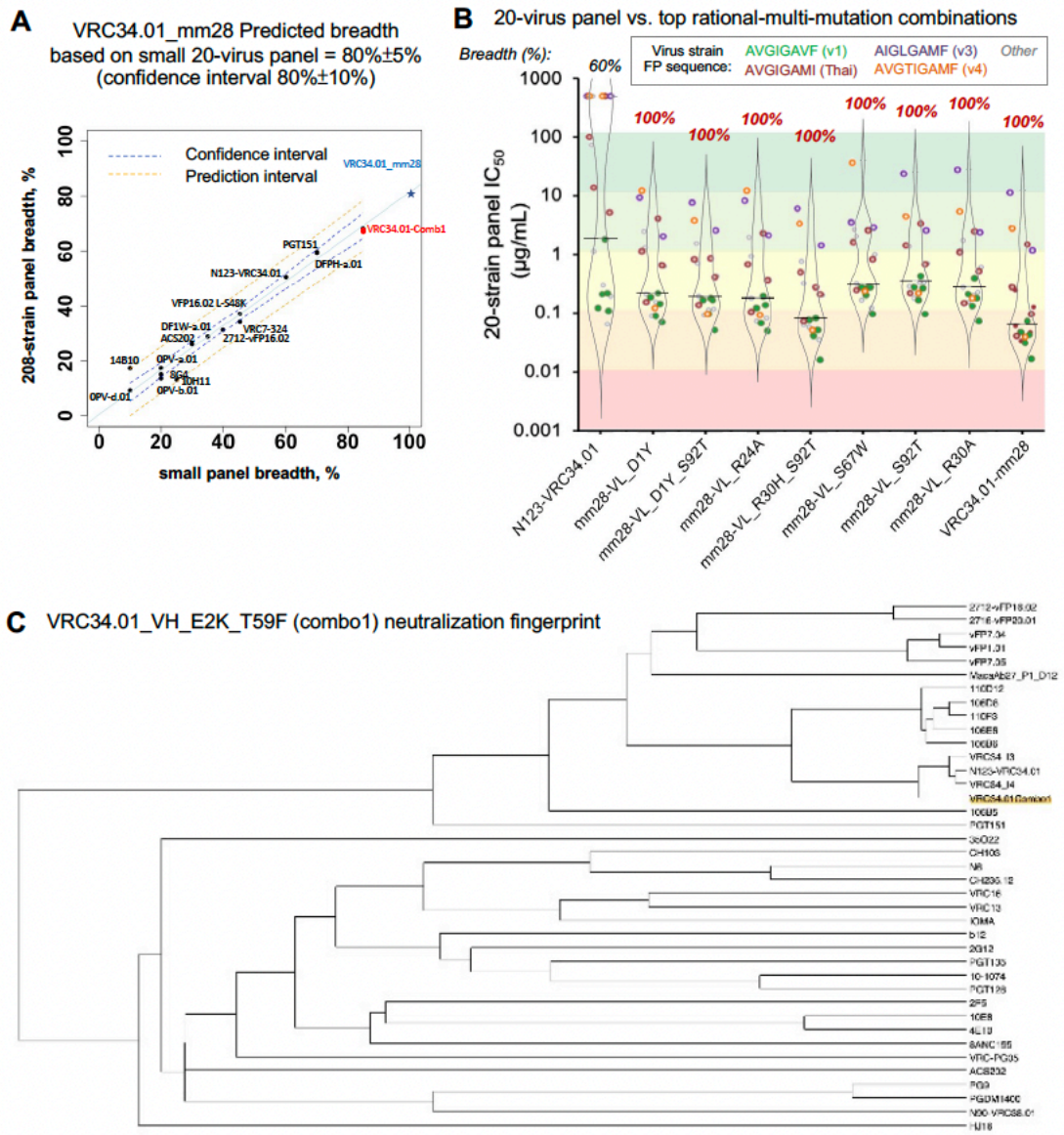


Figure 35. Supplemental Chapter 4. Top VRC34.01 mutant variant characterization.

(A) VRC34.01_mm28 predicted breadth on a 208-virus panel, based on 20-virus panel predictive panel (80%+/-5% confidence interval 80%+/-10%). Because no reported anti-FP antibody has been reported with as high breadth as VRC34.01_mm28, this prediction requires extrapolation beyond current data points.

(B) Multi-mutation 20-virus panel results, including mm28.

(C) VRC34.01_combo1 fingerprint analysis indicates template and mutant variants neutralize similar HIV-1 strains.

Appendix X: Figure 36. Supplemental Chapter 4.

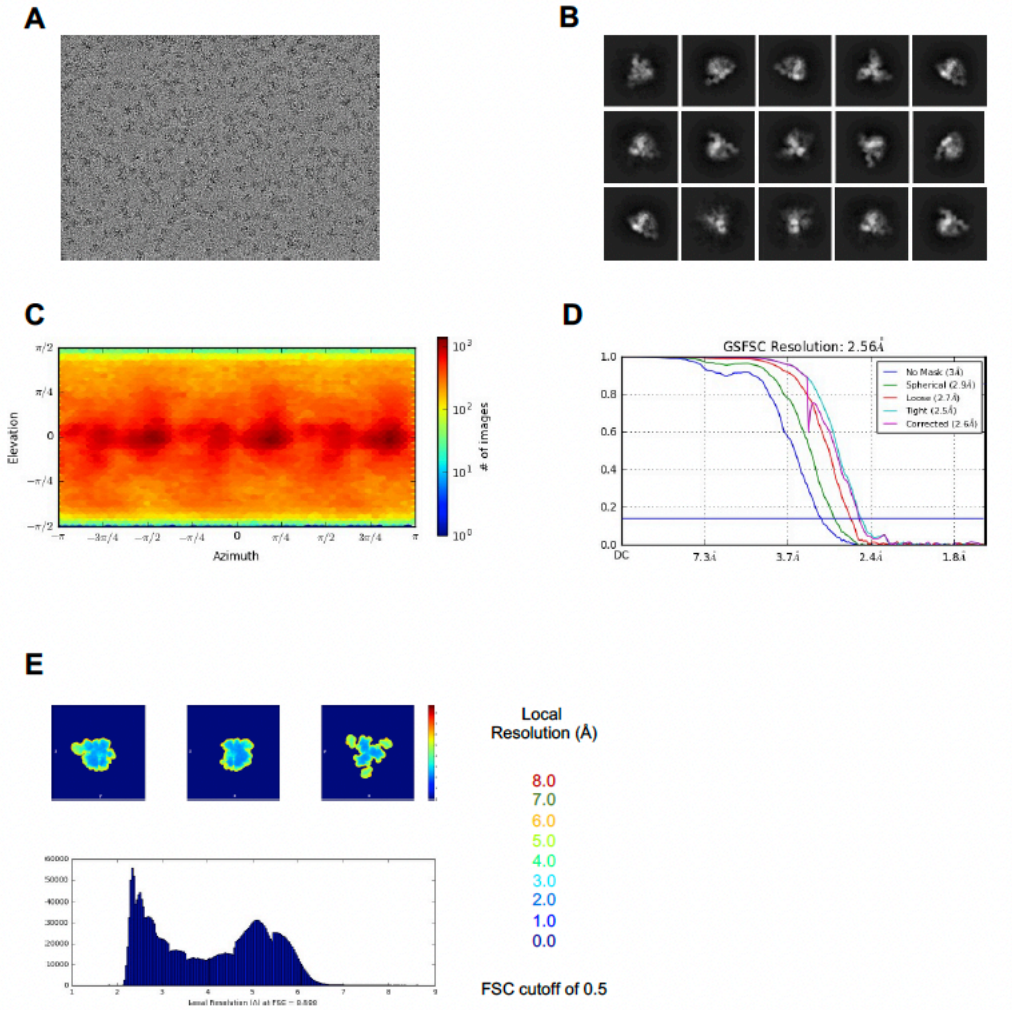


Figure 36. Supplemental Chapter 4. Cryo-EM Details of VRC34-combo.1 in complex with HIV-1, Related to Figure 4.

(A) Representative micrograph are shown.
(B) Representative 2D class averages are shown.
(C) The orientations of all particles used in the final refinement are shown as a heatmap.
(D) The gold-standard Fourier shell correlation resulted in a resolution of 2.56 using non-uniform refinement with C3 symmetry.
(E) The local resolution of the full map is shown generated through cryoSPARC using an FSC cutoff of 0.5.

Appendix Y: Table 11. Supplemental Chapter 4.

Table 11. Supplemental Chapter 4. VRC34.01 mutant-library generation & cloning statistics.

		Single Mutant Libraries		Multi-Mutant Libraries				
	Library Name:	VRC34.01-VH-SSM	VRC34.01-VL-SSM	Library 1	Library 2	Library 3	Library 4	Library 5
	Gene augmentation strategy	NNK codon substitution	MNN codon substitution	Rd3-High VH:VL	Rd3-High-VH-SSM : Template-VL	Template-VH : Rd3-High-VL-SSM	Rd3-High-Shuffled VH:VL	Rd3-High-Shuffled-SSM VH:VL
	*Theoretical Number of Variants	3,904	3,424	*unknown*	*unknown*	*unknown*	*unknown*	*unknown*
E. coli transformations	# Transformants	4.00E+06	3.70E+06	1.80E+07	4.50E+07	4.30E+07	Transf.1 (Shuff-VH) - 9.90E+07, Transf.2 (Shuff-VL) - 8.40E+07	Transf.1 (VH-SSM) - 1.80E+8, Transf.2 (VL-SSM) - 1.55E+08
E. coli transformations	Transformed clones with single mutations	7 out of 12	10 out of 12	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
E. coli transformations	Adujsted Fold Coverage Estimate	598	900	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
Yeast transformations	# Transformants	2.10E+06	2.90E+06	4.50E+06	4.90E+06	5.90E+06	6.90E+06	3.20E+06
Yeast transformations	Theoretical Fold Coverage	538	847	*unknown*	*unknown*	*unknown*	*unknown*	*unknown*

*Theoretical number of variants calculated as (Length of variable gene) * (No. of possible codons)
VRC34.01-VH-SSM: 122 residues * 32 codons = 3,904 possible single mutant variants
VRC34.01-VL-SSM: 107 residues * 32 codons = 3,424 possible single mutant variants

Appendix Z: Table 12. Supplemental Chapter 4.

Table 12. Supplemental Chapter 4. VRC34.01 Variants Expressed as IgG.

Single-mutant variants identified by yeast-library screening

	Antibody IgG	Heavy chain	Kabat ID	Light chain	Kabat ID
1	YD01	VRC34.01-HC-E2K	V2E	VRC34.01 L	*Wild Type*
2	YD02	VRC34.01-HC-E2P	V2P	VRC34.01 L	*Wild Type*
3	YD03	VRC34.01-HC-V3W	Q3W	VRC34.01 L	*Wild Type*
4	YD04	VRC34.01-HC-V3F	Q3F	VRC34.01 L	*Wild Type*
5	YD05	VRC34.01-HC-G8D	G8D	VRC34.01 L	*Wild Type*
6	YD06	VRC34.01-HC-K12A	K12A	VRC34.01 L	*Wild Type*
7	YD07	VRC34.01-HC-K12T	K12T	VRC34.01 L	*Wild Type*
8	YD08	VRC34.01-HC-K13L	K13L	VRC34.01 L	*Wild Type*
9	YD09	VRC34.01-HC-P14Q	P14Q	VRC34.01 L	*Wild Type*
10	YD10	VRC34.01-HC-G15E	G15E	VRC34.01 L	*Wild Type*
11	YD11	VRC34.01-HC-A16D	A16D	VRC34.01 L	*Wild Type*
12	YD12	VRC34.01-HC-F25A	S25A	VRC34.01 L	*Wild Type*
13	YD13	VRC34.01-HC-F25G	S25G	VRC34.01 L	*Wild Type*
14	YD14	VRC34.01-HC-F25L	S25L	VRC34.01 L	*Wild Type*
15	YD15	VRC34.01-HC-A40T	A40T	VRC34.01 L	*Wild Type*
16	YD16	VRC34.01-HC-P41Y	P41Y	VRC34.01 L	*Wild Type*
17	YD17	VRC34.01-HC-T58Q	T57Q	VRC34.01 L	*Wild Type*
18	YD18	VRC34.01-HC-T58V	T57V	VRC34.01 L	*Wild Type*
19	YD19	VRC34.01-HC-T59F	N58F	VRC34.01 L	*Wild Type*
20	YD20	VRC34.01-HC-T59Y	N58Y	VRC34.01 L	*Wild Type*
21	YD21	VRC34.01-HC-T59W	N58W	VRC34.01 L	*Wild Type*
22	YD22	VRC34.01-HC-T59V	N58V	VRC34.01 L	*Wild Type*
23	YD23	VRC34.01-HC-T60L	Y59L	VRC34.01 L	*Wild Type*
24	YD24	VRC34.01-HC-T60V	Y59V	VRC34.01 L	*Wild Type*
25	YD25	VRC34.01-HC-S88G	S84G	VRC34.01 L	*Wild Type*
26	YD26	VRC34.01-HC-T119M	T115M	VRC34.01 L	*Wild Type*
27	YD27	VRC34.01 H	*Wild Type*	VRC34.01-LC-D1G	
28	YD28	VRC34.01 H	*Wild Type*	VRC34.01-LC-D1I	
29	YD29	VRC34.01 H	*Wild Type*	VRC34.01-LC-D1Y	
30	YD30	VRC34.01 H	*Wild Type*	VRC34.01-LC-S9D	
31	YD31	VRC34.01 H	*Wild Type*	VRC34.01-LC-F10I	
32	YD32	VRC34.01 H	*Wild Type*	VRC34.01-LC-R24D	
33	YD33	VRC34.01 H	*Wild Type*	VRC34.01-LC-G28H	
34	YD34	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30A	
35	YD35	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30F	
36	YD36	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30Q	
37	YD37	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30S	
38	YD38	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30W	
39	YD39	VRC34.01 H	*Wild Type*	VRC34.01-LC-R30Y	
40	YD40	VRC34.01 H	*Wild Type*	VRC34.01-LC-K42R	
41	YD41	VRC34.01 H	*Wild Type*	VRC34.01-LC-N45T	
42	YD42	VRC34.01 H	*Wild Type*	VRC34.01-LC-L54R	
43	YD43	VRC34.01 H	*Wild Type*	VRC34.01-LC-S56H	
44	YD44	VRC34.01 H	*Wild Type*	VRC34.01-LC-T65R	
45	YD45	VRC34.01 H	*Wild Type*	VRC34.01-LC-T65S	
46	YD46	VRC34.01 H	*Wild Type*	VRC34.01-LC-S92F	
47	YD47	VRC34.01 H	*Wild Type*	VRC34.01-LC-S92T	

Rational combinations of top single mutants identified by yeast-library screening

	Antibody IgG	Heavy chain	Light chain
1	VRC34.01_Combo1	VH_E2K_T59F	VRC34.01 L
2	VRC34.01_Combo2	VH_E2K_T59Y	VRC34.01 L
3	*VRC34.01_Combo3	VH_E2P_T59F	VRC34.01 L
4	*VRC34.01_Combo4	VH_E2P_T59Y	VRC34.01 L

*Did not express

Multi-mutant variants identified by yeast-library screening

	Antibody IgG	Heavy chain	Light Chain
1	VRC34.01_mm1	VH_E2K_T59F	VL_D1Y
2	VRC34.01_mm2	VH_E2K_T59F	VL_D1Y_S92T
3	VRC34.01_mm3	VH_E2K_T59F	VL_R24A
4	VRC34.01_mm4	VH_E2K_T59F	VL_R30H_S92T
5	VRC34.01_mm5	VH_E2K_T59F	VL_S67W
6	VRC34.01_mm6	VH_E2K_T59F	VL_S92T
7	VRC34.01_mm7	VH_ E2K_K12A_T59F	VL_D1Y
8	VRC34.01_mm8	VH_E2K_K12A_T59F	VL_D1Y_S92T
9	VRC34.01_mm9	VH_E2K_K12T_T59F	VL_D1Y
10	VRC34.01_mm10	VH_E2K_K12T_T59F	VL_D1Y_S92T
11	VRC34.01_mm11	VH_K12A_T59F_S75Y	VL_R24A
12	VRC34.01_mm12	VH_T59F	VL_S67W
13	VRC34.01_mm13	VH_T59F	VL_D1Y_S92T
14	VRC34.01_mm14	VH_T59F	VL_R30H_S92T
15	VRC34.01_mm15	VH_T59F	VL_R30A
16	VRC34.01_mm16	VH_A33P_T59F	VRC34.01 L
17	VRC34.01_mm17	VH_T59Y_S75Y	VRC34.01 L
18	VRC34.01_mm18	VH_T59Y_S75W	VRC34.01 L
19	VRC34.01_mm19	VH_T59W_S75Y	VRC34.01 L
20	VRC34.01_mm20	VH_T59F_S75W	VRC34.01 L
21	VRC34.01_mm21	VH_T59F_S75Y	VL_R24A
22	VRC34.01_mm22	VH_E2K_T59F_S75Y	VRC34.01 L
23	VRC34.01_mm23	VH_E2K_T28N_T59F	VRC34.01 L
24	VRC34.01_mm24	VH_E2K_T59F_S75Y	VL_R30H_S92T
25	VRC34.01_mm25	VH_E2K_T59F_S75Y	VL_S67W
26	VRC34.01_mm26	VH_E2K_T59F_S75Y	VL_S92T
27	VRC34.01_mm27	VH_E2K_T59F_S75W	VRC34.01 L
28	VRC34.01_mm28	VH_E2K_A33P_T59F	VRC34.01 L

Rational combinations of top multi-mutants identified by yeast-library screening

	Antibody IgG	Heavy chain	Light chain
1	mm28_VL_D1Y	VH_E2K_A33P_T59F	VL_D1Y
2	mm28_VL_D1Y_S92T	VH_E2K_A33P_T59F	VL_D1Y_S92T
3	mm28_VL_R24A	VH_E2K_A33P_T59F	VL_R24A
4	mm28_VL_R30H_S92T	VH_E2K_A33P_T59F	VL_R30H_S92T
5	mm28_VL_S67W	VH_E2K_A33P_T59F	VL_S67W
6	mm28_VL_S92T	VH_E2K_A33P_T59F	VL_S92T
7	mm28_VL_R30A	VH_E2K_A33P_T59F	VL_R30A

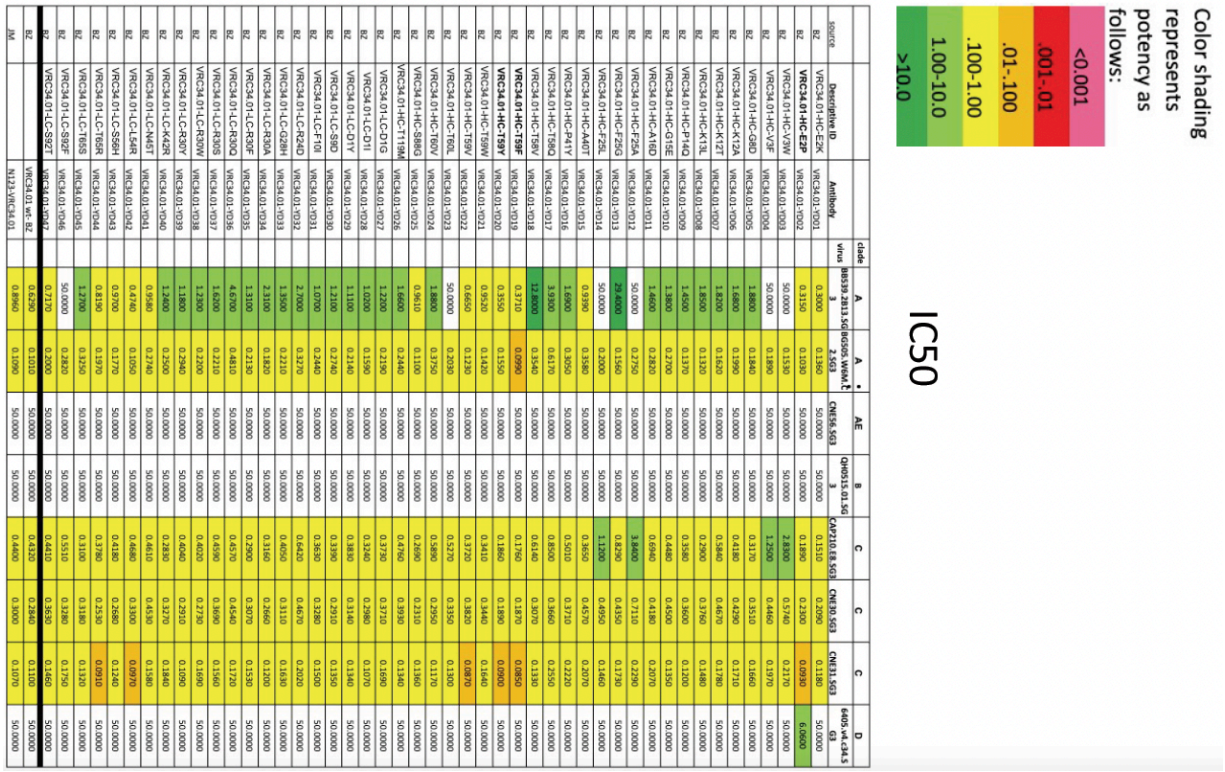
Appendix AA: Table 13. Supplemental Chapter 4.

Table 13. Supplemental Chapter 4. 8-virus predictive panel design.

			Control Antibodies IC50 (ug/mL)								
Strains	**Clade**	**FP seq**	**ACS202**	**N123-VRC34.01**	**PGT151**	**DF1W314**	**0PV12**	**0PV20**	**0PV21**	**110D12**	**2712-vFP16.02**
BG505.W6M.C2	A	**AVGIGAVF – v1**	50	0.211	0.004	0.257	15.2	6.07	20.7	2.21	2.51
CNE30	C	**AVGLGAVF – other**	50	0.191	50	500	500	500	500	5.29	100
BB539.2B13	A	**AVGIGAVF – v1**	50	1.82	0.005	0.566	460	500	122	10.9	100
CAP210.E8	C	**AVGIGAVF – v1**	1.02	0.218	0.032	500	500	101	500	2.66	43.8
CNE31	C	**AVGIGAVF – v1**	0.037	0.122	0.01	314	500	500	500	5.2	100
CNE56	AE	**AVGIGAMI – Thai**	50	500	50	500	3.19	24.6	133	0.65	0.405
6405.v4.c34	D	**AIGLGAMF – v3**	50	500	0.036	28.9	500	500	500	50	100
QH0515.01	B	**AVGTIGAMF – v4**	0.422	500	0.009	500	500	500	500	50	100

Appendix AB: Table 14. Supplemental Chapter 4.

Table 14. Supplemental Chapter 4. 47 single-mutant variants vs. 8 viruses predictive panel IC50 & IC80 values.



Appendix AC: Table 15. Supplemental Chapter 4.

Table 15. Supplemental Chapter 4. 20-virus predictive panel design.

				Control Antibodies IC50 (ug/mL)								
Strains	**Clade**	**FP seq**		**ACS202**	**N123-VRC34.01**	**PGT151**	**DF1W314**	**0PV12**	**0PV20**	**0PV21**	**110D12**	**2712-vFP16.02**
BG505.W6M.C2	A	AVGIGAVF	**FP_v1**	50	0.211	0.004	0.257	15.2	6.07	20.7	2.21	2.51
CNE30	C	AVGLGAVF	other	50	0.191	50	500	500	500	500	5.29	100
BB539.2B13	A	AVGIGAVF	**FP_v1**	50	1.82	0.005	0.566	460	500	122	10.9	100
CAP210.E8	C	AVGIGAVF	**FP_v1**	1.02	0.218	0.032	500	500	101	500	2.66	43.8
CNE31	C	AVGIGAVF	**FP_v1**	0.037	0.122	0.01	314	500	500	500	5.2	100
CNE56	AE	AVGIGAMI	**FP_Thai**	50	500	50	500	3.19	24.6	133	0.65	0.405
8405.v4.c34	D	AIGLGAMF	**FP_v3**	50	500	0.036	28.9	500	500	500	50	100
QH0515.01	B	AVGTIGAMF	**FP_v4**	0.422	500	0.009	500	500	500	500	50	100
C1080.c3	AE	AVGIGAMI	**FP_Thai**	50	5.2	50	70.2	246	321	66.1	50	19
C2101.c1	AE	AVGIGAMI	**FP_Thai**	50	13.8	50	500	500	500	500	50	11.7
C4118.09	AE	AVGIGAMI	**FP_Thai**	0.789	103	50	12	500	500	500	17.1	57.9
A03349M1	D	AIGLGAMF	**FP_v3**	0.147	500	50	13.7	500	500	500	50	100
CNE7	BC	AVGTIGAMF	**FP_v4**	0.179	500	0.021	500	500	500	500	50	100
T266-60	AG	AVGLGAVF	other	50	0.301	0.013	500	500	500	500	1.86	100
DU156.12	C	AVGLGAVL	other	50	500	0.007	500	500	500	500	0.79	78.2
3415.v1.c1	A	AIGMGAVF	other	50	0.126	0.005	500	500	500	500	4.44	100
CAP45.G3	C	AVGIGAVL	other	50	0.064	0.02	500	500	357	500	11.7	6.65
BI369.9A	A	AVGIGAVF	**FP_v1**	50	0.108	0.059	0.118	3.43	1.59	9.59	0.39	0.4
269-12	AG	AIGMGAVF	other	50	72	0.007	500	500	500	500	2.79	100
0077_V1.C16	C	AVGIGAMF	other	50	1.12	0.005	1.64	14.9	1.7	500	0.849	0.643
			20-virus IC50 geomean:	10.36	6.08	0.14	64.26	204.96	172.90	275.83	6.56	24.79
			20-virus breadth	30%	60%	70%	35%	20%	20%	10%	70%	40%

Appendix AD: Table 16. Supplemental Chapter 4.

Table 16. Supplemental Chapter 4. 20 virus panel comprehensive neutralization comparisons.

Top single-mutant variants identified by yeast-library screening (See Fig. 2C)

Virus ID	FP Sequence		N123-VRC34.01 (panel data)	VRC34.01-HC-E2K (YD01)	VRC34.01-HC-E2P (YD02)	VRC34.01-HC-T59F (YD19)	VRC34.01-HC-T59Y (YD20)	VRC34.01-LC-L54R (YD42)
C1080.c3	AVGIGAMI	FP_Thai	5.2	0.168	0.056	0.072	0.149	14.2
C2101.c1	AVGIGAMI	FP_Thai	13.8	>50	0.099	0.119	0.177	5.54
C4118.09	AVGIGAMI	FP_Thai	103	0.165	0.093	0.174	0.672	>50
CNE56	AVGIGAMI	FP_Thai	>500	>50	>50	>50	>50	>50
BB539.2B13	AVGIGAVF	FP_v1	1.82	0.109	0.169	0.274	0.344	0.663
BG505.W6M.C2	AVGIGAVF	FP_v1	0.211	0.038	0.031	0.037	0.055	0.178
BI369.9A	AVGIGAVF	FP_v1	0.108	0.056	0.02	0.024	0.015	0.081
CAP210.E8	AVGIGAVF	FP_v1	0.218	0.177	0.174	0.206	0.209	0.195
CNE31	AVGIGAVF	FP_v1	0.122	0.057	0.077	0.115	0.154	0.202
6405.v4.c34	AIGLGAMF	FP_v3	>500	>50	8.74	>50	>50	>50
A03349M1.vrc4a	AIGLGAMF	FP_v3	>500	49.6	>50	>50	>50	>50
QH0515.01	AVGTIGAMF	FP_v4	>500	>50	>50	>50	>50	>50
CNE7	AVGTIGAMF	FP_v4	>500	1.12	5.08	>50	>50	>50
T266-60	AVGLGAVF	FP_v5	0.301	0.06	0.067	0.06	0.05	0.223
CNE30	AVGLGAVF	FP_v5	0.191	0.185	0.234	0.344	0.316	0.5
DU156.12	AVGLGAVL	FP_v6	>500	13.6	>50	0.121	0.138	0.161
3415.v1.c1	AIGMGAVF	FP_v7	0.126	0.072	0.081	0.075	0.098	0.157
269-12	AIGMGAVF	FP_v7	72	>50	>50	0.254	0.269	>50
CAP45.G3	AVGIGAVL	FP_v8	0.064	0.191	0.07	0.073	0.071	0.086
0077_V1.C16	AVGIGAMF	FP_v9	1.12	46.2	29.3	0.189	0.333	2.24

Rational combinations of top single mutant variants (See Fig. 2C)

Virus ID	FP Sequence		N123-VRC34.01	VRC34.01-BZ	VRC34.01-BZ (duplicate)	VRC34.01	VRC34.01-Combo1 (HC_E2K_T59F)	VRC34.01-Combo2 (HC_E2K_T59Y)	DFPH-a.01-BZ
C1080.c3	AVGIGAMI	FP_Thai	2.19	2.53	0.461	0.42	0.317	0.568	>50
C2101.c1	AVGIGAMI	FP_Thai	2.05	2.48	1.88	0.408	0.812	1.51	>50
C4118.09	AVGIGAMI	FP_Thai	>50	>50	>50	>100	0.3	0.963	>50
CNE56	AVGIGAMI	FP_Thai	>50	>50	>50	>100	>50	>50	0.36
BB539.2B13	AVGIGAVF	FP_v1	0.462	0.667	1.11	1.24	0.126	0.323	37.5
BG505.W6M.C2	AVGIGAVF	FP_v1	0.207	0.221	0.09	0.095	0.065	0.125	5.2
BI369.9A	AVGIGAVF	FP_v1	0.14	0.164	0.049	0.081	0.045	0.062	2.49
CAP210.E8	AVGIGAVF	FP_v1	0.312	0.468	0.393	0.373	0.079	0.217	13.7
CNE31	AVGIGAVF	FP_v1	0.159	0.197	0.12	0.124	0.082	0.14	38.6
6405.v4.c34	AIGLGAMF	FP_v3	>50	>50	>50	>100	2.38	12.1	>50
A03349M1.vrc4a	AIGLGAMF	FP_v3	>50	>50	>50	>100	15.6	43.6	>50
QH0515.01	AVGTIGAMF	FP_v4	>50	>50	>50	>100	18	>50	>50
CNE7	AVGTIGAMF	FP_v4	>50	>50	>50	>100	0.1	0.579	>50
T266-60	AVGLGAVF	FP_v5	0.225	0.327	0.074	0.066	0.063	0.14	0.77
CNE30	AVGLGAVF	FP_v5	0.326	0.386	0.589	0.594	0.121	0.259	45.2
DU156.12	AVGLGAVL	FP_v6	0.668	0.342	0.269	0.214	0.119	0.217	12.7
3415.v1.c1	AIGMGAVF	FP_v7	0.152	0.228	0.159	0.145	0.074	0.121	41
269-12	AIGMGAVF	FP_v7	>50	5.04	>50	7.21	12.8	>50	30.2
CAP45.G3	AVGIGAVL	FP_v8	0.112	0.169	0.129	0.138	0.078	0.13	35
0077_V1.C16	AVGIGAMF	FP_v9	1.67	1.99	2.21	1.78	1.38	2.36	6

Top multi-mutant variants identified by yeast-library screening (See Fig. 3C)

Virus ID	FP Sequence		N123-VRC34.01	VRC34.01-BZ	VRC34.01	VRC34.01_m1	VRC34.01_m2	VRC34.01_m3	VRC34.01_m4	VRC34.01_m5	VRC34.01_m6	VRC34.01_m7	VRC34.01_m8	VRC34.01_m9	VRC34.01_m10	VRC34.01_m11	VRC34.01_m12
C1080.c3	AVGIGAMI	FP_Thai	0.843	0.461	0.42	0.433	0.311	0.346	0.426	0.241	>50	0.284	0.698	0.275	0.306	0.958	0.494
C2101.c1	AVGIGAMI	FP_Thai	0.732	1.88	0.408	0.535	0.124	0.315	0.218	0.768	>50	0.148	0.247	0.122	0.266	0.308	0.132
C4118.09	AVGIGAMI	FP_Thai	6.83	>50	>100	0.715	0.303	0.347	0.446	0.349	>50	0.408	0.896	0.451	0.528	5.04	2.01
CNE56	AVGIGAMI	FP_Thai	>50	>50	>100	>50	>50	>50	19	>50	>50	>50	>50	>50	>50	>50	>50
BB539.2B13	AVGIGAVF	FP_v1	0.571	1.11	1.24	0.235	0.107	0.142	0.211	0.097	>50	0.191	0.417	0.182	0.19	0.82	0.219
BG505.W6M.C2	AVGIGAVF	FP_v1	0.178	0.09	0.095	0.104	0.064	0.096	0.123	0.072	>50	0.079	0.219	0.072	0.088	0.303	0.102
BI369.9A	AVGIGAVF	FP_v1	0.132	0.049	0.081	0.075	0.028	0.042	0.055	0.04	>50	0.037	0.136	0.042	0.053	0.13	0.09
CAP210.E8	AVGIGAVF	FP_v1	0.33	0.393	0.373	0.227	0.092	0.146	0.197	0.121	>50	0.104	0.338	0.1	0.107	0.384	0.257
CNE31	AVGIGAVF	FP_v1	0.166	0.12	0.124	0.184	0.082	0.105	0.172	0.087	>50	0.113	0.291	0.095	0.105	0.251	0.109
6405.v4.c34	AIGLGAMF	FP_v3	>50	>50	>100	5	2.34	2.4	3.97	1.66	>50	2.25	4.18	2.39	2.5	10.7	4.01
A03349M1.vrc4a	AIGLGAMF	FP_v3	>50	>50	>100	29.7	14.6	29.4	20.6	11.9	>50	10.3	>50	26	28.9	>50	>50
QH0515.01	AVGTIGAMF	FP_v4	>50	>50	>100	>50	>50	>50	>50	>50	>50	18.2	>50	>50	39.9	>50	>50
CNE7	AVGTIGAMF	FP_v4	>50	>50	>100	0.131	0.037	0.091	0.086	0.087	>50	0.062	0.159	0.051	0.059	>50	>50
T266-60	AVGLGAVF	FP_v5	0.215	0.074	0.066	0.104	0.054	0.051	0.101	0.078	>50	0.104	0.316	0.054	0.066	0.204	0.112
CNE30	AVGLGAVF	FP_v5	0.351	0.689	0.594	0.281	0.164	0.175	0.277	0.173	>50	0.159	0.444	0.151	0.219	0.749	0.305
DU156.12	AVGLGAVL	FP_v6	4.85	0.269	0.214	>50	0.832	>50	0.967	>50	>50	0.413	0.446	0.383	0.252	0.401	0.218
3415.v1.c1	AIGMGAVF	FP_v7	0.156	0.159	0.145	0.088	0.054	0.073	0.102	0.06	>50	0.075	0.22	0.059	0.071	0.213	0.111
269-12	AIGMGAVF	FP_v7	2.01	>50	7.21	>50	1.67	4.74	1.09	5.38	>50	1.65	1.73	1.16	1.65	1.78	0.835
CAP45.G3	AVGIGAVL	FP_v8	0.142	0.129	0.138	0.112	0.048	0.055	0.086	0.066	>50	0.096	0.211	0.08	0.08	0.111	0.078
0077_V1.C16	AVGIGAMF	FP_v9	0.735	2.21	1.78	0.739	0.24	0.897	0.214	0.32	>50	0.329	0.53	0.585	0.295	0.339	0.139

continued

Virus ID	VRC34.01_m13	VRC34.01_m14	VRC34.01_m15	VRC34.01_m16	VRC34.01_m17	VRC34.01_m18	VRC34.01_m19	VRC34.01_m20	VRC34.01_m21	VRC34.01_m22	VRC34.01_m23	VRC34.01_m24	VRC34.01_m25	VRC34.01_m26	VRC34.01_m27	**VRC34.01_m28**
C1080.c3	0.14	0.287	0.386	0.251	0.353	0.35	0.305	0.187	0.832	0.153	1.81	0.209	0.146	0.159	0.076	0.094
C2101.c1	0.03	0.082	0.213	0.068	0.069	0.091	0.067	0.034	0.217	0.209	2.45	0.066	0.247	0.184	0.147	0.041
C4118.09	0.461	0.987	0.785	0.552	1.25	1.23	1.64	0.877	3.65	0.279	3.26	1.02	0.206	0.256	0.216	0.165
CNE56	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	26.7	>50	>50	1.09
BB539.2B13	0.053	0.227	0.278	0.126	0.142	0.235	0.339	0.158	0.531	0.094	0.924	0.129	0.076	0.095	0.04	0.034
BG505.W6M.C2	0.033	0.147	0.142	0.078	0.065	0.112	0.079	0.065	0.17	0.048	1.04	0.065	0.042	0.042	0.029	0.025
BI369.9A	0.016	0.03	0.015	0.028	0.04	0.043	0.044	0.04	0.103	0.024	0.809	0.029	0.024	0.029	0.021	0.016
CAP210.E8	0.071	0.196	0.17	0.125	0.147	0.171	0.174	0.124	0.316	0.078	1.9	0.122	0.052	0.063	0.045	0.039
CNE31	0.054	0.119	0.089	0.076	0.074	0.088	0.092	0.049	0.202	0.049	0.636	0.074	0.031	0.05	0.025	0.029
6405.v4.c34	1.48	5.33	9.28	1.93	3.64	3.08	5.02	3.45	9.97	1.37	42.6	4.37	1.41	1.24	0.968	0.842
A03349M1.vrc4a	26.6	>50	>50	34.8	>50	>50	>50	>50	>50	17.2	>50	38.5	14.7	12	13.6	10.6
QH0515.01	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	0.929
CNE7	8.95	2.45	>50	0.582	>50	>50	>50	>50	>50	0.027	5.36	>50	0.03	0.054	0.054	0.031
T266-60	0.028	0.061	0.059	0.064	0.052	0.074	0.074	0.042	0.103	0.048	0.472	0.047	0.038	0.056	0.031	0.037
CNE30	0.106	0.234	0.232	0.275	0.191	0.199	0.216	0.142	0.425	0.108	1.44	0.174	0.099	0.086	0.064	0.078
DU156.12	0.067	0.185	0.164	0.185	0.244	0.274	0.208	0.11	0.555	>50	19.2	0.136	>50	0.279	0.699	0.224
3415.v1.c1	0.027	0.091	0.07	0.058	0.047	0.051	0.057	0.036	0.115	0.034	0.79	0.046	0.032	0.042	0.021	0.024
269-12	0.245	0.435	0.838	0.206	0.588	0.428	0.36	0.19	0.63	>50	>50	0.244	>50	1.07	5.51	0.16
CAP45.G3	0.02	0.044	0.036	0.047	0.045	0.049	0.065	0.026	0.096	0.046	0.868	0.044	0.033	0.047	0.036	0.026
0077_V1.C16	0.04	0.093	0.154	0.048	0.091	0.092	0.119	0.056	0.167	0.409	15.5	0.05	0.113	0.106	0.331	0.033

Rational combinations of top multi-mutant variants. (See Fig. S3B)

Virus ID	FP Sequence		N123-VRC34.01	VRC34.01-BZ	VRC34.01	mm28-VL_D1Y	mm28-VL_D1Y_S92T	mm26-VL_R24A	mm28-VL_R30H_S92T	mm28-VL_S67W	mm26-VL_S92T	mm28-VL_R30A	VRC34-I3 combo1	VRC34.01_m28	VRC34.01_m6
C1080.c3	AVGIGAMI	FP_Thai	5.2	0.461	0.42	0.66	0.405	0.363	0.21	0.839	0.686	0.528	0.632	0.099	0.258
C2101.c1	AVGIGAMI	FP_Thai	13.8	1.88	0.408	0.158	0.137	0.107	0.073	0.244	0.221	0.147	0.3	0.041	0.297
C4118.09	AVGIGAMI	FP_Thai	103	>50	>100	1.14	0.829	0.683	0.499	1.64	1.42	1.1	1.2	0.282	0.623
CNE56	AVGIGAMI	FP_Thai	>500	>50	>100	4.08	0.855	2.32	0.277	2.62	3.44	2.45	23.2	1.49	>50
BB539.2B13	AVGIGAVF	FP_v1	1.82	1.11	1.24	0.217	0.162	0.199	0.083	0.277	0.436	0.397	0.564	0.073	0.135
BG505.W6M.C2	AVGIGAVF	FP_v1	0.211	0.09	0.095	0.09	0.096	0.069	0.042	0.206	0.171	0.132	0.179	0.031	0.069
BI369.9A	AVGIGAVF	FP_v1	0.108	0.049	0.081	0.07	0.052	0.051	0.016	0.099	0.097	0.073	0.094	0.017	0.041
CAP210.E8	AVGIGAVF	FP_v1	0.218	0.393	0.373	0.145	0.169	0.136	0.053	0.262	0.267	0.181	0.177	0.043	0.083
CNE31	AVGIGAVF	FP_v1	0.122	0.12	0.124	0.189	0.167	0.123	0.077	0.28	0.28	0.213	0.309	0.049	0.077
6405.v4.c34	AIGLGAMF	FP_v3	>500	>50	>100	2.01	2.54	2.13	1.43	2.9	2.55	2.35	3.15	1.2	1.73
A03349M1.vrc4a	AIGLGAMF	FP_v3	>500	>50	>100	9.41	7.73	8.41	6.16	3.59	24.3	27.8	37.7	11.3	12.9
QH0515.01	AVGTIGAMF	FP_v4	>500	>50	>100	12.5	3.82	12.3	3.46	36.1	4.54	5.41	>50	2.8	>50
CNE7	AVGTIGAMF	FP_v4	>500	>50	>100	0.122	0.098	0.095	0.053	0.24	0.22	0.185	0.241	0.04	0.115
T266-60	AVGLGAVF	FP_v5	0.301	0.074	0.066	0.152	0.112	0.067	0.052	0.265	0.338	0.203	0.168	0.04	0.081
CNE30	AVGLGAVF	FP_v5	0.191	0.689	0.594	0.624	0.421	0.354	0.198	1	0.71	0.596	0.715	0.126	0.228
DU156.12	AVGLGAVL	FP_v6	>500	0.269	0.214	1.32	0.705	0.902	0.319	1.98	1.38	1.69	>50	0.246	>50
3415.v1.c1	AIGMGAVF	FP_v7	0.126	0.159	0.145	0.088	0.09	0.073	0.057	0.164	0.181	0.151	0.181	0.033	0.046
269-12	AIGMGAVF	FP_v7	72	>50	7.21	2.45	1.54	1.95	0.745	2.63	1.4	1.14	>50	0.277	5.54
CAP45.G3	AVGIGAVL	FP_v8	0.064	0.129	0.138	0.147	0.118	0.083	0.034	0.119	0.334	0.223	0.251	0.047	0.064
0077_V1.C16	AVGIGAMF	FP_v9	1.12	2.21	1.78	0.212	0.169	0.177	0.061	0.316	0.29	0.271	0.293	0.06	0.323

Appendix AE: Table 17. Supplemental Chapter 4.

Table 17. Supplemental Chapter 4. 208-virus panel IC50 (ug/mL) neutralization: template VRC34.01 and VRC34.01_combo1.

High Throughput Antibody Screen - FP mAbs			Potency Reference			IC50 (ug/mL) 208-virus Luc/TZM-bl Neutralization assay (17June2021)					
Assay - Luc/TZM-bl			<0.001	.001-.01	.01-.100						
			.100-1.00	1.00-10.0	>10.0						
			FP Sequence								
Virus ID	Clade	Group	512-512b-513-514-514b-515-516-517-517b-518-519	512-513-514-515-516-517-518-519	FP-shorthand		N123-VRC34.01		VRC34.01-Combo1		
620345.c1	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	1	351	R	0.965	S	
BJOX010000.06.2	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	2	>500	R	0.384	S	
BJOX025000.01.1	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	3	222	R	0.414	S	
C1080.c3	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	4	5.2	S	0.119	S	
C2101.c1	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	5	13.8	S	1.44	S	
C3347.c11	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	6	>500	R	>100	R	
C4118.09	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	7	103	R	0.121	S	
CM244.ec1	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	8	>500	R	>100	R	
CNE5	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	9	55.4	R	0.296	S	
CNE55	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	10	>500	R	>100	R	
CNE56	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	11	>500	R	>100	R	
CNE8	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	12	93.5	R	3.19	S	
R1166.c1	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	13	>500	R	>100	R	
R2184.c4	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	14	>500	R	>100	R	
R3265.c6	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	15	>500	R	1.24	S	
TH023.6	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	16	>500	R	>100	R	
TH966.8	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	17	>500	R	66.2	S	
TH976.17	AE	A	A_VG_IGA_MI	AVGIGAMI	FP_Thai	18	>500	R	0.478	S	
0439.v5.c1	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	1	0.441	S	0.104	S	
3365.v2.c20	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	2	0.056	S	0.042	S	
398-F1_F6_20	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	3	0.16	S	0.086	S	
BB201.B42	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	4	0.064	S	0.02	S	
BB539.2B13	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	5	1.82	S	0.098	S	
BG505.W6M.C2	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	6	0.211	S	0.029	S	
BI369.9A	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	7	0.108	S	0.02	S	
KER2008.12	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	8	0.038	S	0.028	S	
KNH1209.18	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	9	0.371	S	0.124	S	
MB201.A1	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	10	0.066	S	0.021	S	
MB539.2B7	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	11	0.306	S	0.209	S	
MI369.A5	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	12	0.256	S	0.032	S	
Q23.17	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	13	0.178	S	0.081	S	
Q259.17	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	14	3.31	S	0.073	S	
Q442.d12	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	15	0.113	S	0.08	S	
QH209.14M.A2	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	16	0.198	S	0.121	S	
RW020.2	A	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	17	1.29	S	0.098	S	
246-F3.C10.2	AC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	18	0.806	S	0.242	S	
3301.V1.C24	AC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	19	0.119	S	0.026	S	
3589.V1.C4	AC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	20	0.113	S	0.03	S	
0815.V3.C3	ACD	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	21	0.036	S	0.027	S	
Q168.a2	AD	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	22	0.162	S	0.073	S	
3988.25	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	23	0.056	S	0.026	S	
45_01dG5	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	24	0.599	S	0.026	S	
89.6.DG	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	25	0.044	S	0.023	S	
Bal.01	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	26	1.41	S	0.084	S	
BL01.DG	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	27	0.096	S	0.022	S	
CNE57	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	28	0.038	S	0.013	S	
HO86.8	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	29	0.574	S	0.065	S	
[illegible]	B	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	30	[illegible]	S	[illegible]	S	
BJOX002000.03.2	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	31	0.251	S	0.049	S	
CH117.4	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	32	0.151	S	0.011	S	
CH119.10	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	33	0.108	S	0.024	S	
CNE15	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	34	>500	R	0.075	S	
CNE19	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	35	0.049	S	0.007	S	
CNE21	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	36	0.055	S	0.032	S	
CNE40	BC	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	37	0.194	S	0.08	S	
286.36	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	38	>500	R	0.165	S	
0921.V2.C14	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	39	>500	R	0.131	S	
25710-2.43	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	40	26.9	S	0.104	S	
3873.V1.C24	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	41	1.11	S	0.127	S	
426c	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	42	0.102	S	0.037	S	
962M651.02	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	43	0.033	S	0.02	S	
BR025.9	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	44	0.805	S	0.049	S	
CAP210.E8	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	45	0.218	S	0.033	S	
Ce1176.A3	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	46	0.216	S	0.036	S	
CNE31	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	47	0.122	S	0.025	S	
CNE58	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	48	0.082	S	0.068	S	
DU172.17	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	49	0.112	S	0.045	S	
MW965.26	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	50	>500	R	0.075	S	
TV1.29	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	51	0.072	S	0.052	S	
ZA012.29	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	52	1.38	S	0.244	S	
ZM109.4	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	53	1.71	S	1.34	S	
ZM135.10a	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	54	298	R	69.5	R	
ZM176.66	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	55	0.88	S	0.159	S	
ZM214.15	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	56	1.15	S	0.074	S	
ZM233.6	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	57	0.058	S	0.015	S	
ZM55.28a	C	T	A_VG_IGA_VF	AVGIGAVF	FP_v1	58	0.232	S	0.068	S	
6095.V1.C10	ACD	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	1	>500	R	75.1	R	
3337.V2.C6	CD	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	2	>500	R	0.712	S	
231965.c1	D	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	3	>500	R	>100	R	
3016.v5.c45	D	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	4	>500	R	>100	R	
6405.v4.c34	D	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	5	>500	R	4.22	S	
A03349M1.vrc4a	D	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	6	>500	R	9.73	S	
A07412M1.vrc12	D	A	A_IG_LGA_MF	AIGLGAMF	FP_v3	7	>500	R	>100	R	
7165.18	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	1	>500	R	>100	R	
ADA.DG	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	2	>500	R	>100	R	
CNE10	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	3	>500	R	>100	R	
CNE12	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	4	>500	R	>100	R	
CNE14	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	5	>500	R	>100	R	
QH0515.01	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	6	>500	R	>100	R	
QH0692.42	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	7	>500	R	>100	R	
SC422.8	B	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	8	>500	R	>100	R	
CNE7	BC	A	A_VGTIGA_MF	AVGTIGAMF	FP_v4	9	>500	R	0.105	S	
3718.v3.c11	A	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	1	2.37	S	0.379	S	
KER2018.11	A	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	2	0.221	S	0.08	S	
6540.v4.c1	AC	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	3	0.305	S	0.074	S	
6545.V4.C1	AC	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	4	0.156	S	0.058	S	
3468.V1.C12	AD	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	5	0.218	S	0.11	S	
242-14	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	6	0.118	S	0.091	S	
DJ263.8	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	7	0.101	S	0.119	S	
T250-4	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	8	0.125	S	0.033	S	
T253-11	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	9	0.395	S	0.257	S	
T255-34	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	10	0.051	S	0.022	S	
T266-60	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	11	0.301	S	0.09	S	
T278-50	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	12	0.023	S	0.169	S	
T280-5	AG	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	13	0.098	S	0.108	S	
X2278.C2.B6	B	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	14	11.9	S	0.785	S	
CH070.1	BC	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	15	269	R	86.1	R	
001428-2.42	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	16	0.07	S	0.035	S	
16936-2.21	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	17	0.085	S	0.03	S	
25925-2.22	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	18	0.145	S	0.036	S	
3168.V4.C10	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	19	2.06	S	0.533	S	
6644.V2.C33	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	20	0.044	S	0.022	S	
CNE30	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	21	0.191	S	0.061	S	
CNE53	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	22	0.116	S	0.017	S	
DU151.02	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	23	6.68	S	0.104	S	
DU422.01	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	24	4.23	S	0.083	S	
ZM215.8	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	25	4.27	S	0.517	S	
ZM249.1	C	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	26	0.067	S	0.026	S	
P1981.C5.3	G	T	A_VG_LGA_VF	AVGLGAVF	FP_v5	27	1.35	S	0.096	S	
25711-2.4	C	A	A_VG_LGA_VL	AVGLGAVL	FP_v6	1	50.7	R	0.145	S	
DU156.12	C	A	A_VG_LGA_VL	AVGLGAVL	FP_v6	2	>500	R	0.045	S	
3415.v1.c1	A	A	A_IG_MGA_VF	AIGMGAVF	FP_v7	1	0.126	S	0.014	S	
269-12	AG	A	A_IG_MGA_VF	AIGMGAVF	FP_v7	2	72	R	>100	R	
P0402.c2.11	G	A	A_IG_MGA_VF	AIGMGAVF	FP_v7	3	>500	R	2.36	S	
BaL.26	B	A	A_VG_IGA_VL	AVGIGAVL	FP_v8	1	353	R	1.32	S	
6471.V1.C16	C	A	A_VG_IGA_VL	AVGIGAVL	FP_v8	2	>500	R	>100	R	
CAP45.G3	C	A	A_VG_IGA_VL	AVGIGAVL	FP_v8	3	0.064	S	0.041	S	
DU123.06	C	A	A_VG_IGA_VL	AVGIGAVL	FP_v8	4	27.3	S	0.044	S	
ZM106.9	C	A	A_VG_IGA_VL	AVGIGAVL	FP_v8	5	17.8	S	0.079	S	
6101.10	B	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	1	0.316	S	0.115	S	
BX08.16	B	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	2	1.57	S	0.424	S	
RHPA.7	B	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	3	1.45	S	0.185	S	
0077_V1_C16	C	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	4	1.12	S	0.497	S	
CAP244.D3	C	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	5	>500	R	42.1	S	
SO18.18	C	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	6	0.262	S	0.021	S	
TZBD.02	C	A	A_VG_IGA_MF	AVGIGAMF	FP_v9	7	2.05	S	0.441	S	
0260.v5.c36	A	A	A_VG_MAA_LF	AVGMAALF	#N/A	1	>500	R	>100	R	
0330.v4.c3	A	A	A_VG_MGA_VF	AVGMGAVF	#N/A	2	0.052	S	0.026	S	
BS208.B1d	A	A	A_VT_MGA_VF	AVTMGAVF	#N/A		#N/A	#N/A	#N/A	#N/A	
MS208.A1d	A	A	A_VT_MGA_VF	AVTMGAVF	#N/A		#N/A	#N/A	#N/A	#N/A	
Q769.d22	A	U	A_VG_FGA_FF	AVGFGAFF	#N/A	3	>500	R	>100	R	
Q769.h5	A	U	A_VG_FGA_FF	AVGFGAFF	#N/A	4	>500	R	>100	R	
UG037.8	A	U	AVFE_LGA_VF	AVVELGAVF	#N/A	5	>500	R	>100	R	
Q461.e2	AD	A	A_VG_MAA_VF	AVGMAAVF	#N/A	6	>500	R	>100	R	
BJOX009000.02.4	AE	A	A_VT_IGA_MI	AVTIGAMI	#N/A	7	>500	R	0.747	S	
BJOX028000.10.3	AE	A	A_VG_IGT_MI	AVGIGTMI	#N/A	8	134	R	0.09	S	
CNE3	AE	A	A_VG_IGT_MI	AVGIGTMI	#N/A	9	15.7	S	0.325	S	
CNE59	AE	A	A_VG_IGA_LL	AVGIGALL	#N/A	10	>500	R	>100	R	
M02138	AE	A	A_VG_IGI_MI	AVGIGIMI	#N/A	11	>500	R	>100	R	
235-47	AG	U	A__A_LGA_VF	A_ALGAVF	#N/A	12	>500	R	>100	R	
263-8	AG	A	A_VG_LGA_LF	AVGLGALF	#N/A	13	>500	R	>100	R	
271-11	AG	U	AVVG_LGA_VF	AVVGLGAVF	#N/A	14	>500	R	>100	R	
928-28	AG	A	A_VG_MGA_LI	AVGMGALI	#N/A	15	>500	R	>100	R	
T251-18	AG	U	A_VG_LGA_FF	AVGLGAFF	#N/A	16	>500	R	>100	R	
T257-31	AG	A	A_VG_MGA_LF	AVGMGALF	#N/A	17	313	R	>100	R	
T33-7	AG	U	AVVG_LGA_VF	AVVGLGAVF	#N/A	18	>500	R	>100	R	
5768.04	B	A	T_VG_IGA_LF	TVGIGALF	#N/A	19	>500	R	>100	R	
6535.3	B	A	A_VGMLGA_MF	AVGMLGAMF	#N/A	20	>500	R	>100	R	
AC10.29	B	A	A_VG_IGA_LF	AVGIGALF	#N/A	21	>500	R	1	S	
BG1168.01	B	A	A_VG_LGA_MF	AVGLGAMF	#N/A	22	4.83	S	0.361	S	
BR07.DG	B	A	A_VG_IGA_LF	AVGIGALF	#N/A	23	>500	R	0.053	S	
CAAN.A2	B	A	A_VG_LGA_MF	AVGLGAMF	#N/A	24	5.5	S	0.169	S	
CNE4	B	A	A_MG_IGA_MF	AMGIGAMF	#N/A	25	0.344	S	0.056	S	
HT593.1	B	A	A_VGIVGA_MF	AVGIVGAMF	#N/A	26	>500	R	>100	R	
HXB2.DG	B	A	A_VG_IGA_LF	AVGIGALF	#N/A	27	0.15	S	0.023	S	
JRCSF.JB	B	A	A_VG_IGA_LF	AVGIGALF	#N/A	28	>500	R	0.633	S	
MN.3	B	A	A__A_IGA_LF	A_AIGALF	#N/A	29	>500	R	>100	R	
PVO.04	B	A	A_VGTLGA_MF	AVGTLGAMF	#N/A	30	>500	R	>100	R	
REJO.67	B	A	A_VG_IGA_LF	AVGIGALF	#N/A	31	325	R	>100	R	
SF162.LS	B	A	A_VT_LGA_MF	AVTLGAMF	#N/A	32	0.183	S	0.053	S	
SS1196.01	B	A	A__A_LGA_LF	A_ALGALF	#N/A	33	>500	R	>100	R	
THRO.18	B	A	D_LG_LGA_LF	DLGLGALF	#N/A	34	>500	R	>100	R	
TRJO.58	B	A	A_IG_IGA_VF	AIGIGAVF	#N/A	35	7.46	S	1.32	S	
TRO.11	B	A	A_VGTLGA_MF	AVGTLGAMF	#N/A	36	>500	R	>100	R	
WITO.33	B	T	A_VT_LGA_VF	AVTLGAVF	#N/A	37	0.26	S	0.067	S	
YU2.DG	B	A	A_VG_LGA_LF	AVGLGALF	#N/A	38	102	R	44.4	S	
CH038.12	BC	U	A__A_LGA_VF	A_ALGAVF	#N/A	39	>500	R	>100	R	
CH181.12	BC	U	A_VGGIGA_VF	AVGGIGAVF	#N/A	40	134	R	>100	R	
CNE20	BC	U	T_VG_IGA_VF	TVGIGAVF	#N/A	41	372	R	2.47	S	
288.38	C	A	A__A_LGA_LF	A_ALGALF	#N/A	42	>500	R	>100	R	
0013095-2.11	C	A	A_VG_IGA_VI	AVGIGAVI	#N/A	43	485	R	0.128	S	
00836-2.5	C	A	A_VG_VGAVMF	AVGVGAVMF	#N/A	44	>500	R	>100	R	
16055-2.3	C	A	A_VG_IGA_VI	AVGIGAVI	#N/A	45	>500	R	0.148	S	
16845-2.22	C	A	A_VG_LGA_MI	AVGLGAMI	#N/A	46	>500	R	>100	R	
26191-2.48	C	U	A__A_LGA_VF	A_ALGAVF	#N/A	47	>500	R	>100	R	
3637.V5.C3	C	A	A_VG_IGA_VI	AVGIGAVI	#N/A	48	>500	R	>100	R	
6322.V4.C1	C	U	A__A_LGA_VF	A_ALGAVF	#N/A	49	>500	R	>100	R	
6631.V3.C10	C	A	A_VG_LGA_LF	AVGLGALF	#N/A	50	>500	R	>100	R	
6785.V5.C14	C	A	A_VGGFGA_MI	AVGGFGAMI	#N/A	51	>500	R	>100	R	
6838.V1.C35	C	A	A_VG_MGA_VF	AVGMGAVF	#N/A	52	0.052	S	0.037	S	
CAP256.206.C9	C	U	AVVG_LGA_VF	AVVGLGAVF	#N/A	53	>500	R	>100	R	
CE703010217.B6	C	A	A_VG_MGA_VF	AVGMGAVF	#N/A	54	>500	R	0.352	S	
TZA125.17	C	A	A_VG_IGA_LF	AVGIGALF	#N/A	55	>500	R	>100	R	
ZM197.7	C	A	A_AG_IGA_VF	AAGIGAVF	#N/A	56	>500	R	>100	R	
ZM53.12	C	A	A_VG_LGA_MF	AVGLGAMF	#N/A	57	>500	R	0.03	S	
3326.V4.C3	CD	A	A_IG_LGA_VF	AIGLGAVF	#N/A	58	31.2	S	8.96	S	
3817.v2.c59	CD	A	A_IG_MGA_LF	AIGMGALF	#N/A	59	>500	R	>100	R	
191821.E8.1	D	A	A_IG_LGA_VF	AIGLGAVF	#N/A	60	16.6	S	1.48	S	
247-23	D	A	A_IG_LGA_LF	AIGLGALF	#N/A	61	>500	R	>100	R	
57128.vrc15	D	A	A_IG_LGA_LF	AIGLGALF	#N/A	62	>500	R	>100	R	
NKU3006.ec1	D	A	A_IG_LGA_LF	AIGLGALF	#N/A	63	>500	R	>100	R	
UG021.16	D	A	A_IG_LGA_LF	AIGLGALF	#N/A	64	>500	R	1.19	S	
UG024.2	D	A	A_IG_LGA_VF	AIGLGAVF	#N/A	65	>500	R	0.439	S	
X1193.c1	G	A	A_VG_MGA_VF	AVGMGAVF	#N/A	66	2.61	S	39.4	S	
X1254.c3	G	U	A_VG_LGA_FF	AVGLGAFF	#N/A	67	>500	R	7.67	S	
X1632.S2.B10	G	A	A_IG_LGT_VF	AIGLGTVF	#N/A	68	2.88	S	0.107	S	
X2088.c9	G	A	E_VT_LGA_LF	EVTLGALF	#N/A	69	>500	R	>100	R	
X2131.C1.B5	G	A	A_VG_MGA_VF	AVGMGAVF	#N/A	70	>500	R	>100	R	
SIVmac251.30.S53	NA	#N/A	#N/A	#N/A	#N/A	71	>500	R	>100	R	
SVA.MLV	NA	#N/A	#N/A	#N/A	#N/A	72	>500	R	>100	R	

	N123-VRC34.01	VRC34.01-Combo1
# Viruses	208	208
Total V5 Neutralized	122	
IC50 <500ug/ml	109	144
IC50 <100ug/ml	105	141
IC50 <50ug/ml	97	137
IC50 <10ug/ml	73	122
IC50 <1.0ug/ml	25	76
IC50 <0.1ug/ml	0	1
% V5 Neutralized		
IC50 <500ug/ml	59	
IC50 <100ug/ml	52	69
IC50 <50ug/ml	50	68
IC50 <10ug/ml	47	66
IC50 <1.0ug/ml	35	59
IC50 <0.1ug/ml	12	37
IC50 <0.01ug/ml	0	0
Median IC50	0.303	0.091
Geometric Mean	0.898	0.156